



Histopathology-Validated Machine Learning Radiographic Biomarker for Noninvasive Discrimination Between True Progression and Pseudo-progression in Glioblastoma

Hamed Akbari, MD, PhD ^{1,2}; Saima Rathore, PhD^{1,2}; Spyridon Bakas, PhD^{1,2,3}; MacLean P. Nasrallah, MD, PhD³; Gaurav Shukla, MD, PhD^{1,2,4,5}; Elizabeth Mamourian^{1,2}; Martin Rozycki, MSc^{1,2}; Stephen J. Bagley, MD, MSCE ⁶; Jeffrey D. Rudie, MD, PhD²; Adam E. Flanders, MD⁷; Adam P. Dicker, MD, PhD⁸; Arati S. Desai, MD⁶; Donald M. O'Rourke, MD⁹; Steven Brem, MD⁹; Robert Lustig, MD⁴; Suyash Mohan, MD²; Ronald L. Wolf, MD, PhD²; Michel Bilello, MD, PhD^{1,2}; Maria Martinez-Lage, MD¹⁰; and Christos Davatzikos, PhD^{1,2}

BACKGROUND: Imaging of glioblastoma patients after maximal safe resection and chemoradiation commonly demonstrates new enhancements that raise concerns about tumor progression. However, in 30% to 50% of patients, these enhancements primarily represent the effects of treatment, or pseudo-progression (PsP). We hypothesize that quantitative machine learning analysis of clinically acquired multiparametric magnetic resonance imaging (mpMRI) can identify subvisual imaging characteristics to provide robust, noninvasive imaging signatures that can distinguish true progression (TP) from PsP. **METHODS:** We evaluated independent discovery ($n = 40$) and replication ($n = 23$) cohorts of glioblastoma patients who underwent second resection due to progressive radiographic changes suspicious for recurrence. Deep learning and conventional feature extraction methods were used to extract quantitative characteristics from the mpMRI scans. Multivariate analysis of these features revealed radiophenotypic signatures distinguishing among TP, PsP, and mixed response that compared with similar categories blindly defined by board-certified neuropathologists. Additionally, inter-institutional validation was performed on 20 new patients. **RESULTS:** Patients who demonstrate TP on neuropathology are significantly different ($P < .0001$) from those with PsP, showing imaging features reflecting higher angiogenesis, higher cellularity, and lower water concentration. The accuracy of the proposed signature in leave-one-out cross-validation was 87% for predicting PsP (area under the curve [AUC], 0.92) and 84% for predicting TP (AUC, 0.83), whereas in the discovery/replication cohort, the accuracy was 87% for predicting PsP (AUC, 0.84) and 78% for TP (AUC, 0.80). The accuracy in the interinstitutional cohort was 75% (AUC, 0.80). **CONCLUSION:** Quantitative mpMRI analysis via machine learning reveals distinctive noninvasive signatures of TP versus PsP after treatment of glioblastoma. Integration of the proposed method into clinical studies can be performed using the freely available Cancer Imaging Phenomics Toolkit. **Cancer** 2020;0:1-12. © 2020 American Cancer Society.

KEYWORDS: glioblastoma, machine learning, pseudo-progression, radiographic biomarker, true progression.

INTRODUCTION

Glioblastoma is the most common malignant primary adult brain tumor.¹ It is associated with a poor prognosis, with median overall survival ranging from 16 to 20 months despite maximal treatment.²⁻⁴ Initial treatment of glioblastoma involves a combined approach of maximal safe surgical resection and adjuvant chemoradiation.³ Surveillance of patients after completion of this initial treatment relies heavily on follow-up serial magnetic resonance imaging (MRI) to detect disease recurrence. As such, distinguishing treatment-induced MRI changes from true progression (TP) of a tumor has critical implications to clinical decision making.

Unfortunately, standardization of radiographic metrics for treatment response and for determining disease progression in glioblastoma has proven quite difficult. The widespread variance in definitions of the terms *progressive disease* and *stable disease* led to the development of the Macdonald criteria in 1990, which rely upon crude radiographic

Corresponding Authors: Hamed Akbari, MD, PhD, Center for Biomedical Image Computing and Analytics, University of Pennsylvania, 3700 Hamilton Walk, 7th Floor, Philadelphia, PA 19104 (hamed.akbari@pennmedicine.upenn.edu); Christos Davatzikos, PhD, Center for Biomedical Image Computing and Analytics, University of Pennsylvania, 3700 Hamilton Walk, 7th Floor, Philadelphia, PA 19104 (christos.davatzikos@pennmedicine.upenn.edu).

¹Center for Biomedical Image Computing and Analytics, University of Pennsylvania, Philadelphia, Pennsylvania; ²Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; ³Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; ⁴Department of Radiation Oncology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; ⁵Helen F. Graham Cancer Center and Research Institute, ChristianaCare, Newark, Delaware; ⁶Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; ⁷Department of Radiology, Thomas Jefferson University Hospital, Philadelphia, Pennsylvania; ⁸Department of Radiation Oncology, Sidney Kimmel Medical College and Cancer Center, Thomas Jefferson University, Philadelphia, Pennsylvania; ⁹Department of Neurosurgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; ¹⁰Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts

Additional supporting information may be found in the online version of this article.

DOI: 10.1002/cncr.32790, **Received:** June 2, 2019; **Revised:** December 10, 2019; **Accepted:** January 22, 2020, **Published online** Month 00, 2020 in Wiley Online Library (wileyonlinelibrary.com)

measurement of areas of contrast enhancement on post-treatment MRI.⁵ With the addition of temozolomide to the standard of care, clinicians soon determined that these criteria were unable to accurately distinguish between TP and the effects of treatment, and the phenomenon of pseudo-progression (PsP) was thus identified. PsP is a subacute treatment-related effect that usually occurs within 3 months of completion of chemoradiation^{6,7} and has imaging characteristics that mimic TP as defined by the Macdonald criteria.⁵

The diagnosis of PsP is usually made on the basis of spontaneous improvement or stabilization of imaging findings over several months (ie, in the setting of continuation of the chemotherapy for at least 6 months). Previous studies have suggested that nearly half (30%-50%) of glioblastoma patients with worsened radiographic findings after standard chemoradiation have PsP rather than TP.⁷⁻¹⁴ Because traditional MRI cannot reliably distinguish TP from PsP, clinicians who care for patients with glioblastoma must frequently choose between declaring TP (and therefore modifying the patient's current therapy) versus proceeding with invasive brain surgery for diagnostic clarity. On histopathologic analysis, such surgeries may reveal recurrent glioblastoma tumor, therapy-related changes (ie, PsP), or a combination of the two.

The Response Assessment in Neuro-Oncology (RANO) Working Group developed new criteria¹⁵ to address some of the limitations of the Macdonald criteria. As clinicians recognized the high prevalence of PsP in the months immediately after completion of chemoradiation, subtle changes in posttreatment imaging are no longer deemed to represent progressive disease, unless there is evidence of clinical deterioration or obvious new disease outside of the treatment field. This classification allows patients to continue maintenance therapy safely, with the goal of achieving some delayed improvement. Nonetheless, some of these patients do actually show TP, and identifying these patients without tissue sampling would allow for earlier change in therapy.

The goal of this study was to noninvasively evaluate radiographic changes in glioblastoma patients treated with chemoradiation by multivariate analysis of preoperative, multiparametric MRI (mpMRI) to identify a radiophenotypic signature to distinguish between TP, mixed response, and PsP. Analysis of radiographic data via advanced computational analytics has been increasingly shown to provide rich and highly informative characterizations of glioblastoma and its surrounding brain tissue,^{11,16-21} extending the evaluation of tissue properties beyond the capabilities

of human visual interpretation. We hypothesize that quantification of subtle yet spatially complex quantitative imaging phenomic (QIP) features extracted from mpMRI can facilitate noninvasive classification of TP versus PsP with sufficient sensitivity and specificity to allow discrimination on an individual patient basis.

