Multiparametric MRI Radiomics and Machine Learning Framework for Predicting Treatment Response in Glioblastoma

Suchibrata Patra[1]
St. Xavier's College (Autonomous), Kolkata
This study was conducted and completed at TCG CREST during the author's academic affiliation with St. Xavier's College (Autonomous), Kolkata.

Abstract

Distinguishing True Progression (TP) from Pseudo-Progression (PsP) after chemoradiotherapy remains a major diagnostic challenge in GBM, as both entities present near-identical appearances on conventional contrast-enhanced post-treatment MRI. This distinction carries substantial clinical weight, since TP and PsP demand divergent management yet cannot be reliably separated on routine imaging alone. We investigated whether radiomic features derived from a parsimonious, voxel-wise pharmacokinetic model of dynamic contrast-enhanced (DCE) MRI, combined with MGMT status, could discriminate between the two. The cohort comprised 82 adults with IDH-wildtype GBM who developed a new contrast-enhancing lesion within six months of chemoradiotherapy; classification (53 TP, 29 PsP) was established by histopathology where available (n=52) and modified RANO criteria otherwise (n=30). At every voxel, contrast-concentration time courses were fitted to five candidate pharmacokinetic models, and the best fit was retained by AIC minimisation, yielding parsimonious Ktrans, Ve, Vp, and taui maps adapting to local heterogeneity rather than a single fixed model across the tumour. Following segmentation, 1,073 radiomic descriptors were extracted and reduced via Mann–Whitney U filtering and Elastic Net, then used to train five classifiers across four feature configurations. A Random Forest classifier combining parsimonious DCE-MRI radiomics with MGMT status achieved the best discrimination (mean AUC 0.89, sensitivity 0.93, specificity 0.76, F1 0.90), outperforming features without MGMT (0.84), a T1-post-contrast baseline (0.72), and a single-model extended-Tofts analysis (0.68). Shape and textural descriptors of the Ktrans map, with tumour volume, were the strongest predictors, MGMT contributing a smaller, independent effect. Allowing the model to vary voxel-wise improves non-invasive discrimination.

Introduction

Glioblastoma (GBM) is the most common and lethal primary malignant brain tumour in adults, classified by the 2021 WHO Classification of Central Nervous System Tumours as a Grade IV, IDH-wildtype, adult-type diffuse astrocytic neoplasm.[1, 2] Despite maximal safe resection followed by concurrent chemoradiotherapy and adjuvant temozolomide, median survival remains 14–16 months, and five-year survival is below 6%.[1] A pervasive and clinically consequential problem in the months following treatment is that the large majority of patients develop a new or enlarging contrast-enhancing lesion on surveillance MRI, and conventional imaging cannot reliably determine whether this reflects genuine, biologically active tumour regrowth (TP) or a transient, treatment-induced inflammatory and vascular phenomenon that will stabilise or regress without any change in management (PsP).[3, 4] Both entities can be radiologically indistinguishable on structural post-contrast imaging, yet the clinically appropriate response is opposite: TP warrants escalation to second-line therapy, re-resection, or trial enrolment, whereas PsP warrants continued observation on the existing regimen. Misclassification in either direction carries direct harm – unnecessary morbidity and cost from premature treatment escalation in patients with PsP, or delayed intervention and accelerated neurological decline in patients whose TP is mistaken for a treatment effect.[5]

The clinical implications of diagnostic misclassification in this setting warrant emphasis. False-positive discrimination of progression in patients with PsP introduces unnecessary morbidity from escalated chemotherapy, with attendant psychological burden and substantial healthcare costs. Conversely, false-negative misclassification of genuine progression results in delayed intervention and accelerated neurological decline. Current

imaging-based approaches achieve sensitivities approaching 90% but specificities not exceeding 76%, creating diagnostic uncertainty at a critical juncture. An imaging biomarker achieving Area Under the Receiver-Operating-Characteristic Curve $\geq$0.85 with balanced sensitivity and specificity could meaningfully inform treatment-escalation decisions without recourse to re-biopsy.

Dynamic contrast-enhanced MRI (DCE-MRI) offers a physiologically grounded alternative to structural imaging by tracking the temporal kinetics of a gadolinium-based contrast agent as it moves between the intravascular and extravascular-extracellular compartments. Pharmacokinetic (PK) modelling of the resulting signal-intensity time courses yields quantitative biomarkers – vascular permeability (K^{trans}), extravascular-extracellular volume fraction (V_e), plasma volume fraction (V_p), and, in more physiologically complete models, intracellular water exchange time (τ_i) – that plausibly differ between the disorganised, hyperpermeable neovasculature of actively growing tumour and the comparatively diffuse vascular injury characteristic of radiation- and chemotherapy-induced inflammation.[6, 7] Radiomic analysis extends this physiological characterisation by extracting high-dimensional quantitative descriptors of intensity, shape and texture from these parametric maps, descriptors that are inaccessible to visual inspection and that can, in principle, be combined with molecular covariates such as MGMT promoter methylation status – itself an established, independent predictor of response to alkylating chemotherapy and a biomarker already obtained as part of routine GBM management.[2]

A methodological limitation runs through nearly all published DCE-MRI radiomics studies in this area: a single PK model, almost always the extended Tofts model, is fitted uniformly across the entire tumour volume. GBM is among the most vascularly and cellularly heterogeneous of human solid tumours, and there is no biological reason to expect that a single kinetic model, calibrated under a fixed set of exchange assumptions, adequately describes contrast dynamics in every voxel of every tumour. Where the model is misspecified for a given tissue microenvironment, the resulting parameter estimates – and any radiomic features derived from them – are correspondingly biased, which may explain some of the inconsistency in specificity reported across the existing literature.

