



## Research article

## Radiomics for pseudoprogression prediction in high grade gliomas: added value of MR contrast agent



Orkhan Mammadov<sup>a</sup>, Burak Han Akkurt<sup>a</sup>, Manfred Musigmann<sup>a</sup>, Asena Petek Ari<sup>a</sup>, David A. Blömer<sup>a</sup>, Dilek N.G. Kasap<sup>a</sup>, Dylan J.H.A. Henssen<sup>b</sup>, Nabila Gala Nacul<sup>a</sup>, Elisabeth Sartoretti<sup>c</sup>, Thomas Sartoretti<sup>c,d</sup>, Philipp Backhaus<sup>e,f</sup>, Christian Thomas<sup>g</sup>, Walter Stummer<sup>h</sup>, Walter Heindel<sup>a</sup>, Manoj Mannil<sup>a,\*</sup>

<sup>a</sup> University Clinic for Radiology, Westfälische Wilhelms-University Muenster and University Hospital Muenster, Albert-Schweitzer-Campus 1, DE-48149 Muenster, Germany

<sup>b</sup> Department of Medical Imaging, Radboud University Medical Center, Radboud University, 6500HB Nijmegen, the Netherlands

<sup>c</sup> Faculty of Medicine, University of Zurich, Zurich, Switzerland

<sup>d</sup> Department of Radiology and Nuclear Medicine, Maastricht University Medical Center, Maastricht University, Maastricht, the Netherlands

<sup>e</sup> Department of Nuclear Medicine, Westfälische Wilhelms-University Muenster and University Hospital Muenster, Albert-Schweitzer-Campus 1, DE-48149 Muenster, Germany

<sup>f</sup> European Institute for Molecular Imaging, Westfälische Wilhelms-University Muenster and University Hospital Muenster, Albert-Schweitzer-Campus 1, DE-48149 Muenster, Germany

<sup>g</sup> Institute of Neuropathology, Westfälische Wilhelms-University Muenster and University Hospital Muenster, Albert-Schweitzer-Campus 1, DE-48149 Muenster, Germany

<sup>h</sup> Department of Neurosurgery, Westfälische Wilhelms-University Muenster and University Hospital Muenster, Albert-Schweitzer-Campus 1, DE-48149 Muenster, Germany

## HIGHLIGHTS

- Radiomics allows for prediction of pseudoprogression in high-grade gliomas.
- Use of contrast media boosts the performance of the Radiomics prediction model.

## ARTICLE INFO

## Keywords:

Artificial intelligence

Glioma

Patient outcome assessment

## ABSTRACT

**Objective:** Our aim is to define the capabilities of radiomics in predicting pseudoprogression from pre-treatment MR images in patients diagnosed with high-grade gliomas using T1 non-contrast-enhanced and contrast-enhanced images.

**Material & methods:** In this retrospective IRB-approved study, image segmentation of high-grade gliomas was semi-automatically performed using 3D Slicer. Non-contrast-enhanced T1-weighted images and contrast-enhanced T1-weighted images were used prior to surgical therapy or radio-chemotherapy. Imaging data was split into a training sample and an independent test sample at random. We extracted 107 radiomic features by use of PyRadiomics. Feature selection and model construction were performed using Generalized Boosted Regression Models (GBM).

**Results:** Our cohort included 124 patients (female: n = 53), diagnosed with progressive (n = 61) and pseudo-progressive disease (n = 63) of primary high-grade gliomas. Based on non-contrast-enhanced T1-weighted images of the independent test sample, the mean area under the curve (AUC), mean sensitivity, mean specificity and mean accuracy of our model were 0.651 [0.576, 0.761], 0.616 [0.417, 0.833], 0.578 [0.417, 0.750] and 0.597 [0.500, 0.708] to predict the development of pseudoprogression. In comparison, the independent test data of contrast-enhanced T1-weighted images yielded significantly higher values of AUC = 0.819 [0.760, 0.872], sensitivity = 0.817 [0.750, 0.833], specificity = 0.723 [0.583, 0.833] and accuracy = 0.770 [0.687, 0.833].

**Conclusion:** Our findings show that it is possible to predict pseudoprogression of high-grade gliomas with a Radiomics model using contrast-enhanced T1-weighted images with comparatively good discriminatory power. The use of a contrast agent results in a clear added value.

\* Corresponding author.

E-mail address: [manoj.mannil@ukmuenster.de](mailto:manoj.mannil@ukmuenster.de) (M. Mannil).

## 1. Introduction

Glioblastoma are among the most common primary malignant brain tumors in adulthood [1]. The tumors are characterized by high vascularity [2], high lethality [3] and invasive growth [4]. The average survival rate in combination with radio- and chemotherapy ranges between 12.6 and 24 months [5]. On average, there are 4–11 cases per 100,000 people in a year [6] and 125,000 to 150,000 new cases are diagnosed annually [7]. Standard treatment consists of surgery, followed by radiation and chemotherapy (temozolomide) according to the Stupp scheme [8, 9]. Magnetic resonance imaging (MRI) is considered one of the most accurate imaging modalities for tumor assessment and response prediction [4]. However, it is difficult to distinguish true tumor progression or tumor recurrence from pseudoprogression [10]. Figure 1 shows an example of true progression and Figure 2 an example of pseudoprogression.