MATERIALS AND METHODS

Study Patient Population

The study population was identified on the basis of retrospective review of the electronic medical record of patients diagnosed with glioblastoma at the Hospital of the University of Pennsylvania between 2011 and 2015. The criteria for inclusion comprised 1) initial gross total resection of the tumor followed by standard radiation therapy and temozolomide chemotherapy, 2) demonstration of new/increasing enhancement areas on follow-up MRI within 6 months after completion of radiation therapy, 3) second resection for histopathological tissue evaluation, and 4) acquisition of mpMRI within 15 days prior to the second resection. We identified 63 patients (men, $n = 38$; women, $n = 25$; mean age, 57.28 years [range, 32.79-81.60 years]) satisfying these inclusion criteria, and we randomly divided them into independent discovery ($n = 40$ [23 TP, 6 PsP, 11 mixed response]) and replication ($n = 23$ [12 TP, 4 PsP, 7 mixed response]) cohorts. Isocitrate dehydrogenase 1 (*IDH1*) was wild-type, mutant, and not otherwise specified for 52, 2, and 9 patients, respectively (Supporting Table S1). This study was approved by the institutional review board of the University of Pennsylvania and was compliant with the Health Insurance Portability and Accountability Act.

Histopathological Tissue Evaluation

After resection, the surgically extracted tissue specimens were entirely fixed in 10% buffered formalin, routinely processed, and embedded in paraffin. Five-micrometer-thick sections of each specimen were cut onto glass slides, stained with hematoxylin and eosin, and assessed by 2 board-certified neuropathologists (M.P.N., M.M.-L.) (blinded to our imaging assessment and the other rater) for presence of apparent tumor features and reactive treatment-related changes.¹¹ The presence or absence of pseudo-palisading necrosis and microvascular proliferation (which are features of recurrent glioblastoma), the presence or absence of dystrophic calcification and vascular hyalinization, and the percentage of geographic necrosis (representative of treatment-related changes) were quantified (Fig. 1). Proliferative activity was determined by quantification of the number of mitotic figures

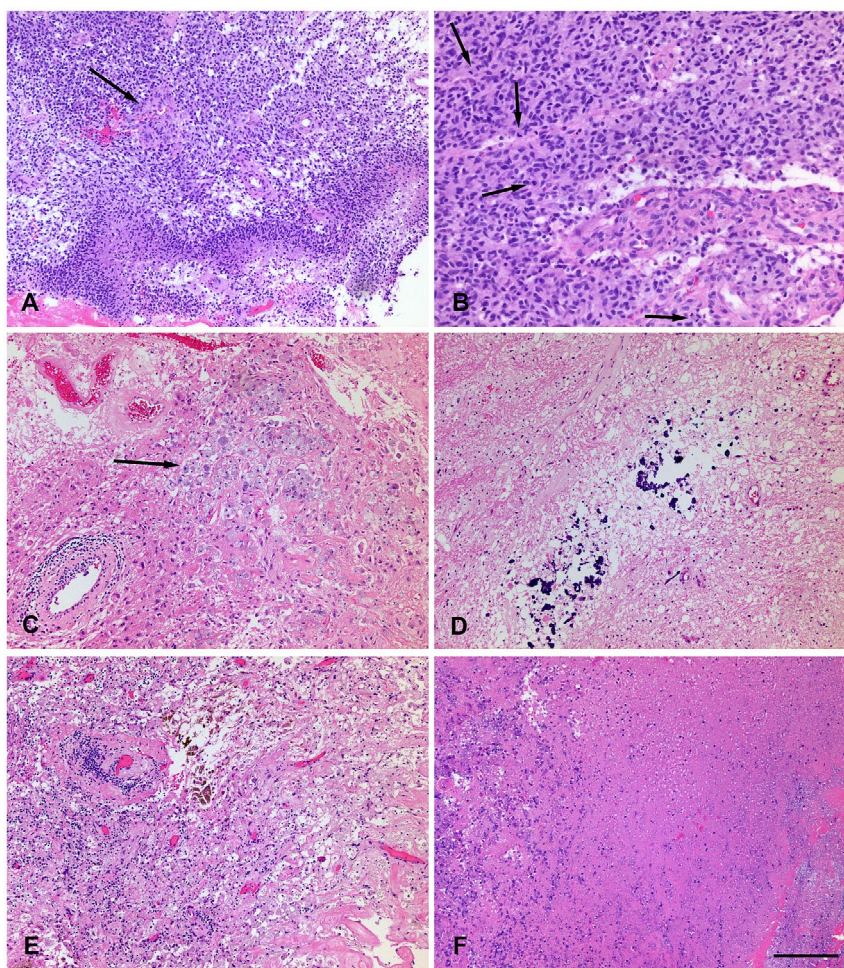


Figure 1. Histological features of (A, B) true progression (malignant tumor) and (C-F) pseudo-progression (treatment-related effects). (A) Highly cellular tumor with pseudo-palisading necrosis and microvascular proliferation (arrow). (B) Mitotically active cellular tumor with microvascular proliferation (mitoses are indicated by arrows). (C) Thickened hyalinized vessel walls with focal lymphocytic inflammation and macrophages (arrow). (D) Dystrophic calcification. (E) Hyalinized vessels, lymphocytes, macrophages and hemosiderin-laden macrophages. (F) Geographic necrosis. Scale bar = 200 μ m (A, C-F) and 100 μ m (B).

in 10 high-power fields and semiquantitative assessment of Ki-67 proliferative index by immunostaining (mouse monoclonal, MIB-1, IR62661; Dako, Carpinteria, California). Based on the combined assessment of these features, the entire resected specimen was assigned a score between 1 and 6 (1 = <10%, 2 = 10%-25%, 3 = 25%-50%, 4 = 50%-75%, 5 = 75%-90%, and 6 = >90% malignant features). A score of 1-2 was defined as PsP, 3-4 as a mixture of TP and PsP, and 5-6 as TP. PsP (score 1-2) may continue the treatment as it was before. The mixture of TP and PsP (score 3-4) may change the treatment or continue the current treatment based on clinical status of the patient. The TP patients (score 5-6) may be recommended for repeat resection. We used linear weighted Cohen's kappa to calculate the interrater agreement.

MRI Acquisition Protocol

All MRI scans were performed on a Magnetom Tim Trio 3 Tesla scanner (Siemens, Erlangen, Germany) using a 12-channel phased array head coil. Routine sequences included axial T1-weighted (T1): matrix = $192 \times 256 \times 192$, resolution = $0.98 \times 0.98 \times 1.00$ mm³, repetition time (TR) = 1760 ms; echo time (TE) = 3.1 ms; T1-weighted contrast enhanced with gadolinium (T1-Gd): matrix = $192 \times 256 \times 192$, resolution = $0.98 \times 0.98 \times 1.00$, TR = 1760 ms, TE = 3.1 ms; T2-weighted (T2): matrix = $208 \times 256 \times 64$, resolution = $0.94 \times 0.94 \times 3.00$, TR = 4680 ms, TE = 85 ms; T2 fluid-attenuated inversion recovery (T2-FLAIR): matrix = $192 \times 256 \times 60$, resolution = $0.94 \times 0.94 \times 3.00$, TR = 9420 ms,

TE = 141 ms; and diffusion tensor imaging (DTI): matrix = $128 \times 128 \times 40$, resolution = $1.72 \times 1.72 \times 3.00$ in 30 gradient directions, from which fractional anisotropy (DTI-FA), radial diffusivity (DTI-RAD), axial diffusivity (DTI-AX), and trace (DTI-TR) maps were calculated. The parameters for dynamic susceptibility contrast MRI (DSC-MRI) were as follows: field of view = 22 cm, matrix = $128 \times 128 \times 20$, resolution = $1.72 \times 1.72 \times 3 \text{ mm}^3$, TR = 2000 ms, TE = 45 ms. The DSC-MRI sequences were acquired as follows: After an initial loading dose of 3 mL MultiHance (gadobenate dimeglumine) was administered to reduce the effect of contrast agent leakage, and another bolus injection was given after 5 minutes with the remaining dose (for a total of 0.3 mL/kg or 1.5 times single dose) during image acquisition (15 patients). With the evolution of clinical protocols, dynamic contrast-enhanced (DCE, also known as permeability) acquisitions have been routinely obtained on more recent studies. In these instances, DCE is obtained first with half of the total contrast and serves as the loading dose to reduce the effect of contrast agent leakage, followed by an additional bolus after a similar delay with the second half of the total contrast volume for the DSC acquisition (total 0.3 mL/kg, 48 patients).