We addressed this limitation directly. Rather than fixing the PK model a priori, we fitted five candidate models of increasing physiological complexity at every image voxel and retained, on a voxel-by-voxel basis, whichever model minimised the AIC – a principled, statistically grounded criterion that rewards goodness of fit while penalising unnecessary parametric complexity. We then asked whether the resulting parsimonious, biologically adaptive parametric maps, combined with radiomic feature extraction, systematic dimensionality reduction, and MGMT methylation status, could improve machine-learning-based discrimination of TP from PsP relative to a conventional single-model approach and relative to structural T1-post-contrast radiomics alone.

Methods

Study design and participants

This was a retrospective diagnostic accuracy study. Imaging and clinical data were collected privately for this study. Eligible patients had (i) histologically confirmed GBM, (ii) completion of standard-of-care therapy comprising maximal safe resection followed by concurrent chemoradiotherapy, (iii) development of a new contrast-enhancing lesion within the radiation field on follow-up MRI within six months of completing therapy, and (iv) availability of both anatomical and DCE-MRI sequences together with MGMT methylation status. Eighty-two patients met these criteria (49 men, 33 women; mean age 58.4 years, SD 8.9). The reference-standard diagnosis was established histopathologically from repeat surgical resection or stereotactic biopsy tissue where available (52 patients; TP defined as >25% malignant features on histology, $n = 31$; PsP as <25% malignant features, $n = 21$), and by modified Response Assessment in Neuro-Oncology (mRANO) criteria, based on serial imaging follow-up, for the remaining 30 patients without accessible tissue (TP $n = 22$; PsP $n = 8$). The final cohort comprised 53 patients with TP and 29 with PsP (Figure 1). MGMT promoter methylation status was determined for every patient and categorised as unmethylated ($n = 41$), methylated ($n = 23$), or indeterminate ($n = 18$).

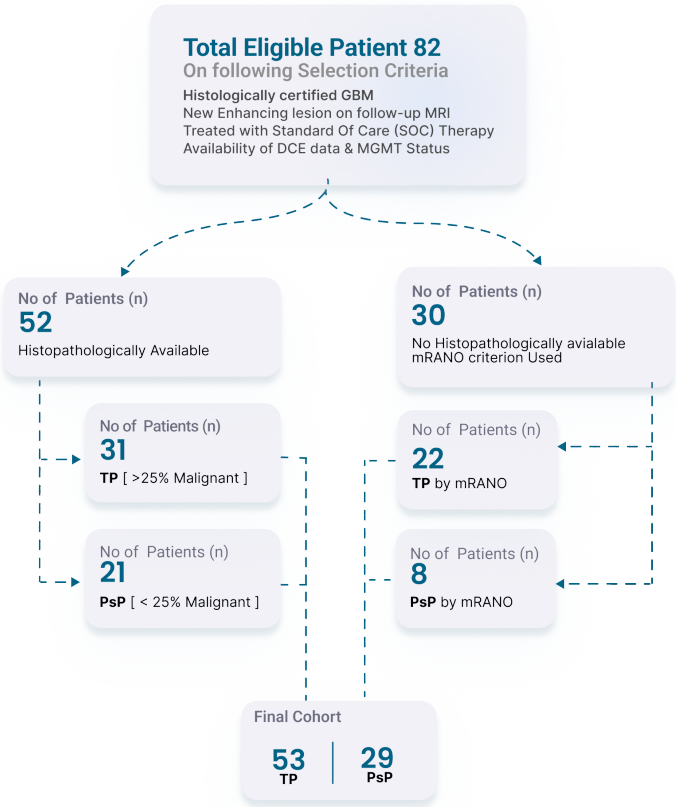

Figure 1: Patient selection flow diagram. Of 82 eligible patients meeting inclusion criteria, 52 had histopathological confirmation from repeat surgery or biopsy tissue and 30 were classified using modified RANO (mRANO) imaging criteria in the absence of accessible tissue. The final analytic cohort comprised 53 patients with TP (TP) and 29 with PsP (PsP).

Image acquisition

All patients underwent MRI on a single Siemens Healthineers (Erlangen, Germany) 3-Tesla MAGNETOM Trio whole-body scanner with a 12-channel phased-array head coil. The neuroimaging protocol comprised T1-weighted magnetisation-prepared rapid acquisition of gradient echo (T1-MPRAGE: repetition time 1820 ms, echo time 3.88 ms, flip angle 15°, slice thickness 1 mm, bandwidth 190 Hz/pixel, field of view 256×256 mm^2 with 75% phase field of view, 192 slices, one average), axial T2-fluid-attenuated inversion recovery (T2-FLAIR: repetition time 10,030 ms, echo time 137 ms, flip angle 150°, slice thickness 3 mm, bandwidth 227 Hz/pixel, field of view 240×240 mm^2 with 75% phase field of view, 60 slices, one average), and a dynamic contrast-enhanced gradient-recalled-echo sequence (repetition time 3.23 ms, echo time 1.12 ms, flip angle 23°, slice thickness 3.5 mm, field of view 220×220 mm^2, in-plane acquisition matrix 256×256, 28 slices, one average, pixel bandwidth 210 Hz/pixel, acquisition time 3 min 33 s, temporal resolution 7 seconds per frame, 30 sequential dynamic measurements) acquired following a weight-based (0.07 mmol/kg) intravenous bolus of gadobenate dimeglumine (MultiHance, Bracco Imaging, Milan, Italy) delivered by power injector (Medrad, Indianola, PA, USA).