Pseudoprogression usually occurs within 3–6 months after completion of multimodal therapy [11] and is defined as a progression of findings on MRI images without clinical correlation, which then regresses over the course of therapy without any changes of therapeutic management [12]. Pseudoprogression appears as contrast-enhancing lesions on T1-weighted images and features an increase in signal intensity on FLAIR images surrounding the resection cavity [13, 14]. Tsakiris et al. showed that the occurrence of pseudoprogression in newly diagnosed glioblastomas is about 36% [15]. A misdiagnosis can result in redundant surgeries and potentially harmful therapeutic changes [15]. In MR imaging conducted for evaluation of tumor progression, the application of gadolinium-based contrast agent is considered common practice, as relevant diagnostic information can be visually observed from contrast-enhanced sequences.

However, with recent studies showing that gadolinium-based contrast-agents may be deposited in the body [16, 17], there is great

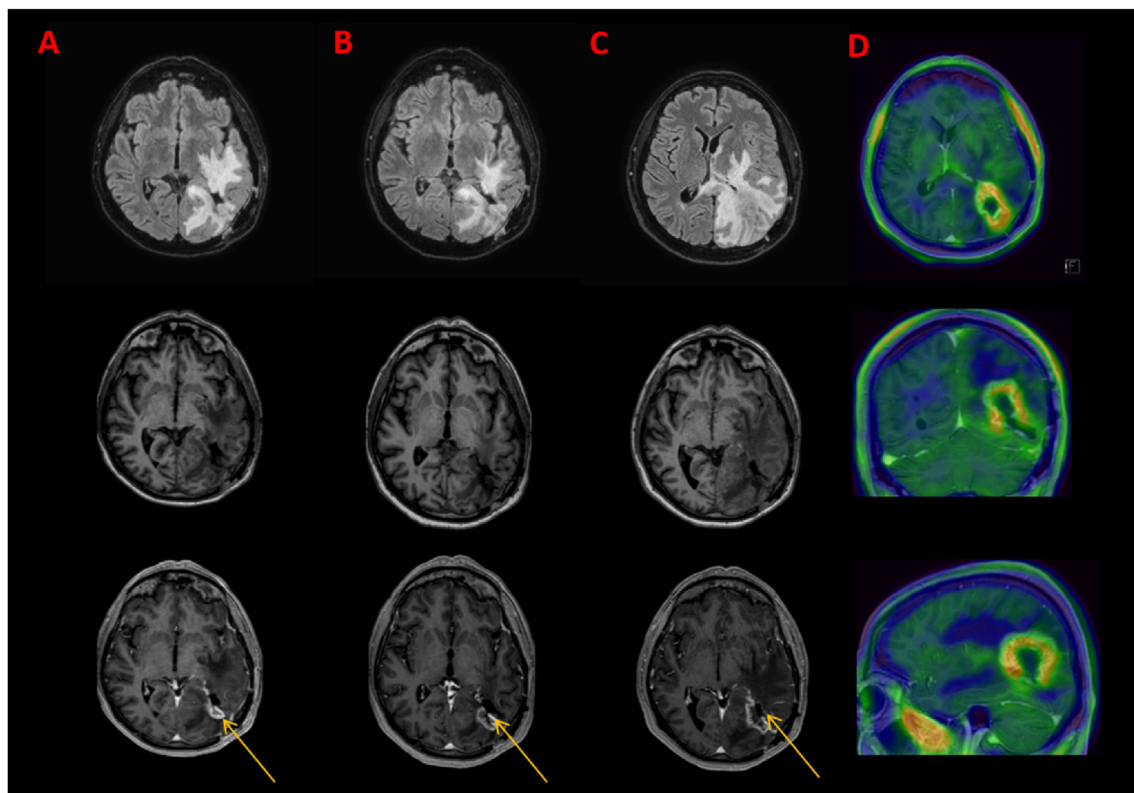
interest in avoiding the use of contrast agent as much as possible. There is a need for a novel, reliable and non-invasive diagnostic tool that allows for the correct pre-treatment prediction of true progression or pseudo-progression, preferably without requiring the image information from contrast-enhanced MRI sequences.

Radiomics analyses objectively quantify medical imaging features [18]. When combined with clinical data and histopathology, it allows for reliable predictions on disease prognosis and response to therapy [19, 20]. There are numerous studies on the benefits of radiomics in medical imaging such as the staging of liver fibrosis, the definition of focal liver lesions [21] or in the detection of prostate cancer [22]. Applied to gliomas, Cooker et al. have described an approximately 90% accuracy of radiomics techniques combined with machine learning when assessing the WHO 2016 grade for newly diagnosed gliomas [23]. Given the potential of radiomics, the aim of the present study was to determine the performance of prognostic models for distinguishing brain tumors with and without progression, and second, to compare the performance of our models using MR images generated without and with the administration of a contrast agent.

## 2. Materials & methods

### 2.1. Study population

The single-center, retrospective study was IRB-approved, performed in compliance with the Declaration of Helsinki and was approved by the local ethics committee (2021-596-f-S). Due to its retrospective nature, written informed consent was waived. We retrospectively screened our databases at the Department of Radiology, Nuclear medicine and Neuropathology for patients with histologically-proven Glioblastoma, who were presented to our tertiary referral hospital between January 2015 and June 2020.



**Figure 1.** Case of true progression of glioblastoma under multimodal therapy. Sequences: T2/FLAIR, T1, T1 Gd-enhanced, PET/MRI fusion. (A–C) MRI was performed 3 month, 6 month and 8 month of therapy regime and shows continuous aggravation of T2/FLAIR Signal (frist column) and increase of contrast enhancing parts (third column, arrows). (D) PET/MRI-fusion images confirming diagnosis of true progression.