MRI Preprocessing

All MRI volume scans of each individual patient were affinely coregistered intrapatient using the Functional MRI of the Brain Software Library (FSL).²² Subsequently all scans were smoothed to remove any high frequency intensity variations (ie, noise),²³ corrected for magnetic field inhomogeneities²⁴ and skull-stripped using FSL BET²⁵ followed by manual revision when needed. We extracted commonly used measurements²⁶ from the acquired DTI volumes (ie, DTI-TR, DTI-AX, DTI-RAD, DTI-FA). The DSC-MRI volumes were used to computationally extract parametric brain maps of the relative cerebral blood volume (rCBV), peak height (PH), and percentage signal recovery (PSR) after considering leakage correction.^{27,28} In addition, all DSC curves were aligned/normalized interpatient for baseline and maximum drop. Principal component analysis (PCA) was also employed to extract a summarized signal of the complete temporal perfusion dynamics encapsulated in the DSC-MRI modality, instead of using isolated measurements such as the rCBV.²⁹ Finally, the machine learning approach we adapted considers all 4 structural MRI images (T1, T1-Gd, T2, T2-FLAIR), the subtraction of T1 from T1-Gd and T2-FLAIR from T2 (following intensity

normalization), 4 DTI-derived measurements, perfusion-derived PCA volumes, and isolated perfusion-derived parametric brain maps. We refer to these image volumes collectively as mpMRI.

Defining Target Tissue

To define the target region of interest (ROI), we first registered the pre- and postoperative images of the time point suspicious for TP using Deformable Registration via Attribute Matching and Mutual-Saliency Weighting (DRAMMS) software.³⁰ We then delineated the ROI describing the resected tissue. The regions in the pre-surgery images that correspond to the resected tissue in postsurgery images was defined as ROI. Therefore, the imaging properties and the pathology features corresponded to the same resected region. Following the definition of these ROIs, all mpMRI sequences were analyzed to extract relevant comprehensive QIP features from the corresponding ROIs to create our predictive model.

Feature Extraction

Considering the complexity of the problem our model attempts to address, we used 2 distinct approaches to ensure the comprehensiveness of the extracted QIP features from the mpMRI volumes, and hence that they describe all aspects of the radiographic appearance. The 2 approaches were distinctively based on deep learning and a priori selected (APS) feature extraction. We used and evaluated these approaches both in combination and in comparison (Fig. 2).

Deep Learning Features

To obtain the deep learning features, we used a pretrained neural network^{31,32} and adapted a convolutional neural network (CNN) model pretrained on 1.2 million 3-channel images of the ImageNet LSVRC-2010 for classifying real-world images into different classes.³¹ The exact CNN model we used (imagenet_vgg_f³¹) was provided by the VLfeat library³² as part of their MatLab toolbox (MatConvNet) for computer vision applications. This CNN is a type of deep feed forward neural network that uses multilayer perceptrons with hidden layers. The hidden layers of CNN comprise convolutional (ie, cross-correlation) layers, pooling layers, fully connected layers, and normalization layers.^{33,34} The convolution layer, which makes CNNs different from other types of deep neural networks, is a main layer of CNN and consists of several adaptive filters (as kernels) with small receptive

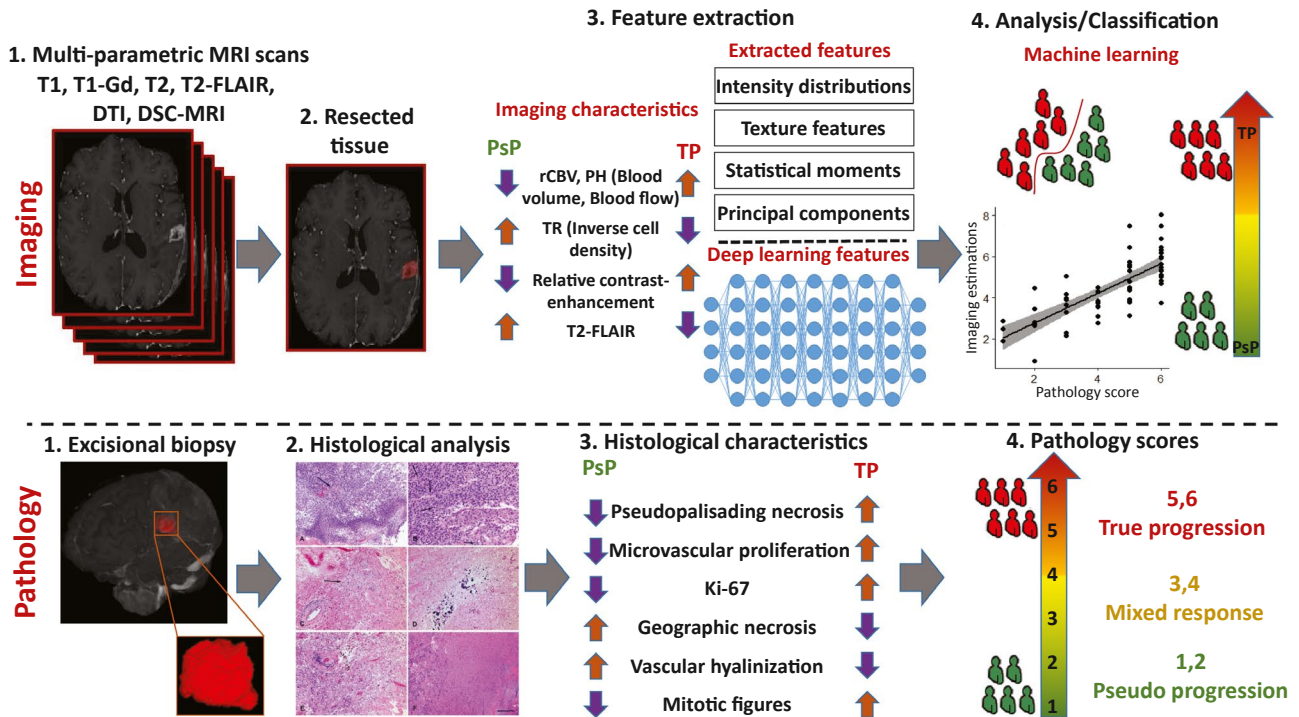


Figure 2. Processing pipeline of the proposed radiomic analysis. Top row (imaging): 1. T1-w pre- and postcontrast, T2-w, T2-FLAIR, diffusion tensor imaging, and dynamic susceptibility contrast magnetic resonance images are acquired. 2. Defining the resected enhancing tissues after registration of the preoperative with postoperative images using Deformable Registration via Attribute Matching and Mutual-Saliency Weighting (DRAMMS) software. 3. Features are extracted from each region, quantifying intensity, shape, principal component analysis, statistics, and texture. 4. Features analysis and classification between pseudo-progression (PsP) versus non-PsP and true progression (TP) versus non-TP. Bottom row (pathology): 1. Excisional biopsy. 2. Histological analysis. 3. Histological characteristics. 4. Pathology scores. Abbreviations: DSC-MRI, dynamic susceptibility contrast magnetic resonance imaging; DTI, diffusion tensor imaging; FLAIR, fluid-attenuated inversion recovery; Gd, gadolinium; MRI, magnetic resonance imaging; PH, peak height; PsP, pseudo-progression; TP, true progression; TR, repetition time.

fields. To apply this pretrained model in our data, we created 7 artificial 3-channel images, each channel of which describes an individual sequence from all the mpMRI images considered (Supporting Table S2).