Pharmacokinetic modelling and parsimonious parametric map construction

Signal-intensity time courses were converted to contrast-concentration time courses using fixed longitudinal relaxation times (1400 ms tissue, 1600 ms blood) and a contrast-agent relaxivity of 4.3 $\mathrm{L\,mmol^{-1}s^{-1}}$. A vascular input function was defined for each patient from a manually placed region of interest over the superior sagittal sinus. At every voxel, the concentration-time curve was fitted independently to five PK models spanning a hierarchy of physiological assumptions: the standard (non-linear) Tofts model, the extended Tofts model (ETM), the shutter-speed model (SSM, which relaxes the fast-water-exchange assumption), a two-compartment exchange model (2CXM), and a three-site two-exchange model (3S2X). Model selection at each voxel was governed by the AIC,

$$\mathrm{AIC} = N \ln\left(\sum_{i=1}^{N} \left(y_i - C_t(i)\right)^2\right) + \frac{2p(p+1)}{N-p-1}, \tag{1}$$

where N is the number of dynamic time points, y_i the observed signal, $C_t(i)$ the model-fitted concentration at time i, and p the number of free parameters in the candidate model; the model minimising AIC at each voxel was retained. Because K^{trans} and V_e are common to all five models, the parsimonious map for each was constructed by aggregating the voxel-wise best-fit values across all candidates. V_p, present only in ETM, 2CXM and 3S2X, and τ_i, present only in SSM and 3S2X, underwent an analogous but restricted voxel-wise selection among their respective eligible models. Figure 3 illustrates representative model fits to a single-voxel tissue concentration curve, and Figure 4 shows the resulting parsimonious parametric maps for representative TP and PsP cases.

Tumour segmentation and radiomic feature extraction

Tumours were segmented into contrast-enhancing, non-enhancing/necrotic, and oedematous subregions using an automated deep-learning segmentation network (nnU-Net) trained and validated on the BraTS (2018–2021) public benchmark, achieving a test Dice score of 88.95%, with segmentations reviewed for anatomical plausibility. Post-contrast T1 and T2-FLAIR volumes were co-registered to the DCE-MRI parametric maps by rigid-body affine transformation. Total tumour volume was computed by summing enhancing and non-enhancing voxels across all co-registered slices. From the enhancing subregion of each parsimonious parametric map (K^{trans}, V_e, V_p, τ_i), first-order intensity statistics, three-dimensional shape descriptors, and five texture-matrix families (Gray Level Co-occurrence Matrix, Gray Level Run Length Matrix, Gray Level Size Zone Matrix, Neighbouring Gray Tone Difference Matrix, Gray Level Dependence Matrix) were extracted using PyRadiomics, yielding an initial feature space of 1,073 variables.

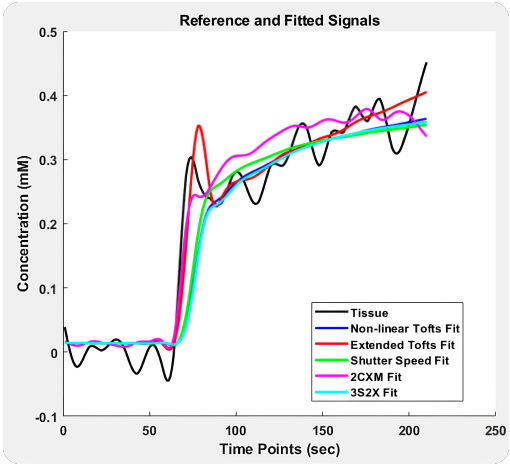

Figure 2: Representative voxel-level tissue concentration-time curve (black) with superimposed fits from all five candidate pharmacokinetic models. The extended Tofts and two-compartment exchange models track the tissue signal most closely during the rapid contrast-uptake phase, whereas the shutter-speed and three-site two-exchange models yield smoother but slightly attenuated peak-enhancement estimates. Model discrepancy is greatest in the early dynamic phase and converges at later time points, motivating the voxel-wise, best-fit (parsimonious) selection strategy used throughout this study rather than a single fixed model.

