






ARTICLE

<https://doi.org/10.1038/s41467-019-11007-0>

OPEN

# Multicenter study demonstrates radiomic features derived from magnetic resonance perfusion images identify pseudoprogression in glioblastoma

Nabil Elshafeey<sup>1</sup>, Aikaterini Kotrotsou <sup>1,2</sup>, Ahmed Hassan<sup>1</sup>, Nancy Elshafei<sup>2,3</sup>, Islam Hassan<sup>2</sup>, Sara Ahmed<sup>2</sup>, Srishti Abrol<sup>1</sup>, Anand Agarwal<sup>1</sup>, Kamel El Salek<sup>1</sup>, Samuel Bergamaschi<sup>4</sup>, Jay Acharya <sup>4</sup>, Fanny E. Moron<sup>5</sup>, Meng Law<sup>4,6</sup>, Gregory N. Fuller <sup>7</sup>, Jason T. Huse<sup>7</sup>, Pascal O. Zinn <sup>8,9,10</sup> & Rivka R. Colen <sup>1,2,10,11</sup>

Pseudoprogression (PsP) is a diagnostic clinical dilemma in cancer. In this study, we retrospectively analyse glioblastoma patients, and using their dynamic susceptibility contrast and dynamic contrast-enhanced perfusion MRI images we build a classifier using radiomic features obtained from both Ktrans and rCBV maps coupled with support vector machines. We achieve an accuracy of 90.82% (area under the curve (AUC) = 89.10%, sensitivity = 91.36%, 67 specificity = 88.24%,  $p = 0.017$ ) in differentiating between pseudoprogression (PsP) and progressive disease (PD). The diagnostic performances of the models built using radiomic features from Ktrans and rCBV separately were equally high (Ktrans: AUC = 94%, 69  $p = 0.012$ ; rCBV: AUC = 89.8%,  $p = 0.004$ ). Thus, this MR perfusion-based radiomic model demonstrates high accuracy, sensitivity and specificity in discriminating PsP from PD, thus provides a reliable alternative for noninvasive identification of PsP versus PD at the time of clinical/radiologic question. This study also illustrates the successful application of radiomic analysis as an advanced processing step on different MR perfusion maps.

<sup>1</sup> Department of Diagnostic Radiology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. <sup>2</sup> Department of Cancer Systems Imaging, The University of Texas MD Anderson Cancer Center, Houston, TX 77054, USA. <sup>3</sup> Department of Restorative and Dental Materials, National Research Centre, Cairo 12622, Egypt. <sup>4</sup> Department of Radiology, University of Southern California, Keck School of Medicine, Los Angeles, CA 90033, USA. <sup>5</sup> Department of Radiology, Baylor College of Medicine, Houston, TX 77030, USA. <sup>6</sup> Alfred Health & Monash University, Melbourne 3004, Australia. <sup>7</sup> Department of Pathology, Anatomical and Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. <sup>8</sup> Department of Neurosurgery, Baylor College of Medicine, Houston, TX 77030, USA. <sup>9</sup> Department of Neurological Surgery, University of Pittsburgh, Pittsburgh, PA 15213, USA. <sup>10</sup> UPMC Hillman Cancer Center, Pittsburgh, PA 15232, USA. <sup>11</sup> Department of Radiology, University of Pittsburgh, Pittsburgh, PA 15213, USA. Correspondence and requests for materials should be addressed to P.O.Z. (email: [pascal.zinn@gmail.com](mailto:pascal.zinn@gmail.com)) or to R.R.C. (email: [rccolen@gmail.com](mailto:rccolen@gmail.com))

The current standard of treatment for glioblastoma is the combination of maximal safe resection, radiation, and chemotherapy; this paradigm was shown to prolong the median overall survival to 14.6 months<sup>1,2</sup>. Immunotherapeutic agents, which utilize the body's innate immune responses to kill cancerous cells, have demonstrated success in preclinical trials<sup>3–5</sup>. However, during treatment, the size of the tumor often increases and/or new inflammatory lesions appear. These transient changes typically stabilize or subside without further treatment<sup>6</sup>, but they are often difficult to distinguish from progressive disease (PD). This PD-mimicking phenomenon is called pseudoprogression (PsP). Patients with PsP have longer overall survival and are less likely to exhibit signs and symptoms of neurological deterioration than those with PD<sup>7,8</sup>.

Early discrimination of PsP from PD is a clinical challenge. Surgical biopsy is used in current clinical practice as a standard procedure for diagnosis of recurrent or residual disease, but biopsy is not only highly invasive but also limited in accuracy depending on the biopsy site, type of resection, and lesion heterogeneity<sup>9</sup>. Furthermore, biopsy is not feasible in every glioblastoma patient who has enhancement on post-treatment imaging<sup>9</sup>.

Conventional magnetic resonance imaging (MRI) and the Response Assessment in Neuro-Oncology (RANO) criteria are used as an alternative to surgical biopsy for distinguishing PsP from PD, but the diagnostic performance of this evaluation varies considerably<sup>10</sup>. By RANO criteria, patients with clinical deterioration, a new lesion, or an increase in existing tumor size ( $\geq 25\%$  increase in the sum of products of perpendicular diameters) on restaging MR scans are defined as having PD, while those with an increase  $< 25\%$  or a decrease of  $< 50\%$  in the sum of products of perpendicular diameters of enhancing lesion are considered to have stable disease, as shown in Supplementary Table 1<sup>10</sup>. Immunotherapy RANO (iRANO) criteria were devised as an update to RANO to monitor patients undergoing immunotherapy<sup>11</sup>. According to the iRANO criteria, patients with evidence of PD (per RANO) within 6 months of immunotherapy who do not develop considerable neurological decline should continue on the current therapy and undergo follow-up imaging in 3 months for confirmation of PD<sup>11</sup>. The drawback of this watchful waiting is that patients with PD will continue on an ineffective therapy, incurring the unnecessary risks of potential toxicity and delay in switching to more effective therapy; furthermore, failure to correctly identify PsP may result in premature discontinuation of an effective treatment<sup>10,11</sup>.

Recently, there have been numerous efforts using MR perfusion to differentiate PD from PsP<sup>12–14</sup>. MR perfusion, which can assess the vascular properties of the post-treatment enhancing lesions, calculates changes in blood volume, blood flow, and vessel wall permeability. These are important characteristics of tumor vessels<sup>15,16</sup>. However, lack of standardization in the post-processing steps of MR perfusion studies has led to a discrepancy in the reported cut-off values determining post-treatment changes from PD and corresponding sensitivity and specificity<sup>17,18</sup>. Therefore, there is a need for an accurate noninvasive quantitative tool to differentiate between PD and pseudoprogression at the time of detection of a brain MRI questionable lesion.

Radiomic analysis is a new automated, high-throughput method that quantifies the tumor phenotype at a microscale level (voxel/pixel level) by using thousands of image-based features obtained by histogram and texture analysis<sup>19–21</sup>. Unlike surgical biopsy, radiomic analysis assesses the entire three-dimensional tumor inclusive of spatial heterogeneity<sup>22,23</sup>. Prior studies have employed radiomic analysis on MR perfusion for predicting the underlying glioblastoma molecular phenotype<sup>20,24–26</sup> and patient outcomes<sup>24,25</sup> as well as distinguishing benign and malignant breast lesions<sup>27</sup>.

However, to date, no multicenter investigations on the use of MR perfusion-based radiomic analysis to distinguish histologically-proven PsP from PD in glioblastoma have been reported.