APS Radiomic Features

We extracted 1040 APS radiomic features using the Cancer Imaging Phenomics Toolkit (www.cbica.upenn.edu/captk).³⁵ Specifically, the features extracted describe the first-order statistical distribution of voxel intensities within each ROI (comprising mean, median, maximum, minimum, skewness, and standard deviation) in all mpMRI sequences, the PCA summarized signal of the intensity distribution histogram of each mpMRI sequence, and texture features (second-order statistics) based on gray-level cooccurrence matrix (GLCM)³⁶ and gray-level run length matrix (GLRLM).³⁷ To obtain these texture features in 3 dimensions, all mpMRI volumes were first quantized to 16 gray levels within the ROI. GLCM and GLRLM were

then populated by taking 13 main directions into account. A neighborhood of $3 \times 3 \times 3$ was considered for GLCM. These features were first computed for each direction independently and then averaged to find their final value. All features were rescaled via z score normalization before machine learning analysis. To identify which of these 1040 extracted features had actual predictive value, we applied a sequential feature selection in the training data until convergence based on a threshold in the accuracy improvement.

Classification and Correlation Approach

A multivariate pattern classification method known as Support Vector Machines (SVM) was used to construct 2 classifiers to predict TP/PsP; 1 classifier to distinguish between PsP (scores 1-2) versus everything else (scores 3-6), and another classifier to distinguish between TP (scores 5-6) versus everything else (scores 1-4). We conducted this multivariate analysis through linear configuration of SVM. The parameter for the soft margin cost function (C)

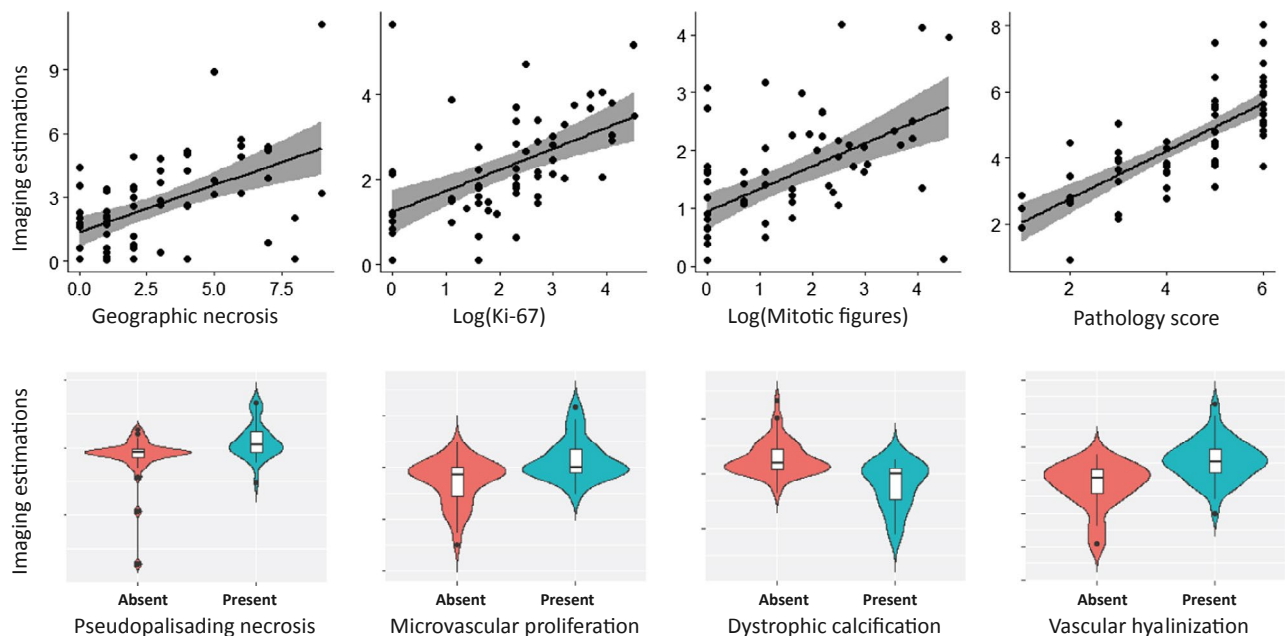


Figure 3. Histopathologic characteristics versus machine learning estimations. The gray bands in the scatter plots represent 95% CIs for the regression lines.

was optimized on the training data, based on a 5-fold cross-validated grid search; $C = 2^\alpha$, where $\alpha \in [-5, 5]$. This parameter controls the influence of each individual support vector that involves a trading error penalty for stability. These classifiers were trained separately for each classification task, and each time 1 of the 2 types of features was used (ie, deep learning and APS features). The classifiers were trained on the discovery cohort ($n = 40$) and tested on the replication cohort ($n = 23$). To confirm the robustness, accuracy, and generalizability of the proposed method in a larger cohort while avoiding optimistically biased estimates of performance, we also evaluated the classifiers in all 63 patients using a leave-one-out cross-validation (LOOCV) schema, where in every iteration of cross-validation, the features were selected using data of $n-1$ patients and tested on the left-out n th patient. All steps including feature selection and model development were performed through cross-validation.

In addition to identifying a noninvasive signature to distinguish between TP and PsP, we also tried to find the correlations between the APS features and the histopathologic characteristics of the resected tissue specimen. This approach should identify complementary information of the extracted features by their correlations with pathological evaluations and quantitative scores. To achieve this, we used all 63 patients and trained separate support vector regression (SVR) models in a LOOCV configuration

for the histopathologic characteristics with continuous values (ie, mitotic figures, Ki-67, geographic necrosis, and the overall histopathology score) and SVM classification models for the histopathologic characteristics with discrete/binary values (ie, pseudo-palisading necrosis, microvascular proliferation, dystrophic calcification, and vascular hyalinization) (Fig. 3).

Interinstitutional Validation

We evaluated our method on an independent testing cohort from a different institute and with different acquisition protocols ($n = 20$ patients [10 TP, 10 PsP]). In particular, we trained the model on the dataset acquired from the University of Pennsylvania and applied the model on an independent cohort acquired from Thomas Jefferson University. Due to the lack of diffusion tensor imaging in the Thomas Jefferson University dataset, we created a model using structural, DSC perfusion and apparent diffusion coefficient imaging sequences from University of Pennsylvania and tested the model on Thomas Jefferson University. Due to lack of pathology, the follow-up serial MRI examinations were used to confirm prediction of PsP and TP by a board-certified neuroradiologist (M.B.). We selected time points that are distinct from therapy changes (systemic therapy and radiation therapy), to reduce the probability of treatment-related changes being measured on the

TABLE 1. Quantitative Evaluation Results

| | PsP vs Non-PsP | | | | | | TP vs Non-TP | | | | | |
|--|----------------|-------|-------------------|-----------|-------------------|-----------|----------------|-------|-------------------|-----------|-------------------|-----------|
| | Accuracy, % | AUC | Sensitivity, % | FNR, % | Specificity, % | FPR, % | Accuracy, % | AUC | Sensitivity, % | FNR, % | Specificity, % | FPR, % |
| Deep learning features (hold-out set) | 87.50 | 0.811 | 60.00 | 40 | 94.74 | 5.26 | 78.26 | 0.867 | 83.33 | 16.67 | 72.73 | 27.27 |
| APS features (hold-out set) | 86.96 | 0.842 | 75.00 | 25 | 89.47 | 10.53 | 78.26 | 0.803 | 83.33 | 16.67 | 72.73 | 27.27 |
| APS features (LOOCV) | 87.30 | 0.919 | 80.00 | 20 | 88.68 | 11.32 | 84.13 | 0.835 | 80.00 | 20 | 89.29 | 10.71 |
| Combined features (hold-out set) | 69.57 | 0.776 | 75.00 | 25 | 68.42 | 31.58 | 78.26 | 0.712 | 83.33 | 16.67 | 72.73 | 27.27 |

Abbreviations: APS, a priori selected; AUC, area under the curve; FNR, false negative rate; FPR, false positive rate; LOOCV, leave-one-out cross-validation; PsP, pseudo-progression; TP, true progression.