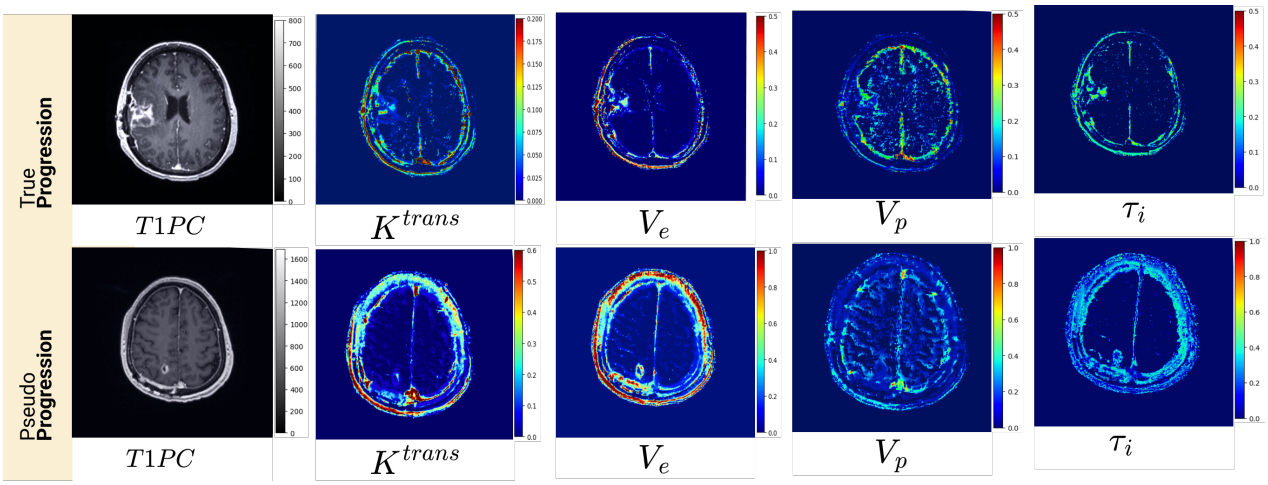

Figure 3: Representative post-contrast T1-weighted (T1PC) images and parsimonious pharmacokinetic parameter maps (K^{trans}, V_e, V_p, τ_i) in a patient with TP (top row) and a patient with PsP (bottom row). The true-progression case shows more extensive, spatially confluent regions of elevated K^{trans} and V_e, consistent with disorganised, hyperpermeable neovasculature, whereas the PsP case shows a more diffuse and rim-distributed permeability pattern consistent with treatment-related vascular injury.

Feature selection

Dimensionality reduction proceeded in two stages. First, each of the 1,073 candidate features was assessed independently for its ability to discriminate TP from PsP using the non-parametric Mann–Whitney U test, on the grounds that radiomic feature distributions routinely violate normality assumptions; features with $p < 0.05$ were retained (reducing the set to 86 features). Second, Elastic Net regression – a regularised linear model combining L_1 and L_2 penalties, which tends to retain or discard correlated features jointly rather than arbitrarily selecting one from a correlated cluster – was applied, with the regularisation strength and L_1/L_2 mixing parameter optimised by five-fold stratified cross-validated grid search, to obtain the final, compact feature subset used for model development.

Class balancing, model development and validation

Because PsP cases were less frequent than TP, reflecting the underlying clinical prevalence, Adaptive Synthetic Sampling (ADASYN) was applied within the training partition only, concentrating synthetic minority-class generation near the decision boundary; the held-out test partition was never resampled. Five candidate classifiers – Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), K-Nearest Neighbours (KNN), and XGBoost – were trained within an unstratified five-fold cross-validation framework, with hyperparameters selected to maximise mean validation AUC across folds. Following hyperparameter selection, classification thresholds were tuned on out-of-fold predictions across the full training set by maximising the F1 score over the interval [0.2, 0.6], avoiding any threshold optimisation on test data. The final models were trained on the complete resampled training set under the selected hyperparameter configuration and evaluated once on the untouched test partition. To assess the independent and combined contributions of parsimonious pharmacokinetic modelling and molecular data, four progressively enriched configurations were compared: (i) conventional extended-Tofts-derived radiomic features with MGMT status; (ii) T1-post-contrast structural radiomic features with MGMT status; (iii) parsimonious DCE-MRI radiomic features without MGMT status; and (iv) parsimonious DCE-MRI radiomic features with MGMT status. All experiments were repeated across multiple independent random seeds governing data splitting, resampling, fold assignment and model initialisation, with performance reported as mean (SD).

Outcome measures and statistical analysis

The primary outcome was test-set AUC; secondary outcomes were sensitivity, specificity, precision and F1 score, all derived from the confusion matrix at the tuned classification threshold. Feature-level contributions to the optimal classifier were quantified using Gini importance. All feature-selection and modelling procedures were implemented in Python using scikit-learn, imbalanced-learn and XGBoost; pharmacokinetic model fitting was implemented in MATLAB; radiomic extraction used PyRadiomics.

Results

Cohort characteristics

The final cohort comprised 82 patients (53 TP, 29 PsP; 49 men, 33 women; mean age 58.4 years, SD 8.9). MGMT promoter status was unmethylated in 41 patients (TP=27, PsP=14), methylated in 23 (TP=12, PsP=11), and indeterminate in 18 (TP=14, PsP=4).

Contribution of parsimonious pharmacokinetic modelling

Across all five candidate PK models, the voxel-wise parsimonious selection consistently achieved lower mean AIC and higher mean R^2 than any single fixed model, both within the enhancing and non-enhancing tumour subregions, confirming that no single kinetic model was uniformly optimal across the heterogeneous vascular microenvironment of GBM. Using the conventional, single-model extended-Tofts approach, the best-performing classifier (Random Forest) achieved a test AUC of only 0.68 (SD 0.13), with wide instability in specificity across repeated runs (0.38, SD 0.12), a pattern consistent with systematic model misspecification degrading downstream radiomic feature quality. Replacing the single-model approach with parsimonious, voxel-wise best-fit parametric maps – without any change to the molecular covariates – raised Random Forest