In this multicenter study, we seek to determine the ability of MR perfusion-based radiomics to discriminate PsP from PD in glioblastoma patients. We evaluated 98 patients with pathologically-proven glioblastoma to identify the MR perfusion-based radiomic signatures of PsP and PD; we then combined these signatures to develop a noninvasive predictive model to robustly differentiate PsP from PD in the clinical setting. With an inclusive and diverse multi-institutional cohort, this pioneering study seeks to distinguish PD from PsP through MR perfusion-based radiomic analysis.

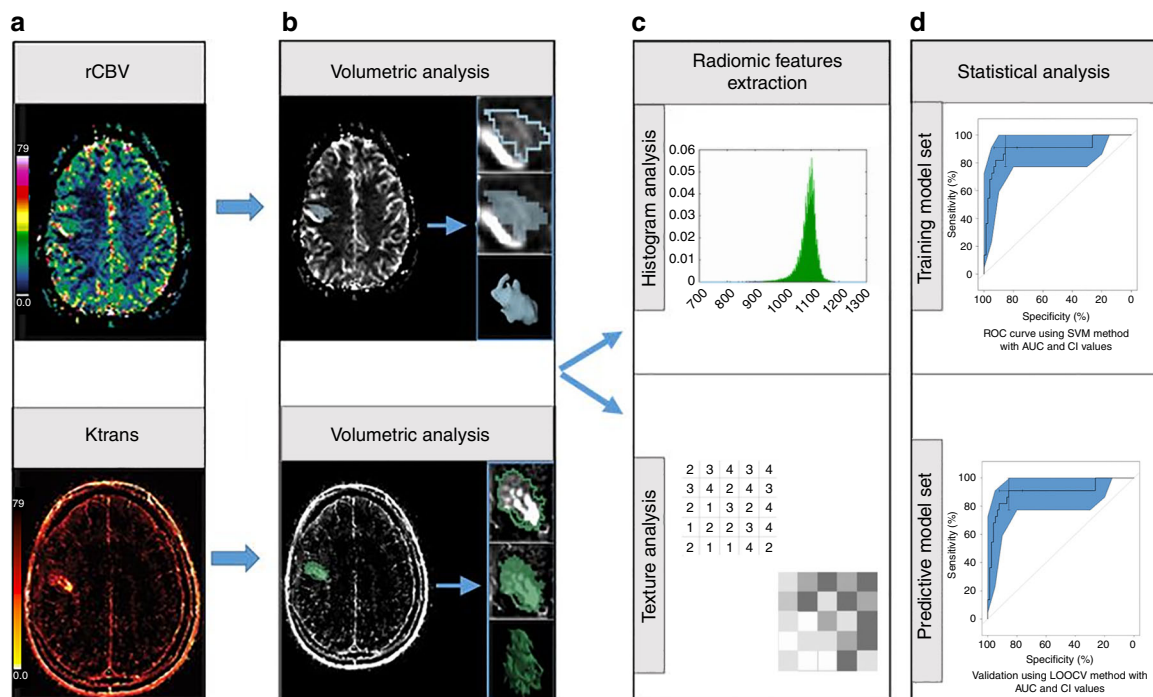
## Results

**Patient characteristics.** We investigated a total of 98 patients from 3 institutions who had histopathologic evidence of PD ( $n = 76$ ; 77.6%) or PsP ( $n = 22$ ; 22.4%; Table 1). Using Mann-Whitney test for non-parametric data and independent sample *t* test for parametric data, there were no statistical significant difference between the two groups in terms of age, sex, or  $K_{trans}$  volume ( $p > 0.05$ ). However, the extracted VOI of rCBV using 3D slicer were differed significantly between PsP and PD groups (1126.91 mm<sup>3</sup> and 1369.44 mm<sup>3</sup>, respectively;  $p = 0.010$ ).

**Predictive value of perfusion parameters analysis.** Perfusion parameter analysis showed that neither  $K_{trans}$  nor rCBV was able to predict PsP from PD groups, and there was no statistically

**Table 1** Baseline selected demographic and clinical characteristics of patients with Pseudo progression/ Glioblastoma grade IV (N = 98)

Characteristic	Pseudoprogression N = 22	Glioblastoma grade IV N = 76
Age, years (SD)	53.50 (14.88)	49.37 (12.97)
Sex, male, N (%)	15 (68.2%)	42 (54.7%)
$K_{trans}$ volume, mm <sup>3</sup> (SD)	3166.99 (4618.17)	4591.56 (6233.76)
rCBV volume, mm <sup>3</sup> (SD)	1126.91 (2771.41)	1396.44 (2245.20)
Surgical type:		
Total resection, N (%)	19(86.4%)	45(59.2%)
Sub-total resection, N (%)	3(13.6%)	29(38.2%)
Biopsy, N (%)	0(0%)	2(2.6%)
Molecular status:		
MGMT		
Methylated, N (%)	3(13.7%)	3(3.9%)
Unmethylated, N (%)	1(4.5%)	10(13.2%)
Non tested, N (%)	18(81.8%)	63(82.9%)
IDH		
Positive, N (%)	2(9%)	7(9.2%)
Negative, N (%)	5(22.7%)	16(21%)
Non tested, N (%)	15(68.3%)	53(69.8%)
Radio-therapy time, days (SD)**	54.21429(36.59)	56(22.24)
Time after RT to PD/PSP, days(SD)***	779.92(766.59)	664.9857 (854.85)
Chemotherapy treatment (Temozolomide), N (%)	22 (100%)	76 (100%)
rCBV relative cerebral blood volume, SD standard deviation *, Significant difference **, 5 PSP and 8 PD with no available data ***, 2 PSP and 5 PD with no available data		



**Fig. 1** Image post-processing radiomic workflow. **a** relative cerebral blood volume (rCBV) and  $K_{trans}$  maps of perfusion MRI are acquired. **b** Segmentation of the region of interest using 3D slicer software. **c** Radiomic feature extraction from the whole tumor volume. **d** Statistical analysis: radiomic and clinical features are analyzed to determine their diagnostic and predictive values

significant difference between the PsP and PD groups (Mann-Whitney test for non-parametric data; Supplementary Table 4).

**Performance of the radiomic model.** To determine the capacity of radiomic features to distinguish PsP from PD, we performed an integrative analysis assessing the predictive performance of perfusion radiomic features. The image post-processing workflow is shown in Fig. 1. We first ranked features based on their relevance to the outcome and within-feature redundancy using the MRMR feature selection technique. Subsequently, SVM with linear kernel and C5.0 models were constructed using the features selected by the MRMR analysis. Finally, the diagnostic performance of the models was evaluated using LOOCV for the SVM model and 10-fold cross-validation for the C5.0 model.

In the analysis using  $K_{trans}$  maps, the top 60 features (as ranked by the MRMR feature selection technique) robustly differentiate between PsP and PD. As shown in Fig. 2, selected radiomic features significantly discriminated between PsP and PD using the C5.0 method (AUC 100%, sensitivity 100%, specificity 100%,  $p$ -value  $1.512e-11$  one sided binomial test). Similar results were obtained using the SVM method (AUC 89.8%, sensitivity 92.5%, specificity 88.89%,  $p$ -value 0.003744 one sided binomial test) (Fig. 2). We then validated the predictive models using LOOCV (AUC 90%, sensitivity 93%, specificity 89%,  $p$ -value 0.004 one sided binomial test, Fig. 2) and 10-fold cross-validation (AUC 100%, sensitivity 100%, specificity 100%,  $p$ -value  $1.512e-11$  one sided binomial test, Fig. 2). Significant features used in the models are shown in Supplementary Table 5.