From the initially detected 193 patients we excluded those with (1) missing or non-diagnostic pre-treatment cerebral magnetic resonance imaging, (2) insufficient diagnostic imaging quality, (3) incomplete clinical data, (4) inconsistent histopathology, (5) insufficient follow-up examinations (e.g. denied treatment/biopsy), and (6) images available for only one of the two T1 sequences without or with contrast agent.

Finally we included 124 patients (male:  $n = 71$ ; female:  $n = 53$ ), diagnosed with progress ( $n = 61$ ) and no progress ( $n = 63$ ) of the brain tumor. The mean age of the patients was 61.02 years. The histopathological and demographic data of the training sample and the independent test sample are summarized in Table 1. A detailed table with molecular subtypes, histologic findings, and type and timing of therapy is provided in the Appendix.

## 2.2. Image acquisition

We required MR images for each patient in our cohort, both without and with contrast administration. We used the T1 MRI sequence to generate the MR images for both, the unenhanced T1-weighted images and the contrast-enhanced T1-weighted images. For feature computation we used the open source software (3D slicer, version 4.11) with a bin width of 25 and resampled voxel sizes to 2,2,2. We extracted a total of 107 radiometric features by hand-delineated regions of interest (ROI) from the unenhanced T1-weighted images and additionally from the

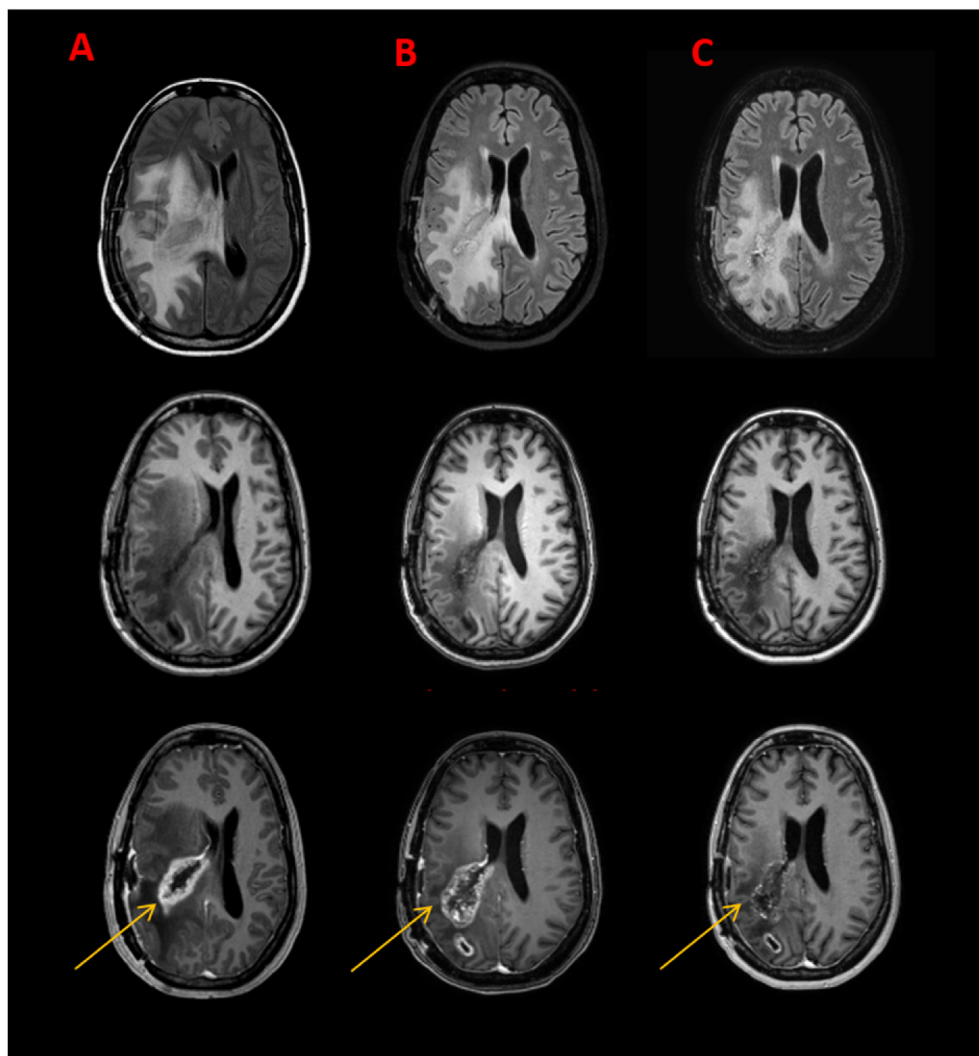
**Table 1.** Histopathological and demographic data.

	Training sample	Independent test sample
Number	100	24
Progress: Number (in %)		
Yes	49 (49.0%)	12 (50.0%)
No	51 (51.0%)	12 (50.0%)
Gender: Number (in %)		
Male	57 (57.0%)	14 (58.3%)
Female	43 (43.0%)	10 (41.7%)
Age (years)	60.67	62.46

contrast-enhanced T1-weighted images of each patient and compared the performance of models using the unenhanced T1-weighted images with corresponding models using the contrast-enhanced T1-weighted images. Detailed MR acquisition parameters are mentioned in the supplementary section.