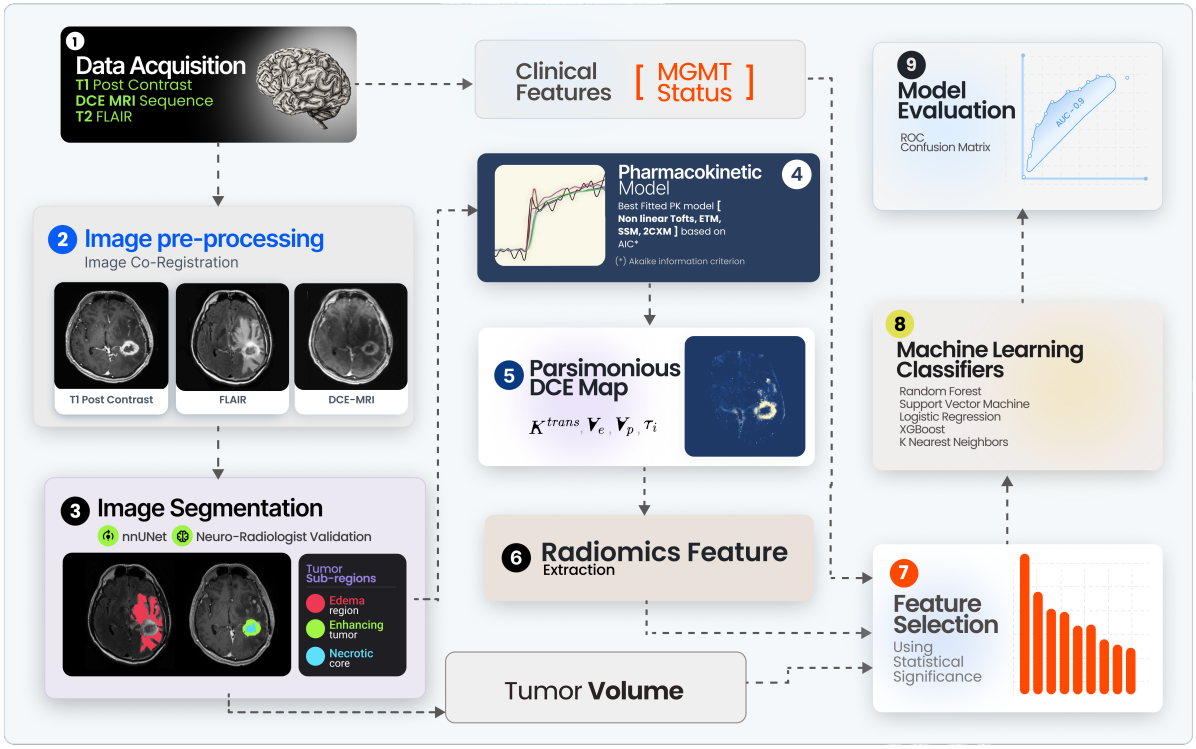

Figure 4: Overview of the end-to-end analytical pipeline: (1) multiparametric MRI and clinical/MGMT data acquisition; (2) image pre-processing and co-registration; (3) automated tumour segmentation with neuroradiologist validation; (4) voxel-wise pharmacokinetic model fitting with AIC-based best-fit selection; (5) construction of parsimonious K^{trans}, V_e, V_p, τ_i parametric maps; (6) radiomic feature extraction; (7) statistically driven feature selection; (8) training of five candidate machine-learning classifiers; and (9) model evaluation by receiver operating characteristic and confusion-matrix analysis.

test AUC to 0.84 (SD 0.09) and sensitivity to 0.93 (SD 0.09), a sixteen-percentage-point absolute improvement in discrimination attributable entirely to more physiologically faithful feature extraction.

Contribution of MGMT methylation status

Incorporating MGMT promoter methylation status alongside the parsimonious DCE-MRI radiomic feature set produced the best-performing model in the entire study. The Random Forest classifier achieved a test AUC of 0.89 (SD 0.07), sensitivity 0.93 (SD 0.09), specificity 0.76 (SD 0.19), precision 0.88 (SD 0.10), and F1 score 0.90 (SD 0.07). The improvement in specificity over the MGMT-free parsimonious configuration – from 0.57 (SD 0.31) to 0.76 (SD 0.19) – was the single largest incremental gain observed across all four configurations, indicating that molecular status substantially reduced false-positive calls of TP, the misclassification with the greatest potential for unwarranted treatment escalation. Table 1 summarises test-set performance of the Random Forest classifier across all four configurations; across every configuration, Random Forest outperformed SVM, LR, KNN and XGBoost.

Table 1: Test-set performance of all five classifiers across the four modelling configurations (mean, SD in parentheses). RF, Random Forest; SVM, Support Vector Machine; LR, Logistic Regression; KNN, K-Nearest Neighbours; XGB, XGBoost.