scans. All MRI scans were performed on a 1.5-Tesla GE Signa HDx scanner (General Electric, Milwaukee, Wisconsin) using an 8-channel phased array head coil. Routine sequences included T1: matrix = 256(4) or 512(16) × 256(4) or 512(16) × 15-30, resolution = 0.39-0.86 × 0.39-0.86 × 6-10 mm³, TR = 12.9-583.3 ms, TE = 4.1-12 ms; T1-Gd: matrix = 512 × 512 × 22-130, resolution = 0.43-0.57 × 0.43-0.57 × 1.5(18) or 7.5(2), TR = 516.7-600, TE = 7.9-12 ms; T2: matrix = 512 × 512 × 20-30, resolution = 0.39-0.49 × 0.39-0.49 × 5, TR = 2466.7-5952 ms, TE = 90.7-102.1 ms; T2-FLAIR: matrix = 512 × 512 × 20-30, resolution = 0.39-0.47 × 0.39-0.47 × 6(14) or 6.5(4) or 7.5(2), TR = 10000-10015 ms, TE = 126-148.5 ms; ADC: matrix = 256 × 256 × 30-37, resolution = 0.93(1) or 1.02(19) × 0.93(1) or 1.02(19) × 5, TR = 8000-10,000 ms, TE = 76.8-101 ms; DSC-MRI: field of view = 22 cm(17) or 24 cm(3), matrix = 128 × 128 × 15-27, resolution = 1.7(16) or 1.9(3) × 1.7(16) or 1.9(3) × 6(2) or 8(14) or 10(3), TR = 9-22.4 ms, TE = 400-2000 ms.

RESULTS

The imaging information captured using both methods was multivariately integrated via SVM to build 2 classification models: TP versus non-TP (mixed response + PsP) and PsP versus non-PsP (mixed response + TP). Table 1 provides a summary of these results. We used linear-weighted Cohen's kappa to calculate interrater agreement. The observed agreement (po) for pathological scores of 1-6 was 0.9103 with random agreement (pe) of 0.6562, Cohen's kappa = 0.7392, and kappa error = 0.1091. The observed agreement (po) for PsP/Mix/TP was 0.9138 with random agreement (pe) of 0.6231, Cohen's kappa = 0.7713, and kappa error = 0.0978.

Using Deep Learning Features

The results were similar when using the deep learning features, which concluded in extracting >28,000 features. Specifically, the classification performance was evaluated on the independent discovery and replication cohorts, and the accuracy for the PsP versus non-PsP model was 87.50% (sensitivity = 60.00%, specificity = 94.74%, area under the curve [AUC] = 0.8105) while the accuracy for the TP versus non-TP model was 78.26% (sensitivity = 83.33%, specificity = 72.73%, AUC = 0.8636) (Fig. 4, Table 1).

Using APS Features

The PsP and TP models developed on the discovery cohort when applied to the replication cohort returned an accuracy of 86.96% (sensitivity = 75.00%, specificity = 89.47%) and 78.26% (sensitivity = 83.33%, specificity = 72.73%), respectively. A receiver operating characteristic analysis also resulted in an AUC of 0.84 and 0.80 for the PsP versus non-PsP and TP versus non-TP (Fig. 4) models, respectively. While using only the APS radiomic features, the accuracy of our model using LOOCV in the pooled cohort was equal to 87.30% (sensitivity = 80.00%, specificity = 88.68%, AUC = 0.9189) for PsP versus non-PsP and 84.13% (sensitivity = 80.00%, specificity = 89.29%, AUC = 0.8347) for TP versus non-TP confirming its generalizability (Fig. 4, Table 1). The most distinctive features for these classifiers can be found in Supporting Table 3.

Integrating Deep Learning and APS Features

In addition to comparing the performance of each feature, we also evaluated the performance of their integration (Fig. 4, Table 1). Specifically, when combining these features, the accuracy for the PsP versus non-PsP model was equal to 69.57% (sensitivity = 75.00%, specificity = 68.42%, AUC, 0.7763) and for the TP versus

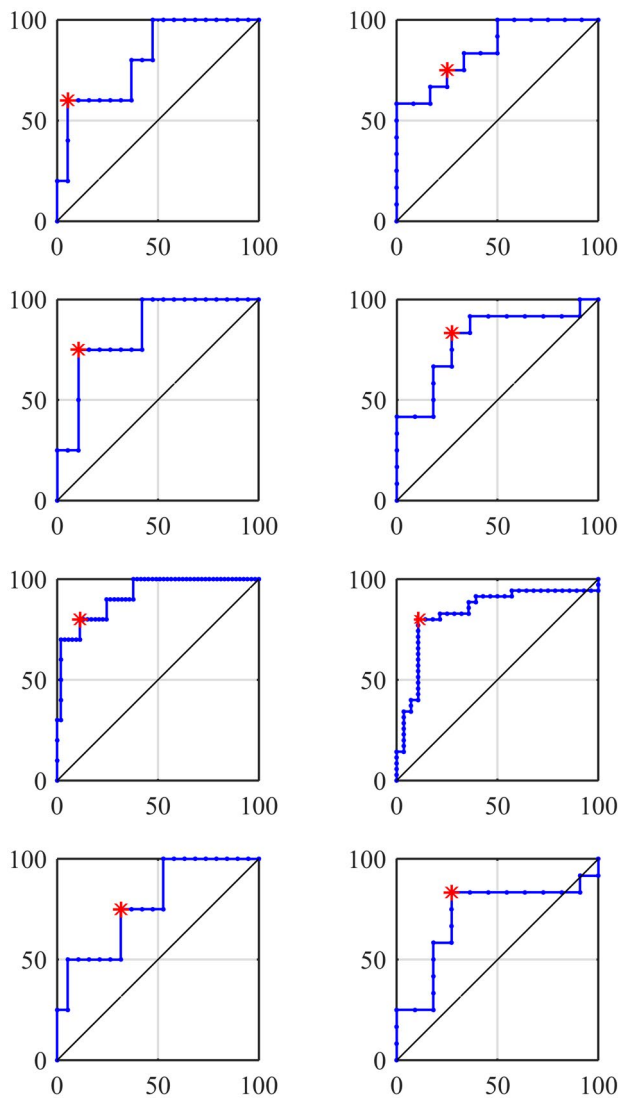


Figure 4. Receiver operating characteristic curves of different experiments for pseudo-progression (PsP) versus non-PsP (left column) and true progression (TP) versus non-TP (right column). First row: deep learning (hold-out set). Second row: a priori selected (APS) features (hold-out set). Third row: APS features (leave-one-out cross-validation). Fourth row: combined features (hold-out set). Vertical axes represent the true positive rate; horizontal axes represent the false positive rate. Asterisks represent the shortest distance from the top left point.

non-TP was 78.26% (sensitivity = 83.33%, specificity = 72.73%, AUC, 0.7121).

Histopathologic Characteristics Versus Machine Learning Estimates

The Pearson correlation coefficients between the SVR scores and logarithm of mitotic figures, logarithm of Ki-67, and geographic necrosis were estimated to be 0.54, 0.53, and 0.51, respectively. The Pearson correlation coefficients between the SVR scores and

pathology scores was 0.76 (Fig. 3). We also evaluated the trained SVM models for the presence or absence of pseudo-palisading necrosis (AUC = 0.7434), microvascular proliferation (AUC = 0.7406), dystrophic calcification (AUC = 0.8292), and vascular hyalinization (AUC = 0.7939) (Fig. 3).