## 2.3. Feature extraction

In both parts of our analysis, we extracted 107 radiometric features by manually-delineated regions of interest (ROI) from each patient's MR images. In order to make the data more normal distribution-like, all



**Figure 2.** Case of pseudoprogression in a GBM patient. Sequences: T2/FLAIRw, T1w, T1w Gd-enhanced. (A–C) MRI performed at the beginning of the therapy as well as 3 month and 4,5 month after the therapy start. FLAIR signal and T1 Gd-enhanced (arrows) signal temporarily increases during therapy.

features underwent a Yeo-Jonson transformation. The features were z-score normalized and subjected to a 95% correlation filter keeping 56 features to account for redundancy between the features. Feature pre-selection and model construction were both performed with the training sample, using Generalized Boosted Regression Models (GBM). A GBM is a combination of a decision tree algorithm and a boosting technique. Usually, GBM prediction models are constructed as an ensemble of weak prediction models, i.e., weak learners.

#### 2.4. Feature pre-selection

We used the “varImp” function in R to identify first the most important variables. This function determines how many times each variable is selected during the building process of the decision trees and how much the prediction error of the model is improved by using each variable. We determined the most important features firstly for the contrast-enhanced images and secondly for the unenhanced images. Table 2 lists the top 15 features for both the contrast-enhanced images and the unenhanced images in descending order of importance.

#### 2.5. Model development

GBM models were then created with an increasing number of the most important features identified previously. In the first step, the model contains only the most important feature, followed by a model with the two most important features, followed by a model with the three most important features, and so on. The model with the highest performance with respect to the hold-out samples used in the cross-validation to determine the tuning parameters included in the GBM model is used as the final model. This step-by-step approach determines the final number of features included in the model. The GBM models contain several tuning parameters: firstly the “tree depth”, secondly the “learning rate”, thirdly the “minimum number of observations in the terminal node” and finally the “number of trees”. The optimal tuning parameters of the GBM models (tree depth = 1 or 2; learning rate = 0.1; minimum number of observations in terminal nodes = 5,7,9,11,13 or 15; number of trees = 50, 60, 70, ...,150) were determined using grid search 10-fold cross-validation, i.e., we divided the training sample 10 times into groups with 90% and 10% of the training data, respectively. The 10 groups that each contain 10% of the training data are denoted as “hold-out samples of the training data”. The technique ensures that the subgroups of the training sample do not overlap. This methodology provides robust results even in combination with small datasets.

The tuning parameters of the GBM model may slightly depend on the data partitioning used in the cross-validation. To determine the stability

of the results, we therefore optimized each of the models with a given number of features 100 times and then tested each of these models with the test sample. The predictive power of the models was analyzed using the area under the curve (AUC) of the receiver operator characteristic (ROC) and the accuracy. All our performance values were determined as means of 100 cycles/repetitions.

#### 2.6. Model analysis

In the first part of our analyses, we used the contrast-enhanced T1-weighted images and determined the maximum performance in discriminating the progression/non-progression of the brain tumors with machine learning algorithms.

In the second part of our analyses, we attempted to obtain comparable results based on the unenhanced T1-weighted images. First, all features were recalculated for the unenhanced images. For these comparative analyses with the unenhanced images, the same approach was used for model construction as in the first part of our analyses. However, we tried two slightly different approaches for variable preselection:

- First, we used the same features as in the first part of our analyses. The order in which each variable was included in the models was also maintained. This means that the selection of the features used and their order still referred to the contrast-enhanced images.
- Subsequently we determined the most important features in relation to the unenhanced T1-weighted images. The order in which these variables were subsequently included in the models was now based on their importance in relation to the unenhanced images.

Finally, the models were again estimated now using these two different sets of features and the unenhanced images. By comparing the results from the first and second parts of our analyses, we were able to determine the added value of the MR contrast agent.

#### 2.7. Statistical analysis

Statistical analysis was performed using R software (version 3.5.3). As mentioned, unenhanced T1-weighted images and contrast-enhanced T1-weighted images before treatment were available for 124 patients respectively. These 124 patients were allocated to a training sample and an independent test sample at random. The training sample was used for the construction of the different models and the optimization of the tuning parameters included in these models. The performance of the models was determined using the test sample, i.e., unknown/independent data. We used a stratified 4:1 ratio. The training sample included

**Table 2.** List of the most important features for the contrast-enhanced images and the unenhanced images in descending order of importance.

Feature number	Features for contrast-enhanced images	Features for unenhanced images
1	T1_GD_1.orig.ngtdm.Strength	T1_nativ_1.orig.shape.Sphericity
2	T1_GD_1.orig.glcmm.ClusterShade	T1_nativ_1.orig.shape.MajorAxisLength
3	T1_GD_1.orig.shape.Elongation	T1_nativ_1.orig.shape.Flatness
4	T1_GD_1.orig.shape.Flatness	T1_nativ_1.orig.ngtdm.Contrast
5	T1_GD_1.orig.shape.MinorAxisLength	T1_nativ_1.orig.shape.Elongation
6	T1_GD_1.orig.shape.Sphericity	T1_nativ_1.orig.glcmm.Idm
7	T1_GD_1.orig.fst.ord.RobustMeanAbsoluteDeviation	T1_nativ_1.orig.fst.ord.Kurtosis
8	T1_GD_1.orig.fst.ord.Uniformity	T1_nativ_1.orig.glcmm.InverseVariance
9	T1_GD_1.orig.glcmm.Idmn	T1_nativ_1.orig.glcmm.Correlation
10	T1_GD_1.orig.glcmm.Correlation	T1_nativ_1.orig.glcmm.Imc1
11	T1_GD_1.orig.glcmm.Idm	T1_nativ_1.orig.glrmm.RunEntropy
12	T1_GD_1.orig.glcmm.Imc2	T1_nativ_1.orig.shape.SurfaceVolumeRatio
13	T1_GD_1.orig.glcmm.MCC	T1_nativ_1.orig.glszm.SizeZoneNonUniformity
14	T1_GD_1.orig.fst.ord.Skewness	T1_nativ_1.orig.shape.LeastAxisLength
15	T1_GD_1.orig.ngtdm.Busyness	T1_nativ_1.orig.glcmm.SumAverage

100 patients and the test sample 24 patients with a balanced distribution between both samples (Table 1) of tumor progress (yes/no) and gender (F = female/M = male). We started our analyses using the contrast-enhanced images. It is important to note that we kept the assignment of the 124 patients to the training sample and the test sample unchanged in both parts of our study. This means that regardless of whether we used the contrast-enhanced (first part of our study) or the unenhanced MR images (second part of our study), the same 100 patients formed the training sample and the remaining 24 patients formed the test sample. P values below <.05 are considered significant.