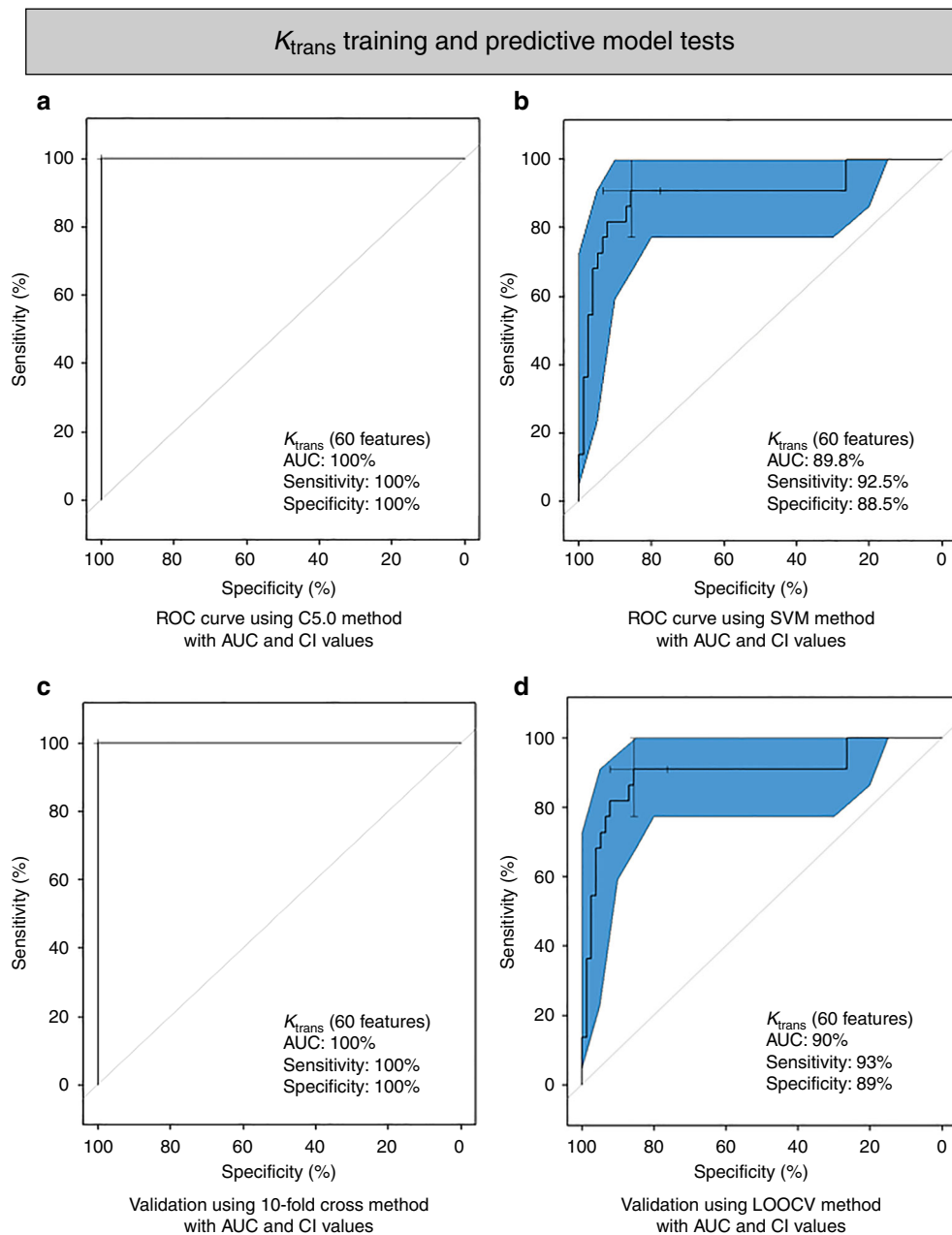
In the analysis using rCBV maps, the top 160 features (as identified by the MRMR feature selection technique) achieved the highest predictive accuracy in differentiating between PsP and PD. Significant features used in the models are shown in Supplementary Table 6. Radiomic models built using SVM significantly distinguished PsP and PD (AUC 94%, sensitivity 92%, specificity 100%,  $p$ -value 0.012 one-sided binomial test); similar results were obtained using the C5.0 classification (AUC

100%, sensitivity 100%, specificity 100%,  $p$ -value  $1.512e-11$  one-sided binomial test). Validation using LOOCV (AUC 94%, sensitivity 92%, specificity 100%,  $p$ -value 0.012 one-sided binomial test, Fig. 3) and 10-fold cross-validation (AUC 100%, sensitivity 100%, specificity 100%,  $p$ -value  $1.512e-11$  one-sided binomial test, Fig. 3) demonstrated to be highly statistically significant.

To determine the added value of combining the radiomic features obtained from both  $K_{trans}$  and rCBV pharmacokinetic maps in predicting PsP from PD, we first ranked all 620 features identified by the MRMR technique and built further predictive models. Using 60 radiomic features (Supplementary Table 7) coupled with SVM, we achieved an accuracy of 90.82% (AUC 89.10%, sensitivity 91.36%, specificity 88.24%,  $p$ -value 0.017 one-sided binomial test); similar results were obtained using C5.0 (AUC 100%, sensitivity 100%, specificity 100%,  $p$ -value  $1.512e-11$  one-sided binomial test). Differences identified by LOOCV predictive modeling (AUC 89%, sensitivity 91.4%, specificity 88%,  $p$ -value 0.02 one-sided binomial test, Fig. 4) and 10-fold cross-validation (AUC 100%, sensitivity 100%, specificity 100%,  $p$ -value  $1.512e-11$  one-sided binomial test, Fig. 4) were both statistically significant.

As the image acquisition parameters were distinct between different institutions, we sought to determine the scalability and generalizability of our radiomics analytical technique. Towards this end, we performed sub-analysis on the 63 MDACC patients, we achieved a high accuracy, sensitivity, and specificity to differentiate between PsP and PD using the three different extracted radiomic models obtained from  $K_{trans}$ , rCBV and combined pharmacokinetic maps as we demonstrated in Supplementary Figs 1–3.

To assess the predictive power of the extracted radiomic features obtained from  $K_{trans}$ , rCBV, and combined pharmacokinetic maps, we collated an independent prospective dataset of seven pathologically proven MDACC patients (Supplementary Table 2). Using the model built with whole dataset on the 98



**Fig. 2** Model building and evaluation using the selected  $K_{trans}$  features (60 features). **a, b** ROC curve depicts the predictive model building using C5.0 ( $P$ -value  $1.512e-11$ ) and SVM methods ( $P$ -value 0.003744) respectively. **c, d** 10-fold cross-validation ROC curve ( $P$ -value  $1.512e-11$ ) and Leave-One-Out Cross-Validation (LOOCV) ROC curve ( $P$ -value 0.004) depicts the performance of the model

patients as well as models built with cross-validation technique (LOOCV and 10-fold CV) for both C5.0 and SVM methods, the prediction results showed that the 7 new prospective patients were correctly classified as PD patients with high probability percentage as mentioned in Supplementary Tables 8–10.