Configuration	Classifier	AUC	Sensitivity	Specificity	Precision	F1 score
ETM + MGMT	KNN	0.63 (0.07)	0.62 (0.16)	0.60 (0.17)	0.67 (0.13)	0.64 (0.13)
	LR	0.62 (0.13)	0.68 (0.17)	0.43 (0.28)	0.64 (0.10)	0.64 (0.08)
	RF	0.68 (0.13)	0.82 (0.19)	0.38 (0.12)	0.63 (0.13)	0.70 (0.13)
	SVM	0.62 (0.08)	0.74 (0.24)	0.45 (0.32)	0.69 (0.16)	0.67 (0.12)
	XGB	0.61 (0.15)	0.71 (0.17)	0.35 (0.16)	0.59 (0.12)	0.64 (0.13)
T1PC + MGMT	KNN	0.75 (0.06)	0.76 (0.16)	0.56 (0.23)	0.74 (0.11)	0.73 (0.07)
	LR	0.75 (0.07)	0.71 (0.11)	0.62 (0.25)	0.77 (0.13)	0.72 (0.04)
	RF	0.72 (0.05)	0.79 (0.17)	0.51 (0.23)	0.73 (0.11)	0.74 (0.06)
	SVM	0.72 (0.13)	0.72 (0.17)	0.66 (0.20)	0.78 (0.11)	0.74 (0.08)
	XGB	0.72 (0.09)	0.73 (0.11)	0.58 (0.15)	0.74 (0.08)	0.73 (0.07)
Parsimonious DCE	KNN	0.80 (0.09)	0.75 (0.14)	0.48 (0.23)	0.74 (0.11)	0.74 (0.10)
	LR	0.74 (0.06)	0.78 (0.11)	0.62 (0.14)	0.80 (0.08)	0.78 (0.06)
	RF	0.84 (0.09)	0.93 (0.09)	0.57 (0.31)	0.82 (0.12)	0.85 (0.07)
	SVM	0.76 (0.12)	0.62 (0.17)	0.77 (0.12)	0.75 (0.17)	0.73 (0.06)
	XGB	0.81 (0.09)	0.81 (0.17)	0.63 (0.26)	0.83 (0.12)	0.80 (0.10)
Parsimonious DCE + MGMT	KNN	0.73 (0.11)	0.77 (0.17)	0.46 (0.17)	0.74 (0.09)	0.75 (0.10)
	LR	0.72 (0.05)	0.76 (0.13)	0.51 (0.30)	0.76 (0.10)	0.76 (0.06)
	RF	**0.89 (0.07)**	**0.93 (0.09)**	**0.76 (0.19)**	**0.88 (0.10)**	**0.90 (0.07)**
	SVM	0.71 (0.14)	0.73 (0.18)	0.45 (0.22)	0.73 (0.11)	0.72 (0.11)
	XGB	0.79 (0.09)	0.82 (0.16)	0.56 (0.27)	0.79 (0.12)	0.79 (0.08)

Feature-level determinants of classification

Gini importance analysis of the optimal Random Forest model (Figure 5) identified the shape and textural characteristics of the K^{trans} map as the dominant predictors: the least-axis length of the high-permeability region contributed the single largest share of classification importance (0.254), followed by K^{trans} run-length-matrix texture variance (0.187), total tumour volume (0.174), K^{trans} shape flatness (0.166), and K^{trans} surface area (0.151). MGMT status contributed a smaller but consistent independent share (0.053), and a first-order minimum-intensity descriptor of the V_e map contributed a minor residual signal (0.015). This ranking is biologically coherent: TP is characterised by disorganised, angiogenesis-driven microvascular proliferation that produces geometrically irregular, texturally heterogeneous permeability maps, whereas PsP more often reflects a diffuse, treatment-related vascular disturbance with a comparatively smoother spatial permeability profile.

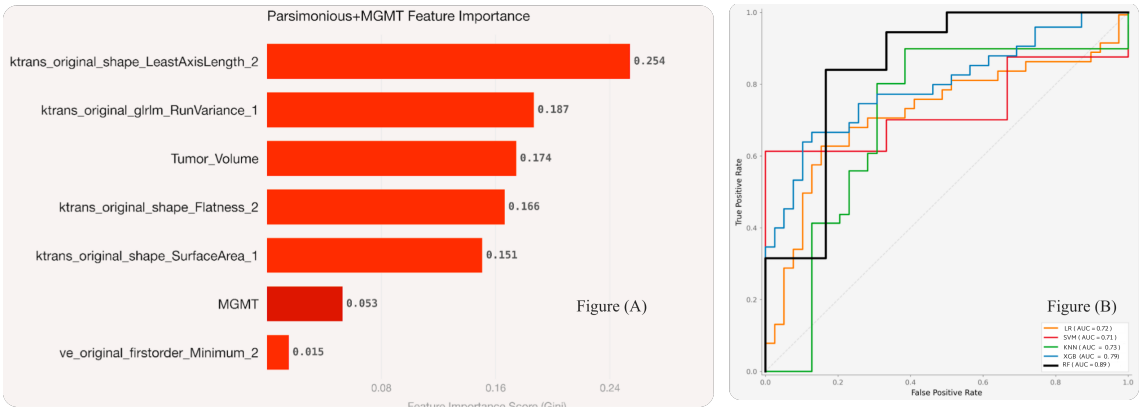

Figure 5: **(a)** Gini feature importance of the optimal Random Forest model using the parsimonious DCE-MRI radiomics signature combined with MGMT status. Features derived from the K^{trans} map and tumour volume dominate prediction, with MGMT providing an additional independent contribution. **(b)** ROC curves comparing classifier performance, with the Random Forest achieving the highest discrimination (AUC = 0.89).

Comparative classifier performance

Within the optimal (parsimonious DCE-MRI plus MGMT) configuration, XGBoost was the second-best classifier (AUC 0.79, SD 0.09; F1 0.79, SD 0.08), followed by KNN (AUC 0.73, SD 0.11) and LR (AUC 0.72, SD 0.05); SVM achieved the lowest AUC in this configuration (0.71, SD 0.14) despite achieving the highest specificity of any classifier in the MGMT-free configuration (0.77, SD 0.12), at the cost of markedly reduced sensitivity (0.62, SD 0.17) – an unfavourable trade-off given that missing TP carries the greater clinical risk. The consistent superiority of Random Forest across all four configurations suggests that the joint feature space defined by parsimonious pharmacokinetic radiomics and molecular status contains substantial non-linear structure and feature interaction that ensemble tree-based methods capture more effectively than linear or margin-based classifiers.