### 3. Results

#### 3.1. Contrast-enhanced images

For the first part of our analyses with the contrast-enhanced T1-weighted images, a GBM model was used for the feature preselection and for the subsequent model construction. Starting with the most important of the original 56 features (“T1\_GD\_1.orig.ngtdm.Strength”), we added one additional feature in every subsequent step. The optimization of each GBM model was repeated 100 times using grid search 10-fold cross-validation. The results averaged over 100 cycles for each model are summarized in Table 3. The performance of the models depends only to a limited extent on the exact number of features used. Models with good discriminatory power are obtained with both the training sample and the independent test sample. The best model in terms of AUC with respect to the hold out-samples of the training data is obtained with approximately six to seven features. This applies accordingly to the independent test data (Figure 3). With the addition of the first features, starting from a model with only one feature, the AUC increases strongly for the training sample and moderately for the independent test data. Models including more than about 6 features do not result in higher AUC values for the training data and even slightly lower values for the independent test data, i.e., in this range the model starts to be overfitted.

The correlation matrix for the best model including the first six features is shown in Figure 4. Most of the correlation coefficients only have small values, i.e., most of the features used in this model are almost independent of each other. In the independent testing group, the mean AUC, mean sensitivity, mean specificity and mean accuracy of this model were 0.819 [0.760, 0.872], 0.817 [0.750, 0.833], 0.723 [0.583, 0.833] and 0.770 [0.687, 0.833] and in the training sample 0.923 [0.883, 0.983], 0.779 [0.694, 0.910], 0.888 [0.824, 0.952] and 0.835 [0.780, 0.926] respectively. The values in the brackets indicate the 95%

confidence intervals. Hence, this final GBM model shows good prediction performance in both training and test group. In the left part of Figure 5 the ROC curve for the test group is shown. As our results show, brain tumor progression/non-progression can be predicted with comparatively good discriminatory power using machine learning algorithms based on contrast-enhanced T1-weighted images.

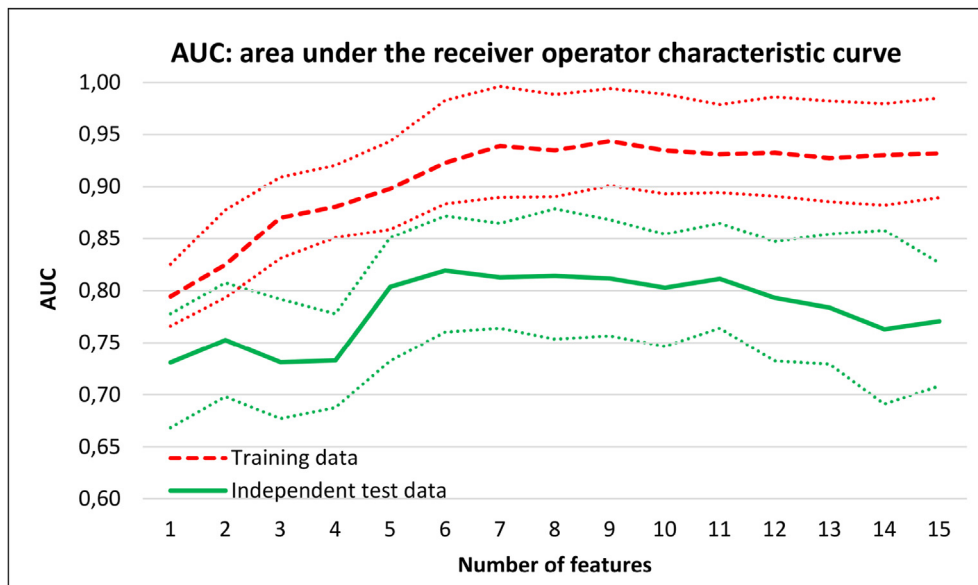
#### 3.2. Non-contrast-enhanced images

For the second part of our analyses, we used the unenhanced T1-weighted images. As already described, the calculations were first performed with the variables previously determined using the contrast-enhanced T1-weighted images. The values of the variables were recalculated using the corresponding data of the unenhanced images. We also kept the model approach of a GBM model. Table 4 shows the results for these GBM models as a function of the number of variables used. The models were optimized according to the previous analyses by maximizing the AUC, using cross-validation. Comparable to the first part of the analyses, the “best” models also had approximately 6 variables. However, the achieved discriminatory power values were clearly below the corresponding values that were previously determined using the contrast-enhanced T1-weighted images. The accuracy values are not even above those of a random model.