Biologically Interpreted Features

In our attempt to understand the underlying biological processes that are likely to give rise to the imaging signature of TP, we performed a detailed histogram analysis of all the imaging features. The Cohen's d effect sizes were as follows: T1, 0.6477; normalized T1-Gd, 0.3103; FLAIR, 0.2599; rCBV, 0.2829; PH, 0.2907; PSR, 0.5020; TR, 0.1442; and FA, 0.3319 for the difference between TP and PsP tumors. Figure 5 shows the imaging characteristics of PsP (dashed lines) and TP (solid lines) in each modality across all patients. The main results of comparing the QIP features of TP with those of PsP patients identified TP tumors as having the following characteristics:

1. regions of higher blood volume and flow, which points toward hypervascularity, hyperperfusion, and increased angiogenesis, based on the combination of rCBV and PH;
2. regions of higher cellularity, suggestive of increased proliferation, as well as different tissue microarchitecture, based on the combination of measures extracted from DTI, namely DTI-TR and DTI-FA;
3. regions of more compromised blood–brain barrier, based on the combination of measures extracted from T1-Gd and T1-Gd subtracted T1, also consistent with infiltrating tumor characteristics; and
4. regions of lower water concentration, based on the combination of T2-FLAIR and DTI-TR, consistent with dense and nonnecrotic tissue.

Interinstitutional Validation

We evaluated the trained model on the dataset acquired from University of Pennsylvania on an independent testing cohort from Thomas Jefferson University with different acquisition protocols. Seven out of 10 TP patients and 8 out of 10 PsP patients were diagnosed correctly by the model, for an overall accuracy of 75% (AUC, 0.80). The training performance was 79% (AUC, 0.80).

DISCUSSION

In this study of glioblastoma patients treated with standard chemoradiation, we used advanced feature extraction

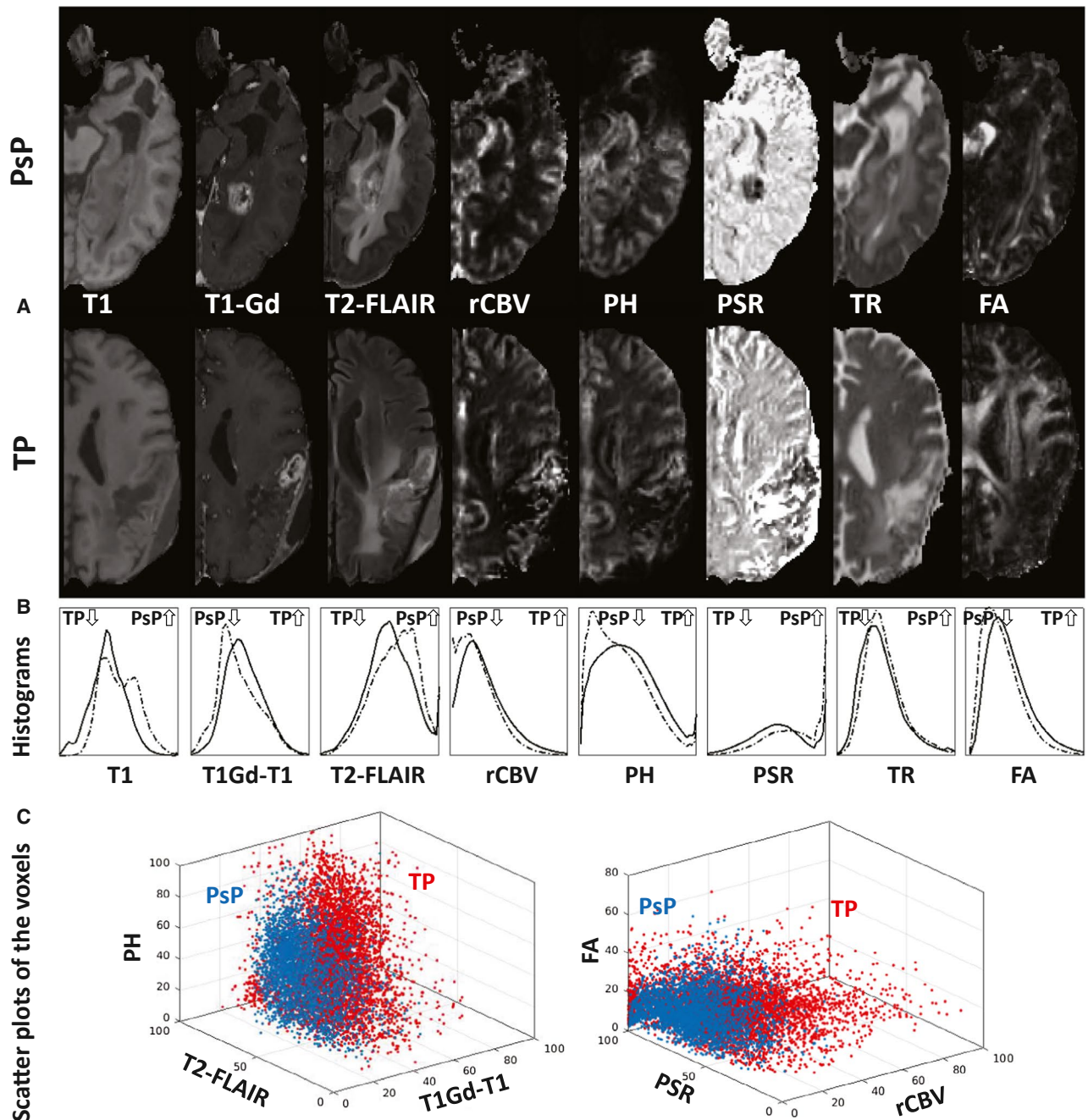


Figure 5. (A) Two examples of the imaging modalities in pseudo-progression (PsP) and true progression (TP) patients. (B) Histograms of the most distinctive modalities according to progression status. The histograms were created using information from all patients with PsP (dashed lines) and TP (solid lines). (C) Scatter plots of the voxels of patients with highest ratios of PsP (blue) and TP (red). Abbreviations: FA, fractional anisotropy; FLAIR, fluid-attenuated inversion recovery; Gd, gadolinium; PH, peak height; PsP, pseudo-progression; PSR, percentage signal recovery; rCBV, relative cerebral blood volume; TP, true progression; TR, repetition time.

and machine learning techniques to comprehensively capture the radiographic characteristics of a given ROI using structural MRI, the temporal dynamics of DSC-MRI, and DTI-derived modalities. Notably, our approach identified selected radiomic features within the given ROI

in post-chemoradiation MRI that are significantly and robustly correlated with the histopathology of resected tissue, thereby offering noninvasive means of discriminating between TP and PsP in posttreatment glioblastoma. Critically, we have made these methods and models freely

available through the Cancer Imaging Phenomics Toolkit (www.cbica.upenn.edu/captk), a publicly available open-source software platform (Supporting Fig. S1), to facilitate clinical use and further validation of these results in other studies. This software has been designed for research purposes only and has neither been reviewed nor approved for clinical use by the US Food and Drug Administration or any other federal/state agency and should not be used as the primary source of information for making clinical decisions.

Using advanced computational methodologies, our proposed noninvasive signature can quantify subtle imaging characteristics within an ROI that confer an estimate of the likelihood of TP versus PsP. It is important to emphasize that no single imaging feature was sufficiently discriminative by itself in our modeling. Rather, several QIP features were integrated by our multivariate approach to generate a discriminative score able to capture differences between TP and PsP. These results emphasize the importance of comprehensive multivariate analysis as opposed to imaging threshold based on isolated features, and the use of computational methods complementing traditional human interpretation. One of the important strengths of this signature is that it is generated from images acquired in the standard-of-care surveillance of glioblastoma patients. Thus, additional testing (invasive or otherwise) is not required, which is an advantage when considering clinical generalizability. Importantly, the QIP features are strongly associated with histopathologic scoring (Pearson correlation coefficient = 0.76, $P = 5.5 \times 10^{-13}$), and the predictive models have been validated in an independent replication cohort, unseen during their training, as well as via LOOCV.