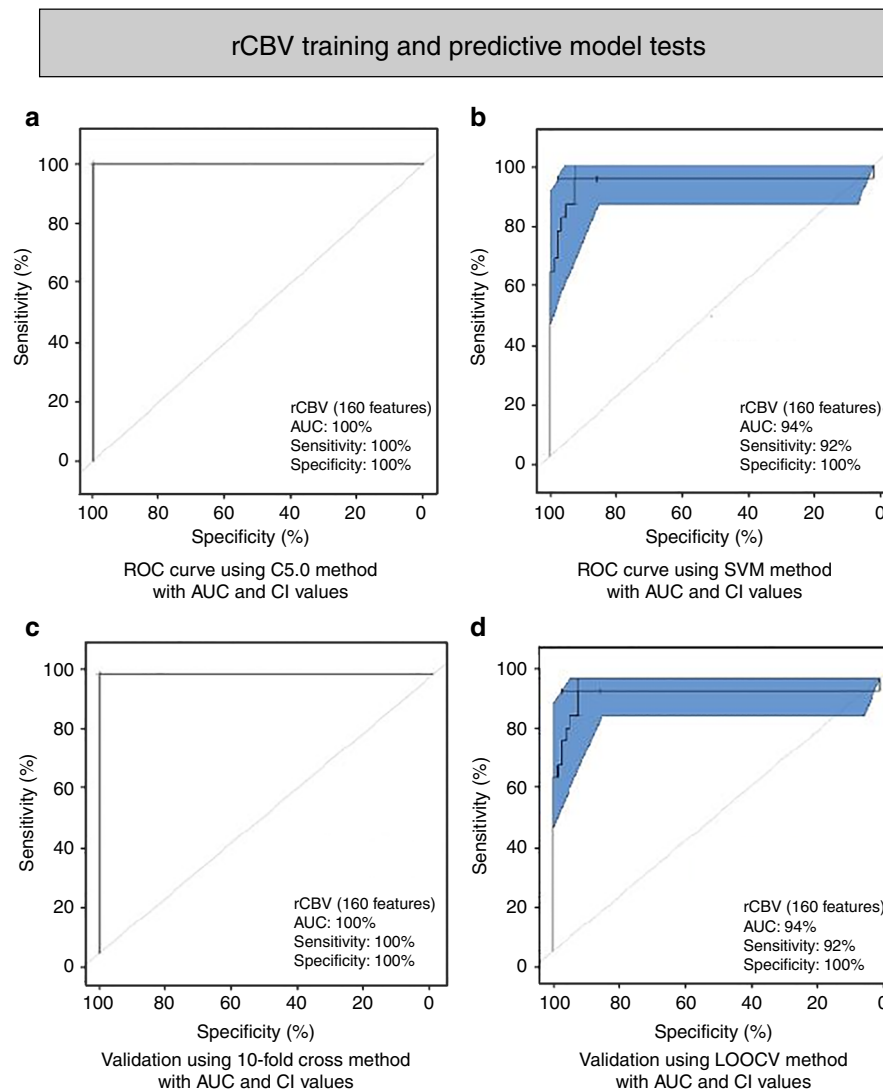
## Discussion

In this study, we evaluated the ability of MR perfusion-based radiomics to non-invasively discriminate between PsP and PD in glioblastoma patients at the time of clinical/radiologic question. We demonstrated that, although PsP and PD have similar radiographic appearances, they harbor distinct radiomic information that is hidden within MR perfusion images and can be extracted and used to build a clinically-relevant predictive model that discriminates PsP from PD.

Evaluation of post-treatment changes in glioblastoma patients using conventional MRI remains a challenge for clinicians and radiologist, as post-treatment MR images demonstrate confounding appearances<sup>28–30</sup>. Advanced MRI techniques, including perfusion MRI, have been proposed as alternatives to current imaging assessment criteria; however, reported cut-off values determining post-treatment changes versus PD vary widely across different studies. In the meta-analysis by Wan et al., they demonstrated substantial discrepancy between the cut-off values (see Wan et al.; Table 1)<sup>18</sup>, which they attributed to the differences in the acquisition parameters of DSC and DCE among different institutions. Further, the reported accuracy, sensitivity, and specificity of the perfusion studies are not clinically reliable<sup>17</sup>.

In our patient cohort, ROI-based perfusion parameter analysis using  $K_{trans}$  and rCBV pharmacokinetic maps alone did not





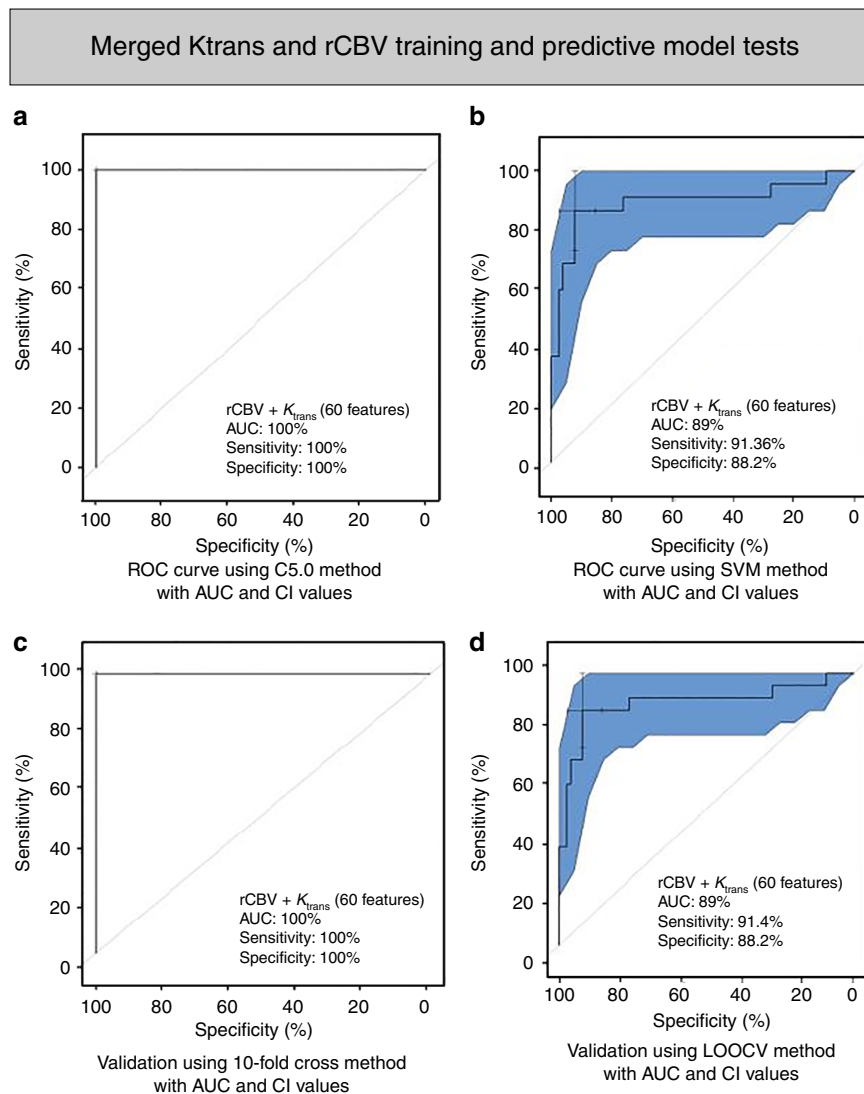
**Fig. 3** Model building and evaluation using the selected rCBV features (160 features). **a, b** ROC curve depicts the predictive model building using C5.0 ( $P$ -value  $1.512e-11$ ) and SVM methods ( $P$ -value 0.012) respectively. **c, d** 10-fold cross-validation ROC curve ( $P$ -value  $1.512e-11$ ) and Leave-One-Out Cross-Validation (LOOCV) ROC curve ( $P$ -value 0.012) depicts the performance of the model

detect any differences between PsP and PD patient groups (Supplementary Table 4). However, our proposed radiomic model achieved high accuracy, sensitivity, and specificity in differentiating between PsP from PD compared to previous methods at the time of the clinical question of PsP versus PD; therefore, radiomics can be used by radiologists and oncologists for clinical decision-making (Figs. 2–4).