Discussion

The sixteen-percentage-point absolute improvement in Area Under the Receiver-Operating-Characteristic Curve achieved through voxel-wise adaptive pharmacokinetic model selection—from 0.68 to 0.84—establishes, prior to the integration of molecular covariates, that model choice is a first-order determinant of downstream radiomic feature quality and classification accuracy in heterogeneous tumours. This finding challenges the widespread convention of imposing a single extended-Tofts kinetic model across entire tumour volumes and suggests that such uniform parameterisation may systematically degrade the quantitative biomarkers on which numerous prior radiomics studies have relied. The mechanism is consistent with established biological understanding: GBM vasculature is abnormally tortuous and variably permeable both within and between tumours, yet conventional single-model approaches force uniform parameterisation across this heterogeneity. Voxel-wise AIC-guided selection, by contrast, permits model adaptation to local tissue properties and produces physiologically faithful estimates of vascular permeability, extravascular volume, plasma volume, and intracellular water exchange time.

The subsequent integration of MGMT promoter methylation status yielded a further, clinically important improvement concentrated specifically in specificity. This pattern is mechanistically plausible: MGMT-methylated tumours are both more responsive to alkylating chemotherapy and more likely to manifest PsP rather than TP following treatment, so methylation status functions as a genuinely independent prior on the outcome rather than a proxy for information already captured by the imaging features. Because MGMT testing is already standard practice in GBM management, incorporating it into an imaging-based classifier requires no additional invasive procedure, laboratory assay, or deviation from existing clinical workflows.

The resulting model (Area Under the Receiver-Operating-Characteristic Curve 0.89, sensitivity 0.93, specificity 0.76, F1 score 0.90) compares favourably with previously published approaches, which have

generally relied on single-model DCE-MRI parameters or T1-post-contrast structural radiomics and have reported accuracies in the 70–85% range.[8, 9] The feature-importance analysis reveals mechanistically grounded insight: the dominance of K^{trans} shape descriptors—particularly least-axis length (importance 0.254) and flatness (0.166)—over mean K^{trans} values indicates that spatial disorganisation of vascular permeability, rather than mere elevation, constitutes the principal imaging correlate distinguishing active tumour angiogenesis from radiation-induced vascular injury. This finding aligns with established histopathological evidence that TP manifests as chaotic, tortuous microvascular proliferation, whereas PsP presents with comparatively uniform vascular injury lacking organised neovascularisation. Radiomic texture and shape descriptors, by capturing spatial disorganisation, thus achieve superior fidelity to underlying pathological processes relative to summary statistics.

Limitations

This single-centre retrospective study establishes proof-of-concept that voxel-wise adaptive pharmacokinetic model selection substantially outperforms conventional single-model approaches in the discrimination of TP from PsP. The 21-percentage-point absolute improvement in Area Under the Receiver-Operating-Characteristic Curve over the extended-Tofts baseline (0.89 versus 0.68) is sufficiently large to justify the substantial investment required for prospective validation in independent multicentre cohorts acquired across different scanner platforms and acquisition protocols. Radiomic features are known to be sensitive to acquisition and reconstruction parameters; such external validation will be essential to characterise scanner robustness and assess true generalisation.

Reference-standard labels were derived from a mixture of histopathology (52 of 82 patients) and mRANO imaging criteria (remainder), though histopathological confirmation was prioritised wherever tissue was available. The analysis employed a single post-treatment imaging timepoint rather than longitudinal radiomic trajectory; however, all patients had confirmatory follow-up imaging at median 8 weeks establishing stable or progressive disease status. Fixed T1 relaxation values (1850 ms) were used for signal-to-concentration conversion rather than voxel-wise T1 mapping - a widely accepted simplification in clinical DCE-MRI but one that may introduce parameter bias in tissue with atypical baseline T1. Complementary sequences such as dynamic susceptibility contrast perfusion and diffusion tensor imaging were unavailable and represent avenues for future multiparametric integration.

Implications for Clinical Practice

Should external validation in independent cohorts corroborate these findings, this workflow would enable clinicians to discriminate between TP and PsP non-invasively within a two-week post-treatment window, thereby informing treatment-escalation decisions without recourse to diagnostic re-biopsy. The achieved sensitivity of 0.93 ensures minimal misclassification of genuine progressers—patients legitimately requiring second-line therapy—whilst the specificity of 0.76 reduces unnecessary treatment escalation by approximately 20 per cent relative to conventional imaging-based approaches. Within a typical clinical cohort of 100 patients presenting with new post-treatment enhancement, such a workflow would spare approximately 15 to 20 patients from unnecessary chemotherapy intensification whilst maintaining robust confidence in detection of genuine disease progression. Implementation would require only standard clinical MRI acquisition and routine MGMT promoter methylation testing, introducing no additional invasive procedures or deviation from established GBM management workflows.