We quantitatively evaluated the performance of our approach to distinguish between TP and PsP using models trained both independently on, and in combination with, 2 distinct feature extraction approaches: one based on deep learning and another using APS radiomic features. The comparison between the deep learning-based model and the APS radiomic features-based model yielded similar results, based on their evaluation on a relatively small independent replication cohort (Table 1). Interestingly, when combining the 2 feature types to create an integrated model, we noted a drop (>17%) in the classifier distinguishing PsP from the rest, while the accuracy of the classifier distinguishing TP from the rest remained stable. It is worth noting that the evaluation of the APS features-based model using a LOOCV scheme revealed a better performing model. A larger cohort of patients could possibly allow for further conclusions, and for developing a deep learning classifier instead of using deep

learning to create features for training an SVM model. However, considering the current results, we tend to be in favor of the model trained using APS radiomic features due to the benefit of interpretability. Larger datasets might allow deep learning achieve more reproducible and accurate results in future studies.

The feature engineering approach we consider for estimating TP can offer potential insights into biological mechanisms via each MRI sequence that may uniquely reflect radiographic phenotypes of TP versus PsP. Specifically, regions of TP in our data showed increased contrast uptake on the T1-Gd sequences, which, consistent with existing literature, may be indicative of regional angiogenesis and associated with compromise of the blood–brain barrier in areas of tumor infiltration.³⁸ T2 and T2-FLAIR sequences provide information relevant in the assessment of areas of nonenhancing and necrotic tumor, as well as the extent of the peritumoral edematous/invaded tissue.³⁹ Our results identified regions of TP demonstrating lower signal intensity on T2 and T2-FLAIR, which indicate relatively lower water content and may thus reflect higher levels of tumor infiltration. This finding is consistent with the hypothesis that the TP regions harbor a higher ratio of malignant cells to water content. DTI maps the diffusion process of water in the brain, affected in part by tumor cellularity⁴⁰ and by the integrity of white matter structures, as well as the underlying microstructure of tissue (eg, via the DTI-FA measurements). Here, regions of TP showed lower DTI-TR and increased DTI-FA that may be expected in areas of higher cellular concentration (ie, tumor cell proliferation). DTI-AX and DTI-RAD were also consistent with overall diffusivity captured by the DTI-TR volume. DSC imaging reflects various aspects of perfusion in the brain,²⁹ which provide quantitative measures of regional microvasculature, perfusion hemodynamics, and permeability of blood vessels.^{38,41} Specifically, when brain tumors exceed a critical volume, the resultant ischemia triggers the secretion of angiogenic factors that promote vascular proliferation, leading to the formation and maintenance of tumor vessels.^{42–48} These new, immature vessels tend to be tortuous and leaky.⁴⁹ In our analysis, the second principal component of the DSC signal (PC2) is inversely related to the magnitude of the signal drop, in relation to the baseline. Our results indicate a relatively lower PC2 in TP regions, which may be indicative of a higher degree of blood–brain barrier compromise and leaky neovasculature. We also observed PC3, which reflects the steepness of the complete perfusion signal drop and its recovery rate. We found TP regions to show relatively higher values

of PC3 that may suggest a relative time delay in the contrast agent reaching the TP tissue, possibly due to higher flow resistance, tortuosity, and other characteristics of tumor vasculature.^{29,38,43,50} While these proposed biological associations are limited by the macroscopic nature of MRI, it must be pointed out that one does not require understanding of the mechanism to develop an effective signature—it merely requires rigorous validation to have potential clinical use.

The limitations of our study include the fact that it relied on data from a single institution and could benefit from validation in multi-institutional data. Furthermore, the discovered signature relies upon features extracted from advanced mpMRI, which may not be routinely acquired in all clinical departments. Sample size is also a potential limitation, as the strength of deep learning methods is often improved as the number of subjects increases. The limited number of patients relative to the number of features used in deep learning methods, in particular, may increase the risk of overfitting. We addressed this potential pitfall by cross-validation of all steps when using APS features (ie, feature selection, SVM and SVR parameters selections, and training and testing on different patients). Multi-institutional, prospective validation of our signature is necessary to establish interinstitutional reproducibility.

In conclusion, advanced computational analyses are increasingly used in the clinical evaluation of human gliomas and their response to treatment. The present study extracts subtle but informative QIP features from the temporal dynamics of DSC-MRI, DTI, and structural MRI modalities and integrates them via multivariate machine learning to develop an imaging signature that may noninvasively distinguish TP from PsP. Accurate stratification of the entities of TP and PsP may facilitate appropriate triage of patients to continuing maintenance, therapy, or evaluating them for new intervention, which carries great importance in the era of increasing personalization of therapy.

FUNDING SUPPORT

This study was supported by National Institutes of Health grants on “Predicting brain tumor progression via multiparametric image analysis and modeling” (R01-NS042645) and “Cancer imaging phenomics software suite: application to brain and breast cancer” (U24-CA189523) and a grant by the State of Pennsylvania to the University of Pennsylvania’s Abramson Cancer Center.

CONFLICT OF INTEREST DISCLOSURES

Spyridon Bakas has received research grants from the National Institute of Neurological Disorders and Stroke and the National Cancer Institute Informatics Technology for Cancer Research. Adam P. Dicker has received

personal fees from Roche, Varian, EMD Serono, Janssen, Cybrexa, Oncohost, ThirdBridge, Noxopharm, and Self Care Catalysts. Donald M. O’Rourke has received personal fees from Soma Therapeutics and the Wistar Institute, holds US patent US7625558 and multiple provisional patents on CART cells, and has received a Novartis grant for CART cells. Suyash Mohan has received research grants from NovoCure, the Penn Center for Precision Medicine, and Galileo CDS. The other authors made no disclosures.

AUTHOR CONTRIBUTIONS

Hamed Akbari: conceptualization, methodology, MATLAB programming, data interpretation, validation, analysis, data collection, writing and editing. **Saima Rathore:** CaPTk programming, data interpretation, validation, analysis, writing and editing. **Spyridon Bakas:** data interpretation, validation, data collection, writing and editing. **MacLean P. Nasrallah:** data interpretation, neuropathology, validation, data collection, writing and editing. **Gaurav Shukla:** data interpretation, radiation oncology, validation, data collection, writing and editing. **Elizabeth Mamourian:** data interpretation, validation, data collection, writing and editing. **Martin Rozycki:** data interpretation, validation, data collection, writing and editing. **Stephen J. Bagley:** data interpretation, validation, writing and editing. **Jeffrey D. Rudie:** data interpretation, validation, writing and editing. **Adam E. Flanders:** data interpretation, validation, data collection, writing and editing. **Adam P. Dicker:** data interpretation, validation, data collection, writing and editing. **Arati S. Desai:** data interpretation, validation, writing and editing. **Donald M. O’Rourke:** data interpretation, neurosurgery, validation, writing and editing. **Steven Brem:** data interpretation, neurosurgery, validation, writing and editing. **Robert Lustig:** data interpretation, radiation oncology, validation, writing and editing. **Suyash Mohan:** data interpretation, neuroradiology, validation, writing and editing. **Ronald L. Wolf:** data interpretation, neuroradiology, validation, writing and editing. **Michel Bilello:** data interpretation, neuroradiology, validation, writing and editing. **Maria Martinez-Lage:** conceptualization, methodology, data interpretation, neuropathology, validation, data collection, writing and editing, supervision. **Christos Davatzikos:** conceptualization, methodology, data interpretation, validation, analysis, writing and editing, supervision, funding acquisition.

REFERENCES

- Ostrom QT, Gittleman H, Liao P, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2010–2014. *Neuro Oncol.* 2017;19(suppl 5):v1-v88.
- Stupp R, Taillibert S, Kanner A, et al. Effect of tumor-treating fields plus maintenance temozolomide vs maintenance temozolomide alone on survival in patients with glioblastoma: a randomized clinical trial. *JAMA.* 2017;318:2306–2316.
- Stupp R, Hegi ME, Mason WP, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol.* 2009;10:459–466.
- Gilbert MR, Dignam JJ, Armstrong TS, et al. A randomized trial of bevacizumab for newly diagnosed glioblastoma. *N Engl J Med.* 2014;370:699–708.
- Macdonald DR, Cascino TL, Schold SC Jr, Cairncross JG. Response criteria for phase II studies of supratentorial malignant glioma. *J Clin Oncol.* 1990;8:1277–1280.
- Taal W, Brandsma D, de Bruin HG, et al. Incidence of early pseudo-progression in a cohort of malignant glioma patients treated with chemoradiation with temozolomide. *Cancer.* 2008;113:405–410.
- O’Brien BJ, Colen RR. Post-treatment imaging changes in primary brain tumors. *Curr Oncol Rep.* 2014;16:397.
- Chamberlain MC, Glantz MJ, Chalmers L, Van Horn A, Sloan AE. Early necrosis following concurrent Temodar and radiotherapy in patients with glioblastoma. *J Neurooncol.* 2007;82:81–83.
- Rowe LS, Butman JA, Mackey M, et al. Differentiating pseudoprogression from true progression: analysis of radiographic, biologic, and clinical clues in GBM. *J Neurooncol.* 2018;1–8.