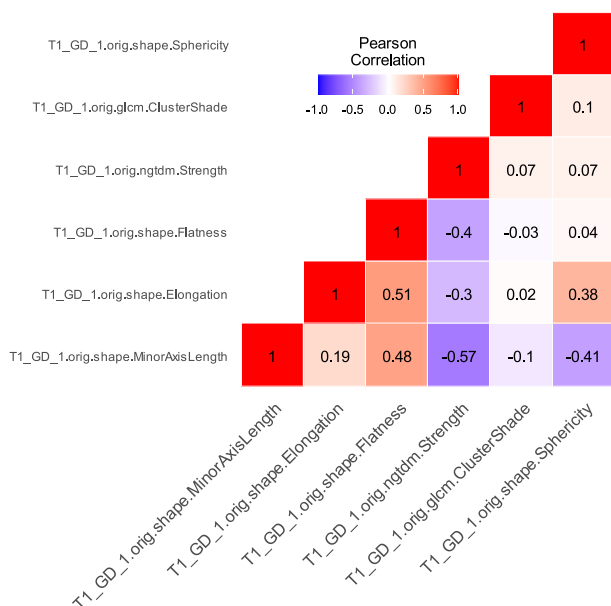
According to Table 2, it is obvious that different features are important for the unenhanced images than for the contrast-enhanced images. We therefore repeated the optimization of our GBM models. However, we now used the most important variables in relation to the unenhanced images. This preselection of variables was again performed using a GBM model. The results are summarized in Table 5. Compared to the results in Table 4, slightly higher discriminatory power values were obtained, but these values remain significantly below the values obtained with the contrast-enhanced T1-weighted images. The highest accuracy values with the independent test data were slightly above 60%, and thus close to the value of 50%, which would result from a purely random experiment. The independent test data with T1 non-contrast images using the same GBM methodology with 6 features yielded values of 0.651 [0.576, 0.761] for the mean AUC, 0.616 [0.417, 0.833] for the mean sensitivity, 0.578 [0.417, 0.750] for the mean specificity and 0.597 [0.500, 0.708] for the mean accuracy. The ROC curve for the test group is shown in the right part of Figure 5. We thus obtained our best results with both the contrast-enhanced and the non-contrast-enhanced images with 6 features each. We compared the two results in relation to the AUC using the DeLong test [24]. We obtained a p-value < 2.2e-16, which means that the

**Table 3.** Classification results per group using the contrast-enhanced T1-weighted images. AUC: area under the receiver operator characteristic curve. Sens.: sensitivity. Spec.: specificity. Acc.: accuracy.

Number of features	Training data				Independent test data			
	AUC	Sens.	Spec.	Acc.	AUC	Sens.	Spec.	Acc.
1	0.7943	0.6853	0.7014	0.6935	0.7308	0.7533	0.4417	0.5975
2	0.8247	0.6853	0.8290	0.7586	0.7525	0.6842	0.7475	0.7158
3	0.8701	0.7037	0.8153	0.7606	0.7314	0.6867	0.7067	0.6967
4	0.8805	0.7241	0.8294	0.7778	0.7329	0.7108	0.7075	0.7092
5	0.8979	0.7590	0.8529	0.8069	0.8035	0.8142	0.6325	0.7233
6	0.9225	0.7788	0.8884	0.8347	<b>0.8192</b>	0.8167	0.7225	0.7696
7	0.9388	0.8002	0.9041	0.8532	0.8128	0.8017	0.7175	0.7596
8	0.9345	0.7978	0.8980	0.8489	0.8142	0.7992	0.7083	0.7538
9	<b>0.9435</b>	0.8273	0.9059	0.8674	0.8117	0.8000	0.7383	0.7692
10	0.9347	0.8114	0.8965	0.8548	0.8028	0.7750	0.7650	0.7700
11	0.9308	0.8129	0.8912	0.8528	0.8114	0.7725	0.7825	0.7775
12	0.9322	0.8163	0.8902	0.8540	0.7930	0.7283	0.7808	0.7546
13	0.9272	0.8137	0.8857	0.8504	0.7837	0.7408	0.7733	0.7571
14	0.9299	0.8198	0.8800	0.8505	0.7632	0.7175	0.7475	0.7325
15	0.9318	0.8282	0.8851	0.8572	0.7707	0.7333	0.7442	0.7388



**Figure 3.** Mean AUCs (100 cycles) for the GBM models using the contrast-enhanced T1-weighted images with different number of features. Dotted lines: 95% confidence interval.



**Figure 4.** Pearson Correlation for the GBM model with 6 features using the contrast-enhanced T1-weighted images.

discriminatory power of the two models (with 6 features each) is significantly different with extremely high probability.

### 3.3. Test of further machine learning models

It should be noted that in addition to the method described here using a GBM model, we have tried numerous other machine learning methods for both feature preselection and model estimation. However, all these calculations using the unenhanced images resulted in much lower discriminative powers than those we were able to achieve with the contrast-enhanced images. In detail, we tried a total of 9 different methods for the feature preselection. We used “distance correlation”, “linear discriminant analysis (LDA)”, “univariate analysis”, “Lasso regression”, “Ridge regression”, “elastic net”, “random forest”, “bagged trees” and “naïve Bayes”. The subsequent model estimation was then