The radiomic model used in this study and proposed for ongoing clinical applications was built using imaging data from a multi-institutional cohort of patients for whom histopathologic confirmation of PsP or PD was available. The model demonstrated high discriminatory power, with AUCs greater than 89% (Figs. 2–4). Despite differences in the MR acquisition parameters of DSC and DCE imaging among the institutions (Supplementary Table 3), the radiomics predictive model demonstrated robust performance. Additionally, the discriminatory power was highly robust for the individual pharmacokinetic maps and after fusing the features from both the  $K_{trans}$  and rCBV maps. These results suggest that our radiomic model is well-calibrated and can be directly applied in any clinical setting without the need to adjust the acquisition protocol or include both DSC and DCE sequences.

Further, the majority of the features included in our models are second-order (Supplementary Tables 5–7); of the key features selected (60 for  $K_{trans}$ , 160 for rCBV and 60 for fused model), the number of second-order features was 54 (90%) for the  $K_{trans}$ -based model (Supplementary Table 5), 150 (94%) for the rCBV-based model (Supplementary Table 6), and 54 (90%) for the fused model (Supplementary Table 7). This finding is concordant with the fact that histogram features depict the data distribution within a selected region, and therefore their value is directly affected by the differences in acquisition parameters. Conversely, second-order features capture the inter-relationship between neighboring voxels, and thus are not sensitive to the absolute data value. Hence, by incorporating second-order radiomic features, the model is more robust and ‘immune’ to the MR acquisition parameters.

The robustness of our extracted radiomic features is further demonstrated by the similarity in calculated AUC, sensitivity, and specificity for the two classification algorithms investigated: (i) SVM with linear kernel and (ii) C5.0. A previous study investigated the performance and stability of various feature selection and classifier methods in radiomic applications<sup>31</sup>. They reported



**Fig. 4** Model building and evaluation using the selected merged  $K_{trans}$  and rCBV features (60 features). **a, b** ROC curve depicts the predictive model building using C5.0 ( $P$ -value  $1.512e-11$ ) and SVM methods ( $P$ -value 0.017), respectively. **c, d** 10-fold cross-validation ROC curve ( $P$ -value  $1.512e-11$ ) and Leave-One-Out Cross-Validation (LOOCV) ROC curve ( $P$ -value 0.02) depicts the performance of the model

variable accuracy for different combinations of feature selection and classifier methods. Although evaluation of classifier methods and feature selection techniques was beyond the scope of this work, our findings indicate that our extracted radiomic features were stable and performed equally well independent of the classifier method<sup>31</sup>.

Further, the key features selected for both  $K_{trans}$ - and rCBV-based models were entropy, sum of squares, and autocorrelation (Supplementary Tables 5–7). These features hold information about the spatial heterogeneity of the studied region and are more sensitive when assessing tumor microenvironment and the presence of tumor cells. The findings of the present study confirmed and extended previous work by Zinn et al. who showed that MRI features (Diffusion and radiomic features) capture spatial heterogeneity<sup>25,32,33</sup>. Our results show that patients with PD have high entropy and sum of squares and low autocorrelation compared to PsP patients. Thus, selected volumes in PD patients were characterized by non-uniformity and increased complexity compared to PsP patients. This may reflect the heterogeneity seen microscopically<sup>34–36</sup>.

Limitations of the current study mostly result from the requirement for histopathologic confirmation of PsP or PD. First,

the percentage of patients with PsP is low, and our study design did not include patients for whom histopathologic confirmation was not available. Because biopsy is an invasive procedure and clinicians usually prefer a watchful approach over biopsy, it is expected that only a small fraction of patient who have undergone biopsy will have PsP. Second, we performed semi-automated segmentation of the VOI, which is a time-consuming process and also requires an in-depth assessment of inter-rater variability and stability of the extracted radiomic features. A user-friendly tool that performs reliable automatic segmentations is needed, as it would allow for easy translation of the radiomic analysis in the clinical setting. Additionally, being a retrospective study, important information, such as molecular and genetic characteristics, was not available to review. For example, the contribution of MGMT methylation and *IDH1/2* mutations, was not investigated in this work, since this information was not available for the majority of the studied population. Finally, the total number of patients included in the study was relatively small and did not allow for a large validation prospective dataset using as a separate cohort. A much larger study with a larger patient cohort that will allow independent validation is needed; further the prospective validation component is currently performed in our institution.

In summary, our study presents a radiomic model based on MR perfusion data for noninvasive, individualized prediction of PsP in glioblastoma patients at the point-of-care (at the time of the clinical question). We demonstrate the high predictive accuracy of radiomics in differentiating PsP from PD. Our radiomic model can be easily integrated into the clinical setting, as it is a post-processing approach that does not require changes in the current imaging protocol and will allow clinicians to make more informed decisions for optimal patient care. After outlining the VOI on DCE and DSC maps, the model will indicate whether the patient has PsP or PD. Our radiomic model will complement the methods now available to clinicians by offering a comprehensive evaluation of the imaging data. We expect that it will dramatically decrease the number of invasive procedures performed to confirm the absence of tumor in patients; Most importantly, our radiomics method can help ensure that patients with PsP will continue on an effective treatment or discontinue an ineffective treatment without the need for delays caused by watchful waiting. Conversely, early identification of PD will allow the prompt transition to a more effective treatment.

## Methods

**Study population.** This retrospective multi-institutional clinical study was conducted in compliance with U.S. Health Insurance Portability and Accountability Act (HIPAA) regulations. All necessary institutional ethics review approvals were obtained by the participating institutions (The University of Texas MD Anderson Cancer Center, Baylor College of Medicine, and University of Southern California Keck School of Medicine).

A total of 651 patients with histopathologic diagnosis of GBM and an enhancing lesion on MRI that may indicate PD or treatment-related changes were received from three institutions; The University of Texas MD Anderson Cancer Center (MDACC) (598 patients), Baylor College of Medicine (BCM) (13 patients), and University of Southern California (USC) Keck School of Medicine (40 patients). A total of 98 patients were included in the final analysis based on the following inclusion criteria: (1) minimum age of 18 years; (2) enhancing lesion on MRI that may indicate PD or treatment-related changes; (3) conventional and advanced MRI, including dynamic susceptibility contrast (DSC) and dynamic contrast-enhanced (DCE) perfusion imaging; (4) pathologic confirmation of PD or PsP within 3 months after detection of a questionable enhancing lesion on MRI (time of question); (5) adequate image quality with no artifact affecting the radiomic analysis. 553 patients did not meet the inclusion criteria and therefore excluded. 8 patients (1 BCM, 7 USC) did not have dynamic susceptibility contrast-enhanced T2\*-weighted (DSC) and dynamic contrast-enhanced (DCE) perfusion; 85 patients (82 MDACC, 2 BCM, and 1 USC) were excluded due to poor quality of imaging/perfusion data; 460 (453 MDACC, 4 USC, and 3 BCM) were excluded due to absence of histopathological confirmation of PsP versus PD.