10. Young R, Gupta A, Shah A, et al. Potential utility of conventional MRI signs in diagnosing pseudoprogression in glioblastoma. *Neurology*. 2011;76:1918-1924.
11. 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. *Am J Neuroradiol*. 2016;37:28-36.
12. Verma N, Cowperthwaite MC, Burnett MG, Markey MK. Differentiating tumor recurrence from treatment necrosis: a review of neuro-oncologic imaging strategies. *Neuro Oncol*. 2013;15:515-534.
13. Ismail M, Hill V, Stasevych V, et al. Shape features of the lesion habitat to differentiate brain tumor progression from pseudoprogression on routine multiparametric MRI: a multisite study. *Am J Neuroradiol*. 2018;39:2187-2193.
14. Prasanna P, Rogers L, Lam T, et al. Disorder in pixel-level edge directions on T1WI is associated with the degree of radiation necrosis in primary and metastatic brain tumors: preliminary findings. *Am J Neuroradiol*. 2019.
15. Wen PY, Macdonald DR, Reardon DA, et al. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *J Clin Oncol*. 2010;28:1963-1972.
16. Kickingereder P, Isensee F, Tursunova I, et al. Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study. *Lancet Oncol*. 2019;20:728-740.
17. Galbán CJ, Chenevert TL, Meyer CR, et al. Prospective analysis of parametric response map-derived MRI biomarkers: identification of early and distinct glioma response patterns not predicted by standard radiographic assessment. *Clin Cancer Res*. 2011;17:4751-4760.
18. Gevaert O, Mitchell LA, Achrol AS, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features. *Radiology*. 2014;273:168-174.
19. Ellingson BM. Radiogenomics and imaging phenotypes in glioblastoma: novel observations and correlation with molecular characteristics. *Curr Neurol Neurosci Rep*. 2015;15:1-12.
20. Zhang B, Chang K, Ramkissoon S, et al. Multimodal MRI features predict isocitrate dehydrogenase genotype in high-grade gliomas. *Neuro Oncol*. 2017;19:109-117.
21. Akbari H, Bakas S, Pisapia JM, et al. In vivo evaluation of EGFRvIII mutation in primary glioblastoma patients via complex multiparametric MRI signature. *Neuro Oncol*. 2018;20:1068-1079.
22. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. *NeuroImage*. 2012;62:782-790.
23. Smith SM, Brady JM. SUSAN—a new approach to low level image processing. *Int J Comput Vis*. 1997;23:45-78.
24. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*. 1998;17:87-97.
25. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002;17:143-155.
26. Soares J, Marques P, Alves V, Sousa N. A hitchhiker's guide to diffusion tensor imaging. *Front Neurosci*. 2013;7:31.
27. Boxerman J, Schmainda K, Weisskoff R. Relative cerebral blood volume maps corrected for contrast agent extravasation significantly correlate with glioma tumor grade, whereas uncorrected maps do not. *Am J Neuroradiol*. 2006;27:859-867.
28. Cha S, Lupo J, Chen M-H, et al. Differentiation of glioblastoma multiforme and single brain metastasis by peak height and percentage of signal intensity recovery derived from dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging. *Am J Neuroradiol*. 2007;28:1078-1084.
29. Akbari H, Macyszyn L, Da X, et al. Pattern analysis of dynamic susceptibility contrast-enhanced MR imaging demonstrates peritumoral tissue heterogeneity. *Radiology*. 2014;273:502-510.
30. Ou Y, Akbari H, Bilello M, Da X, Davatzikos C. Comparative evaluation of registration algorithms in different brain databases with varying difficulty: results and insights. *IEEE Trans Med Imaging*. 2014;33:2039-2065.
31. Chatfield K, Simonyan K, Vedaldi A, Zisserman A. Return of the devil in the details: delving deep into convolutional nets [preprint]. 2014; arXiv:14053531.
32. Vedaldi A, Fulkerson B. VLFeat: an open and portable library of computer vision algorithms. Paper presented at: Proceedings of the 18th ACM International Conference on Multimedia; October 2010; Firenze, Italy.
33. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [preprint]. 2014; arXiv:14091556.
34. Yousefi B, Kalhor D, Usamentiaga Fernández R, Lei L, Castanedo CI, Maldague XP. Application of deep learning in infrared non-destructive testing. Paper presented at: Proceedings of QIRT 2018, Berlin, Germany, June 2018:1-9.
35. Davatzikos C, Rathore S, Bakas S, et al. Cancer imaging phenomics toolkit: quantitative imaging analytics for precision diagnostics and predictive modeling of clinical outcome. *J Med Imaging (Bellingham)*. 2018;5:011018.
36. Haralick RM, Shanmugam K. Textural features for image classification. *IEEE Trans Syst Man Cybern Syst*. 1973;6:610-621.
37. Galloway MM. Texture analysis using grey level run lengths. NASA STI/Recon Technical Report. 1974;75.
38. 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;78:572-580.
39. Kurki T, Lundbom N, Valtonen S. Tissue characterisation of intracranial tumours: the value of magnetisation transfer and conventional MRI. *Neuroradiology*. 1995;37:515-521.
40. Lu S, Ahn D, Johnson G, Cha S. Peritumoral diffusion tensor imaging of high-grade gliomas and metastatic brain tumors. *Am J Neuroradiol*. 2003;24:937-941.
41. Wintermark M, Sesay M, Barbier E, et al. Comparative overview of brain perfusion imaging techniques. *J Neuroradiol*. 2005;32:294-314.
42. Lev MH, Hochberg F. Perfusion magnetic resonance imaging to assess brain tumor responses to new therapies. *Cancer Control*. 1998;5:115-123.
43. Bullitt E, Zeng D, Gerig G, et al. Vessel tortuosity and brain tumor malignancy: a blinded study. *Acad Radiol*. 2005;12:1232-1240.
44. Hicklin DJ, Ellis LM. Role of the vascular endothelial growth factor pathway in tumor growth and angiogenesis. *J Clin Oncol*. 2005;23:1011-1027.
45. Essock-Burns E, Lupo JM, Cha S, et al. Assessment of perfusion MRI-derived parameters in evaluating and predicting response to antiangiogenic therapy in patients with newly diagnosed glioblastoma. *Neuro Oncol*. 2011;13:119-131.
46. Jensen RL, Mumert ML, Gillespie DL, Kinney AY, Schabel MC, Salzman KL. Preoperative dynamic contrast-enhanced MRI correlates with molecular markers of hypoxia and vascularity in specific areas of intratumoral microenvironment and is predictive of patient outcome. *Neuro Oncol*. 2014;16:280-291.
47. McDonald DM, Choyke PL. Imaging of angiogenesis: from microscope to clinic. *Nat Med*. 2003;9:713-725.
48. Swami M. Cancer: enhancing EGFR targeting. *Nat Med*. 2013;19:682-682.
49. Thompson G, Mills S, Coope D, O'Connor J, Jackson A. Imaging biomarkers of angiogenesis and the microvascular environment in cerebral tumours. *Br J Radiol*. 2011;84:S127-S144.
50. Jain RK, Di Tomaso E, Duda DG, Loeffler JS, Sorensen AG, Batchelor TT. Angiogenesis in brain tumours. *Nat Rev Neurosci*. 2007;8:610-622.