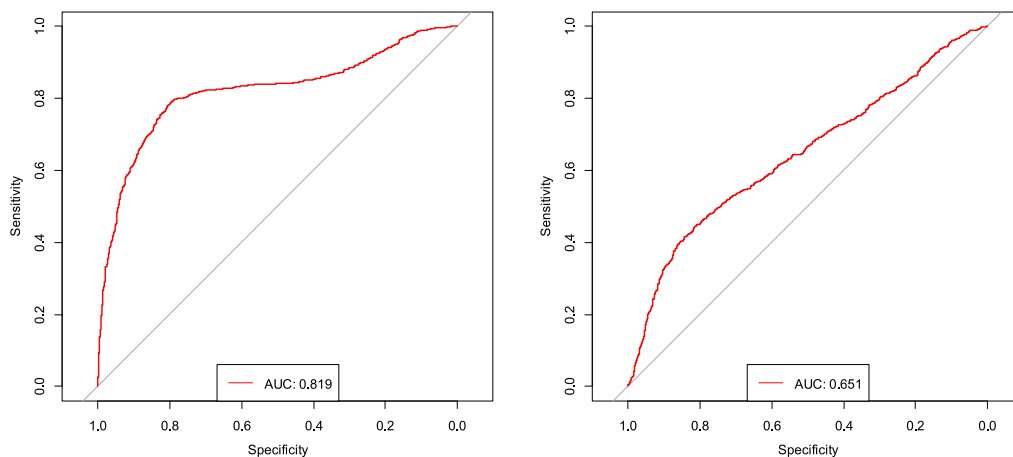
carried out with a total of 7 different model approaches, namely “linear discriminant analysis (LDA)”, “Lasso regression”, “Ridge regression”, “elastic net”, “random forest”, “bagged trees” and “naïve Bayes”, resulting in a total number of 63 possible combinations. Our two best combinations were firstly “random forest” for variable preselection with linear discriminant analysis as model and secondly “bagged trees” for variable preselection in combination with “random forest” as model. With these two combinations, the independent test data yielded values for AUC and accuracy slightly greater than 0.7. However, even these best values are below the corresponding values obtained with the contrast-enhanced T1-weighted images. In addition, it must be noted, that although the two described combinations led to comparatively good results using the independent test data, the corresponding performance values with the hold-out samples of the training data were lower. Therefore, a certain random effect cannot be ruled out here either.

### 4. Discussion

The diagnosis of glioblastoma is based on histology and several molecular markers. The appearance and development of neurological symptoms allow the estimation of the growth dynamics of gliomas [25].

Due to the increased perfusion and higher blood volume of gliomas, clinically best validated technique on brain tumor growth and response to treatment is CBV measurement derived from DSC-MRI [26]. Radiomics is ready to contribute to the imaging arsenal. Morphological and textural signatures derived from the high-throughput extraction of quantitative MR image metrics at the voxel level can be used by Radiomics techniques to make an accurate diagnosis and evaluate tumor response [27].

Our analyses show that it is possible to predict development of pseudoprogression of high-grade gliomas with machine learning algorithms using contrast-enhanced T1-weighted images with comparatively good discriminatory power before treatment. However, without the use of a contrast agent, the prediction quality of the tested learning algorithms is significantly reduced using the same T1 sequence. Our models using MR images without contrast agent yielded a discriminatory power that was only conditionally higher than that of a random model. The mean AUC could be increased from 0.651 to 0.819 and the mean accuracy from 0.597 to 0.770 by using the contrast agent. It is obvious that the use of a contrast agent significantly contributes to the discriminatory power achieved. We verified the increase in discriminatory power using the DeLong test.



**Figure 5.** ROC curves (test group) for GBM models with 6 features for the prediction of tumor progress using the contrast-enhanced T1-weighted images (left figure) and the unenhanced T1-weighted images (right figure).

**Table 4.** Classification results per group using the unenhanced T1-weighted images, features determined with the contrast-enhanced T1-weighted images. AUC: area under the receiver operator characteristic curve. Sens.: sensitivity. Spec.: specificity. Acc.: accuracy.

Number of features	Training data				Independent test data			
	AUC	Sens.	Spec.	Acc.	AUC	Sens.	Spec.	Acc.
1	0.7061	0.6027	0.6914	0.6479	0.6308	0.4458	0.7517	0.5988
2	0.7108	0.6188	0.6802	0.6501	0.6187	0.4033	0.7775	0.5904
3	0.8027	0.6759	0.7418	0.7095	0.5387	0.4267	0.6133	0.5200
4	0.8091	0.6876	0.7508	0.7198	0.5378	0.3742	0.5733	0.4738
5	0.8571	0.7469	0.7976	0.7728	0.5649	0.5875	0.4817	0.5346
6	0.8829	0.7549	0.8043	0.7801	0.6099	0.5825	0.5650	0.5738
7	0.8810	0.7604	0.8041	0.7827	0.5956	0.5942	0.5533	0.5738
8	0.8780	0.7614	0.7908	0.7764	0.6246	0.5833	0.5842	0.5838
9	0.8754	0.7688	0.7898	0.7795	0.6254	0.5875	0.5642	0.5758
10	0.8784	0.7700	0.8037	0.7872	0.5856	0.5392	0.5675	0.5533
11	0.8729	0.7716	0.7869	0.7794	0.6158	0.5450	0.5858	0.5654
12	0.8602	0.7527	0.7806	0.7669	0.5826	0.5025	0.5650	0.5338
13	0.8737	0.7633	0.7982	0.7811	0.5865	0.5300	0.5692	0.5496
14	0.8664	0.7612	0.7988	0.7804	0.5594	0.4575	0.5717	0.5146
15	0.8718	0.7667	0.8051	0.7863	0.5542	0.4475	0.5850	0.5163

Jang et al. first developed a machine learning algorithm that showed acceptable performance in distinguishing between real progress and pseudoprogression [28]. The T1 contrast-enhanced sequence and various clinical data, such as molecular characteristics, age, gender, and time after completion of therapy were selected as inputs to the model. However, due to the small data set, this model required further validation. In 2020, Jang et al. optimized their previous machine learning model with more clinical data [29]. Sun et al. also investigated the ability of radiomics features on T1 contrast-enhanced images to discriminate true progression from pseudoprogression using clinical data. Their radiomics model showed an AUC of 0.72, and a sensitivity of 78,36% [30]. Another study combined the clinical features and the MGMT promoter methylation status. Here a radiomics model was built on T1-weighted, T2-weighted images and apparent diffusion coefficient (ADC) maps [31] resulting in AUC of 0.80, a sensitivity of 78.2%, specificity of 66.7%, and an accuracy of 73.7%. Compared to other studies available in the literature, we included more datasets in order to establish our machine learning models. Furthermore, we did not rely on any input data outside the objectively quantified MR images. Our study is the first study that can unequivocally show that contrast agent is beneficial to predict the response.