The clinical, radiologic, and pathologic data were retrieved from the medical records. Our final study cohort consisted of: 63 (50PD and 13PSP) patients from MDACC, 7 (6PD and 1PSP) patients from BCM, and 28 (21PD and 7PSP) patients from USC. Based on the pathologic diagnosis, 76 (77.6%) patients had PD and 22 (22.4%) patients had treatment-related changes considered to be PsP. Detailed demographic, clinical characteristics and treatment information is presented in Table 1.

For further validation of our results, a prospective study of 7 MDACC patients with pathologically proven PD were included using the same aforementioned inclusion and exclusion criteria. (Supplementary Table 2 showed the demographic, clinical and pathological data)

**Histopathologic evaluation.** All patients had undergone biopsy or surgical resection, and resected entire tumor/tissue block specimens were histopathologically examined. The pathology status as PD or PSP was decided and reviewed in consensus by two experienced board-certified pathologists (J.T.H. and G.N.F.) blinded to the MRI data. PD was defined as samples with either solely recurrent GBM tissue or mixture of radiation necrosis and recurrent GBM, while PSP was determined by the presence of radiation necrosis only or mixture between radiation necrosis and reactive gliosis.

**MRI acquisitions and imaging post processing.** For conventional MRI acquisition, images were acquired using a 1.5 or 3.0 Tesla MR Scanner. The MRI protocol for MDACC included an axial T1-weighted sequence (repetition time [TR], 700 ms; echo time [TE], 12 ms; slice thickness, 5 mm; acquisition matrix 352 × 224), an axial fluid attenuation inversion recovery (FLAIR) sequence (TR, 10,000 ms; TE, 140 ms; slice thickness, 5 mm; acquisition matrix, 256 × 256), and an axial post-contrast T1-weighted sequence acquired 5 min after the contrast injection (TR, 750 ms; TE, 13 ms; slice thickness, 5 mm; acquisition matrix, 384 ×

256). The MRI protocol for USC included an axial T1-weighted sequence (TR, 600 ms; TE, 8 ms; slice thickness, 5 mm; acquisition matrix, 320 × 192), an axial FLAIR sequence (TR, 8000 ms; TE, 151.2 ms; slice thickness, 5 mm; acquisition matrix, 320 × 224), and an axial post-contrast T1-weighted sequence acquired 5 min after the contrast injection (TR, 615 ms; TE, 17 ms; slice thickness, 5 mm; acquisition matrix, 320 × 192). The MRI protocol for BCM included an axial T1-weighted sequence (TR, 750 ms; TE, 11 ms; slice thickness, 5 mm; acquisition matrix, 384 × 192), an axial FLAIR sequence (TR, 2541 ms; TE, 16.4 ms; slice thickness, 5 mm; acquisition matrix, 384 × 192), and an axial post-contrast T1-weighted sequence acquired 5 min after the contrast injection (TR, 750 ms; TE, 11 ms; slice thickness, 5 mm; acquisition matrix, 384 × 192). While for advanced MRI acquisition: The parameters of both DCE and DSC for each institution are shown in Supplementary Table 3. Pre-contrast T1 mapping images were acquired with six flip angles (2, 5, 10, 15, 20, and 25 degrees). Images were acquired before and after injection of passively targeted gadolinium-diethylenetriamine penta-acetic acid (0.1 mmol per kg).

The advanced imaging analysis and pharmacokinetic map calculation were performed using Nordic ICE (NordicNeuroLab, Bergen, Norway). Images were segmented using 3D Slicer (version 4.3.1, <https://www.slicer.org>), an open-source software program widely used for image visualization and segmentation<sup>37,38</sup>. The image post-processing workflow is shown in Fig. 1.

DCE data were analyzed by using Nordic ICE, using the dual compartment modified Tofts and Kermode pharmacokinetic model as described elsewhere<sup>39,40</sup>. We manually selected the arterial input function in the middle cerebral artery located on the ipsilateral side, then the pharmacokinetic  $K_{trans}$  parameter (reflects local blood flow, endothelial permeability, endothelial surface area, proportional blood volume within a given voxel) was obtained<sup>41</sup>. For DSC analysis, the arterial input function was manually selected in the middle cerebral artery located on the ipsilateral side. Using the integrated DSC module, which incorporates a relative cerebral blood volume (rCBV) leakage-correction algorithm and manual noise thresholding, we obtained the amount of blood in a given volume of tissue expressed as mL per 100 mL of tissue<sup>41</sup>.

**Perfusion parametric maps analysis.** We performed a perfusion parameter analysis to determine whether perfusion parameters can differentiate between PD and PsP. Both capillary permeability ( $K_{trans}$ ) and leakage-corrected rCBV maps were rigidly registered based on mutual information with the post-contrast T1WI using Nordic ICE; the resulting registrations were visually inspected to ensure adequate alignment. Subsequently, circular regions-of-interest (ROIs) with an area of 10.55 mm<sup>2</sup> were manually delineated on the area with the highest parameter value for each parameter map, ensuring that necrotic areas and large blood vessels were excluded. The mean values were extracted from the ROIs for each parameter map. To reduce sampling error, four ROIs were selected and the average of the mean values was obtained<sup>42</sup>. This method has been demonstrated to provide the optimal interobserver and intra-observer reproducibility<sup>43</sup>.

**Quantitative radiomic analysis and model building.** Post-contrast T1WI were coregistered with  $K_{trans}$  and rCBV maps to ensure appropriate tumor delineation; cystic, necrotic regions and intralesional macrovessels were excluded. The coregistered images for both  $K_{trans}$  and leakage-corrected rCBV maps were semi-automatically segmented by an experienced user (N.E., 4 years experience) and manually reviewed slice-by-slice by an experienced board-certified radiologist (R.R.C., 9 years experience), both of whom were blinded to histopathologic assessment. Three-dimensional volumes of interest (VOIs) were subsequently extracted for further analysis. A total of 310 radiomic features were extracted from each pharmacokinetic parametric map, consisting of 10 histogram-based features<sup>44</sup> and 300 s-order Haralick features<sup>45–47</sup>. Accounting for the  $K_{trans}$  and leakage-corrected rCBV maps analyzed in this study, a total of 620 radiomic features were obtained per patient<sup>26,48,49</sup>. Radiomic analysis was performed using our in-house developed software implemented in MATLAB (version 2017b; MathWorks Inc; Natick, MA).

Histogram-based features describe the  $K_{trans}$  and rCBV values distribution within the VOI; the features obtained were minimum, maximum, mean, standard deviation, skewness, kurtosis, and the percentiles 1%, 5%, 95%, and 99%. The second-order Haralick features were computed from the gray-level co-occurrence matrix (GLCM) after requantization of the image gray levels; the gray levels implemented in this analysis were 8, 16, 32, 64, and 256. The second-order Haralick features describe mathematical relationships between co-occurring voxels separated by a given distance in a specific direction. For the purposes of this analysis, we investigated a distance of 1 voxel and 4 in-plane directions. Twenty GLCM-based features were obtained per gray level: autocorrelation, contrast, correlation, cluster shade, cluster prominence, dissimilarity, energy, entropy, homogeneity, maximum probability, variance, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation 1, information measure of correlation 2, inverse difference moment, and normalized inverse difference moment. Rotation invariant measures of the features were obtained by calculating the average, range, and angular variance of the features across the four in-plane directions. Taking into account the number of gray levels and the number of rotation invariant measures, a total of 300 were calculated per map. Fig. 1 shows the pipeline used to get the radiomic perfusion model.