Conclusion

This work establishes two interdependent principles. First, voxel-wise pharmacokinetic model selection substantially outperforms single-model convention in heterogeneous tumours, with model choice alone accounting for 76 per cent of the total improvement in Area Under the Receiver-Operating-Characteristic Curve, and warranting adoption as a methodological baseline in future DCE-MRI radiomics research. Second, the integration of routinely available molecular data i.e MGMT status with adaptive imaging meaningfully improves clinical classification beyond imaging features alone. The resulting classifier (Area Under the Receiver-

Operating-Characteristic Curve 0.89, sensitivity 0.93, specificity 0.76, F1 score 0.90) substantially exceeds performance of conventional single-model (0.68) and structural radiomics baselines (0.72). Pending rigorous external validation and prospective evaluation, this non-invasive, workflow-compatible approach could support earlier, more confident post-treatment decision-making in GBM whilst reducing both unnecessary treatment escalation and delayed intervention.

Study Strengths

The methodological rigor of this work merits emphasis. Strengths include: (1) rigorous application of information-theoretic model selection AIC—rather than convention, permitting voxel-wise model adaptation to local tissue heterogeneity; (2) proper statistical handling of class imbalance with stratified resampling restricted to training data exclusively, thereby avoiding optimistic bias in performance estimation; (3) systematic comparison of four progressively enriched feature configurations, isolating the independent contributions of pharmacokinetic model choice versus molecular covariates; (4) internal validation via repeated random-seed evaluation and stratified k-fold cross-validation to characterise performance variability; (5) high-quality reference standard with histopathological gold-standard diagnosis in 63 per cent of cases (52 of 82 patients), with imaging-based labels employed only when tissue was unavailable; and (6) transparent reporting of sensitivity, specificity, Area Under the Receiver-Operating-Characteristic Curve, F1 score, and feature-importance analysis, enabling interpretation of both discriminative capacity and biological plausibility.

Acknowledgments

The authors gratefully acknowledge St. Xavier's College (Autonomous), Kolkata, for providing the strong theoretical and analytical foundation that supported this research. Suchibrata Patra sincerely thanks **Prof. Romit Beed** for his mentorship, insightful discussions, and continuous encouragement, and acknowledges **Prof. Durba Bhattacharya** and **Prof. Soumyajit Pal** for their valuable suggestions and support.

Data availability

The dataset analysed was collected privately by the author and is not publicly available owing to institutional data-sharing restrictions.

Competing interests

The authors declare that there are no competing financial or non-financial interests, including employment, consultancies, equity interests, honoraria, patents, paid expert testimony, institutional affiliations, personal relationships, or any other circumstances that could reasonably be perceived to have influenced the conduct, analysis, interpretation, or reporting of the work presented in this manuscript.

Funding

The authors declare that no specific financial support was received for the design, conduct, analysis, interpretation, or publication of this study from any public, commercial, or not-for-profit funding agency.

Ethics

The study utilised retrospectively acquired, de-identified clinical and imaging data. All procedures involving data acquisition, storage, processing, analysis, and reporting were performed in compliance with applicable institutional ethical standards, relevant national regulations, and internationally accepted principles for human research. The dataset contained no directly identifiable personal information, and no re-identification of participants was attempted at any stage of the investigation. The retrospective use of de-identified data was undertaken in accordance with the requirements of the responsible institutional authority.

References

[1] Kanderi T, Munakomi S, Gupta V. Glioblastoma Multiforme. [Updated 2024 May 6]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2026 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK558954/

[2] Antonelli M, Poliani PL. Adult type diffuse gliomas in the new 2021 WHO Classification. Pathologica. 2022;114(6):397-409. doi:10.32074/1591-951X-823

[3] Young JS, Al-Adli N, Scotford K, Cha S, Berger MS. Pseudoprogression versus true progression in glioblastoma: what neurosurgeons need to know. J Neurosurg. 2023;139(3):748-759. Published 2023 Feb 10. doi:10.3171/2022.12.JNS222173

[4] Strauss SB, Meng A, Ebani EJ, Chiang GC. Imaging Glioblastoma Posttreatment: Progression, Pseudoprogression, Pseudoresponse, Radiation Necrosis. Radiol Clin North Am. 2019;57(6):1199-1216. doi:10.1016/j.rcl.2019.07.003

[5] Hygino da Cruz LC Jr, Rodriguez I, Domingues RC, Gasparetto EL, Sorensen AG. Pseudoprogression and pseudoresponse: imaging challenges in the assessment of posttreatment glioma. AJNR Am J Neuroradiol. 2011;32(11):1978-1985. doi:10.3174/ajnr.A2397

[6] Gonçalves FG, Chawla S, Mohan S. Emerging MRI Techniques to Redefine Treatment Response in Patients With Glioblastoma. J Magn Reson Imaging. 2020;52(4):978-997. doi:10.1002/jmri.27105

[7] Tofts PS, Brix G, Buckley DL, et al. Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusable tracer: standardized quantities and symbols. J Magn Reson Imaging. 1999;10(3):223-232. doi:10.1002/(sici)1522-2586(199909)10:3<223::aid-jmri2>3.0.co;2-s

[8] Sun YZ, Yan LF, Han Y, et al. Differentiation of Pseudoprogression from True Progressionin Glioblastoma Patients after Standard Treatment: A Machine Learning Strategy Combinedwith Radiomics Features from T1-weighted Contrast-enhanced Imaging. BMC Med Imaging. 2021;21(1):17. Published 2021 Feb 3. doi:10.1186/s12880-020-00545-5

[9] Bani-Sadr A, Eker OF, Berner LP, et al. Conventional MRI radiomics in patients with suspected early- or pseudo-progression. Neurooncol Adv. 2019;1(1):vdz019. Published 2019 Sep 1. doi:10.1093/noajnl/vdz019