To investigate the predictive performance of extracted radiomic features in discriminating between PD and PsP (outcome), we first identified the radiomic features (feature set) that are associated with outcome. To minimize irrelevant and redundant radiomic features, we used the Maximum Relevance Minimum Redundancy (MRMR) feature selection technique; MRMR ranks the features based on maximizing the relevance and avoiding excess redundancy<sup>50</sup>. The MRMR approach was implemented using the mRMRe R package.

Different sizes of feature sets were investigated to ensure selection of the feature set size with best prediction accuracy. We started with the MRMR-ranked 310 features and evaluated different feature set sizes ranging from 310 to 20 in steps of 10.

MRMR-ranked radiomic features were used for classification and model building. Two different supervised learning algorithms were used in this analysis for classification and predictive model building: Support Vector Machine (SVM) and decision tree algorithm C5.0. In the C5.0 algorithm, a decision tree is used while modeling the classification process<sup>51</sup>. SVM is a supervised machine learning algorithm that is used for solving classification problems by transforming the feature space to a higher dimension space so that a separating hyperplane maximizes the distance between the two classifiers<sup>52</sup>. To increase model accuracy in our analysis, we tested the classifier in SVM with three different kernels (linear, polynomial, and radial basis)<sup>53</sup>.

We used leave-one-out cross-validation (LOOCV) and k-fold cross-validation (k = 10) to evaluate the accuracy, sensitivity, and specificity of the models<sup>54</sup>. Receiver Operating Characteristic (ROC) curves were plotted; and the area under the curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value, and p-value were reported for each cross-validation and prediction output.

**Statistical analysis.** Statistical analyses were performed using R software, version 3.4.3 (<https://www.r-project.org>)<sup>55</sup>. The statistical significance was two-sided in this study, with significance level <0.05. Prior to univariate analysis, the fitness of numeric dataset to normal distribution was determined by the Kolmogorov–Smirnov test. Differences in non-parametric data, such as demographic and clinical characteristics and individual perfusion parameters (rCBV and  $K_{trans}$ ), between the PD and PsP groups were assessed using the Mann–Whitney U test.

## Data availability

The data is available within the Article or the Supplementary Information. The imaging data that support the findings of this study are available from the corresponding author upon request.

## Code availability

For the MR perfusion analysis, we used Nordic ICE, a commercially available processing platform (<https://www.nordicneurolab.com/products/nordicICE.html>). For volumetric analysis we used 3D Slicer (version 4.3.1, <https://www.slicer.org/>), an open-source software program widely used for image visualization and segmentation. The pipeline for the radiomic analysis has been introduced in Zinn et al.<sup>20</sup>. For the biostatistical analysis, we used R software (version 3.4.0, R Foundation for Statistical Computing, Vienna, Austria); package mRMRe (version 2.0.7) and C5 decision tree (version 0.1.2) and SVM (version 1.7–0.1)<sup>50–52</sup>.

Received: 18 January 2019 Accepted: 7 June 2019

Published online: 18 July 2019

## References

- Holland, E. C. Glioblastoma multiforme: The terminator. *Proc. Natl Acad. Sci. USA* **97**, 6242–6244 (2000).
- Ostrom, Q. T. et al. CBTRUS Statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2009–2013. *Neuro-Oncol.* **18**, v1–v75 (2016).
- Wainwright, D. A. et al. Durable therapeutic efficacy utilizing combinatorial blockade against IDO, CTLA-4, and PD-L1 in mice with brain tumors. *Clin. Cancer Res. : Off. J. Am. Assoc. Cancer Res.* **20**, 5290–5301 (2014).
- Antonios, J. P. et al. Detection of immune responses after immunotherapy in glioblastoma using PET and MRI. *Proc. Natl Acad. Sci. USA* **114**, 10220–10225 (2017).
- Zeng, J. et al. Anti-PD-1 blockade and stereotactic radiation produce long-term survival in mice with intracranial gliomas. *Int. J. Radiat. Oncol., Biol., Phys.* **86**, 343–349 (2013).
- Lin, A. L. et al. Molecular and histologic characteristics of pseudoprogression in diffuse gliomas. *J. Neuro-Oncol.* **130**, 529–533 (2016).
- Chamberlain, M. C., Glantz, M. J., Chalmers, L., Van Horn, A. & Sloan, A. E. Early necrosis following concurrent Temodar and radiotherapy in patients with glioblastoma. *J. Neuro-Oncol.* **82**, 81–83 (2007).
- de Wit, M. C. Y., de Bruin, H. G., Eijkenboom, W., Sillevs Smitt, P. A. E. & van den Bent, M. J. Immediate post-radiotherapy changes in malignant glioma can mimic tumor progression. *Neurology* **63**, 535–537 (2004).
- Hygino da Cruz, L. C., Rodriguez, I., Domingues, R. C., Gasparetto, E. L. & Sorensen, A. G. Pseudoprogression and pseudoresponse: imaging challenges in the assessment of posttreatment glioma. *Am. J. Neuroradiol.* **32**, 1978–1985 (2011).
- Wen, P. Y. et al. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *J. Clin. Oncol.* **28**, 1963–1972 (2010).
- Okada, H. et al. Immunotherapy response assessment in neuro-oncology: a report of the RANO working group. *Lancet. Oncol.* **16**, e534–e542 (2015).
- Young, R. J. et al. MRI perfusion in determining pseudoprogression in patients with glioblastoma. *Clin. Imaging* **37**, 41–49 (2013).
- Prager, A. J. et al. Diffusion and perfusion MRI to differentiate treatment-related changes including pseudoprogression from recurrent tumors in high-grade gliomas with histopathologic evidence. *Am J Neuroradiol.* <https://doi.org/10.3174/ajnr.A4218> (2015).
- Ye, J. et al. Differentiation between recurrent gliomas and radiation necrosis using arterial spin labeling perfusion imaging. *Exp. Ther. Med.* **11**, 2432–2436 (2016).
- Essig, M. et al. Perfusion MRI: the five most frequently asked technical questions. *Ajr. Am. J. Roentgenol.* **200**, 24–34 (2013).
- Law, M. Advanced imaging techniques in brain tumors. *Cancer Imaging* **9**, S4–S9 (2009).
- Huang, R. Y., Neagu, M. R., Reardon, D. A. & Wen, P. Y. Pitfalls in the neuroimaging of glioblastoma in the era of antiangiogenic and immunol/targeted therapy—detecting illusive disease, defining response. *Fron. Neurol* **6**, 33 (2015).
- Wan, B. et al. The diagnostic performance of perfusion MRI for differentiating glioma recurrence from pseudoprogression: A meta-analysis. *Medicine* **96**, e6333 (2017).
- Abrol, S., Kotrotsou, A., Salem, A., Zinn, P. O. & Colen, R. R. Radiomic phenotyping in brain cancer to unravel hidden information in medical images. *Top. Magn. Reson. Imaging : TMRI* **26**, 43–53 (2017).
- Zinn, P. O. et al. A Coclinical radiogenomic validation study: conserved magnetic resonance radiomic appearance of periostin-expressing glioblastoma in patients and xenograft models. *Clin. Cancer Res.* <https://doi.org/10.1158/1078-0432.ccr-17-3420> (2018).
- Kotrotsou, A., Zinn, P. O. & Colen, R. R. Radiomics in brain tumors: an emerging technique for characterization of tumor environment. *Magn. Reson. Imaging Clin. North Am.* **24**, 719–729 (2016).
- Castellano, G., Bonilha, L., Li, L. M. & Cendes, F. Texture analysis of medical images. *Clin. Radiol.* **59**, 1061–1069 (2004).
- Davnull, F. et al. Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? *Insights into Imaging* **3**, 573–589 (2012).
- Kong, D. S. et al. Quantitative radiomic profiling of glioblastoma represents transcriptomic expression. *Oncotarget* **9**, 6336–6345 (2018).
- Rao, A. et al. A combinatorial radiographic phenotype may stratify patient survival and be associated with invasion and proliferation characteristics in glioblastoma. *J. Neurosurg.* **124**, 1008 (2016).
- Zinn, P. O. et al. Distinct radiomic phenotypes define glioblastoma TP53-PTEN-EGFR mutational landscape. *Neurosurgery* **64**, 203–210 (2017).
- Parekh, V. S. & Jacobs, M. A. Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI. *NPJ Breast Cancer* **3**, <https://doi.org/10.1038/s41523-017-0045-3> (2017).
- Shah, R. et al. Radiation necrosis in the brain: imaging features and differentiation from tumor recurrence. *Radiographics* **32**, 1343–1359 (2012).
- Kumar, A. J. et al. Malignant gliomas: MR imaging spectrum of radiation therapy- and chemotherapy-induced necrosis of the brain after treatment. *Radiology* **217**, 377–384 (2000).
- Ulmer, S. et al. Clinical and radiographic features of peritumoral infarction following resection of glioblastoma. *Neurology* **67**, 1668–1670 (2006).
- Parmar, C., Grossmann, P., Bussink, J., Lambin, P. & Aerts, H. J. W. L. Machine learning methods for quantitative radiomic. *Biomark. Sci. Rep.* **5**, 13087 (2015).
- Zinn, P. O. et al. Diffusion weighted magnetic resonance imaging radiophenotypes and associated molecular pathways in glioblastoma. *Neurosurgery* **63**(Suppl 1), 127–135 (2016).
- Zinn, P. O. et al. A coclinical radiogenomic validation study: conserved magnetic resonance radiomic appearance of periostin-expressing glioblastoma in patients and xenograft models. *Clin. Cancer Res.* <https://doi.org/10.1158/1078-0432.CCR-17-3420> (2018).
- Qin, J.-b et al. Grading of gliomas by using radiomic features on multiple magnetic resonance imaging (MRI) sequences. *Med. Sci. Monit.* **23**, 2168–2178 (2017).



35. Hall-Beyer, M. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. *Int. J. Remote Sens.* **38**, 1312–1338 (2017).
36. Tiwari, P. et al. Computer-extracted texture features to distinguish cerebral radionecrosis from recurrent brain tumors on multiparametric MRI: a feasibility study. *Am. J. Neuroradiol.* **37**, 2231 (2016).
37. Zinn, P. O. et al. Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme. *PLoS ONE* **6**, e25451 (2011).
38. Zinn, P. O. et al. A novel volume-age-KPS (VAK) glioblastoma classification identifies a prognostic cognate microRNA-gene signature. *PLoS ONE* **7**, e41522 (2012).
39. Tofts, P. S. Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *J. Magn. Reson. Imaging* **7**, 91–101 (1997).
40. Tofts, P. S. 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* **10**, 223–232 (1999).
41. Zhang, J. et al. Clinical applications of contrast-enhanced perfusion MRI techniques in gliomas: recent advances and current challenges. *Contrast Media Mol. Imaging* **2017**, 7064120 (2017).
42. Law, M. et al. Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging. *Am. J. Neuroradiol.* **24**, 1989–1998 (2003).
43. Wetzel, S. G. et al. Relative cerebral blood volume measurements in intracranial mass lesions: interobserver and intraobserver reproducibility study. *Radiology* **224**, 797–803 (2002).
44. Papoulis Athanasios, S. U. P. Probability, random variables, and stochastic processes. *McGraw-Hill* **4** (1991/2).
45. Clausi, D. A. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Can. J. Remote Sens.* **28**, 45–62 (2002).
46. Haralick, R. M., Shanmugam, K. & Dinstein, I. Textural features for image classification. *IEEE Trans. Syst., Man, Cybern. SMC* **3**, 610–621 (1973).
47. Soh, L. K. & Tsatsoulis, C. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Trans. Geosci. Remote Sens.* **37**, 780–795 (1999).
48. Zinn, P. O. et al. 139 Clinically applicable and biologically validated mri radiomic test method predicts glioblastoma genomic landscape and survival. *Neurosurgery* **63**, 156–157 (2016).
49. Colen, R. R. et al. Radiomics to predict immunotherapy-induced pneumonitis: proof of concept. *Invest. New Drugs*, <https://doi.org/10.1007/s10637-017-0524-2> (2017).
50. Hanchuan, P., Fuhui, L. & Ding, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**, 1226–1238 (2005).
51. Frank, E., Wang, Y., Inglis, S., Holmes, G. & Witten, I. H. Using model trees for classification. *Mach. Learn.* **32**, 63–76 (1998).
52. Vojt et al. in *Proceedings of the 28th International Conference on International Conference on Machine Learning* 665–672 (Omnipress, Bellevue, Washington, USA, 2011).
53. Platt, J. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Adv. large margin Classif.* **10**, 61–74 (1999). 3/26.
54. Refaeilzadeh, P., Tang, L. & Liu, H. in *Encyclopedia of Database Systems* (eds Ling Liu & M. Tamer Özsu) 532–538 (Springer US, 2009).
55. Team, R. C. R: A Language and Environment for Statistical Computing. doi: citeulike-article-id:12956656 (2013).

## Acknowledgements

This work was supported by the John S. Dunn Sr. Distinguished Chair in Diagnostic Imaging fund, MD Anderson Cancer Center startup funding, Multi-Investigator Imaging (MI2) Research Award, Radiological Society of North America scholar grant (RSCH11506). The Cancer Prevention & Research Institute of Texas (RP160150) and by the U.S. National Cancer Institute through the MD Anderson Cancer Center Support Grant (P30CA016672) (RRC).

## Author contribution

N. Elshafeey and R.R.C conceived the study; S. Ahmed, S. Abrol, A.H., I.H., A.A., K.S., J.A. and S.B. collected patient clinical data; N. Elshafeey and A.K. designed and performed the imaging analysis; N. Elshafeey, N. Elshafei, and A.K. performed the statistical analysis; J.T.H. and G.N.F. pathology evaluation and review; N. Elshafeey and A.H. prepared the figures and tables; F.M. and M.L. provided patient data from collaborating institutes; N. Elshafeey, A.K., P.O.Z., F.M., M.L. and M.L. provided support for the interpretation of the results; N. Elshafeey and A.K. drafted the manuscript; N. Elshafeey, A.K., F.M. and R.R.C critically revised the article for important intellectual content. All authors read and approved the final version of the article.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-019-11007-0>.

**Competing interests:** Meng Law receives Honorarium and Research Grant Support from Bracco Diagnostics. The remaining authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Peer review information:** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019