Although previous studies have shown that gadolinium-based contrast agents may be deposited in the body, there is still no scientific consensus on whether gadolinium is dangerous or harmful. Therefore, relevant guidelines state that no patient should be denied gadolinium if the clinical indication justifies contrast administration [32]. Our study shows that even in radiomics, the contrast-enhanced T1w sequence continues to provide important diagnostic information. Thus, our results further provide evidence, that the administration of contrast agent in patients with suspicion of tumor progression is decisive and should not be omitted [33, 34]. In pediatric patients and patients with impaired renal function due to rapidly repeated measurements, ASL techniques are discussed as an alternative [26].

There are some limitations in our study. Firstly, this analysis was based on a retrospective data set, which has inherent limitations. Secondly, we used MR images of different vendors and could not account for differences in scanning technique. Finally, despite our greatest efforts, we cannot rule out overfitting of our results. For this reason, larger prospective trials are needed for further validation.

In conclusion, our study shows the capabilities of a Radiomics analysis based on T1 weighted MR images in predicting the occurrence of pseudoprogression in high-grade gliomas. We observed an added value in

**Table 5.** Classification results per group using the unenhanced T1-weighted images, features determined with the unenhanced T1-weighted images. AUC: area under the receiver operator characteristic curve. Sens.: sensitivity. Spec.: specificity. Acc.: accuracy.

Number of features	Training data				Independent test data			
	AUC	Sens.	Spec.	Acc.	AUC	Sens.	Spec.	Acc.
1	0.7458	0.5898	0.7298	0.6612	0.5524	0.4983	0.6442	0.5713
2	0.8442	0.7429	0.7616	0.7524	0.6100	0.5958	0.5700	0.5829
3	0.8607	0.7665	0.7492	0.7577	0.5458	0.5633	0.4958	0.5296
4	0.8818	0.8004	0.7843	0.7922	0.6319	0.5925	0.6150	0.6038
5	0.9200	0.8349	0.8441	0.8396	0.6387	0.7350	0.5425	0.6388
6	0.9059	0.8369	0.8157	0.8261	<b>0.6505</b>	0.6158	0.5775	0.5967
7	0.9139	0.8557	0.8220	0.8385	0.6300	0.5783	0.5550	0.5667
8	0.9202	0.8706	0.8212	0.8454	0.5605	0.5033	0.5733	0.5383
9	0.9605	0.9253	0.8747	0.8995	0.5644	0.5358	0.5450	0.5404
10	0.9567	0.9200	0.8645	0.8917	0.5609	0.5083	0.5892	0.5488
11	0.9578	0.9147	0.8788	0.8964	0.6421	0.5542	0.6083	0.5813
12	0.9543	0.9127	0.8708	0.8913	0.6372	0.5742	0.6233	0.5988
13	0.9621	0.9220	0.8825	0.9019	0.6177	0.5842	0.6058	0.5950
14	0.9663	0.9276	0.8990	0.9130	0.6205	0.5792	0.6058	0.5925
15	0.9559	0.9145	0.8743	0.8940	0.6111	0.5867	0.6000	0.5933

administering gadolinium-based contrast media for higher diagnostic accuracy.

## Declarations

### Author contribution statement

Orkhan Mammadov, Burak Han Akkurt, Dylan J.H.A. Henssen, Nabila Gala Nacul, Elisabeth Sartoretti, Thomas Sartoretti: Analyzed and interpreted the data; Wrote the paper.

Manfred Musigmann: Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Asena Petek Ari, David A. Blömer, Dilek N.G. Kasap, Philipp Backhaus, Christian Thomas, Walter Stummer: Contributed reagents, materials, analysis tools or data.

Walter Heindel: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data.

Manoj Mannil: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

### Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### Data availability statement

Data will be made available on request.

### Declaration of interests statement

The authors declare no conflict of interest.

### Additional information

Supplementary content related to this article has been published online at <https://doi:10.1016/j.heliyon.2022.e10023>.

## References

- [1] M.W. Ronellenfitsch, P.S. Zeiner, M. Mittelbronn, H. Urban, T. Pletsch, D. Reuter, C. Senft, J.P. Steinbach, M. Westphal, P.N. Harter, Akt and mTORC1 signaling as

- predictive biomarkers for the EGFR antibody nimotuzumab in glioblastoma, *Acta Neuropathol. Commun.* 6 (1) (2018) 81.
- [2] T. Cloughesy, G. Finocchiaro, C. Belda-Iniesta, L. Recht, A.A. Brandes, E. Pineda, T. Mikkelsen, O.L. Chinot, C. Balana, D.R. Macdonald, M. Westphal, K. Hopkins, M. Weller, C. Bais, T. Sandmann, J.M. Bruey, H. Koeppen, B. Liu, W. Verret, S.C. Phan, D.S. Shames, Randomized, double-blind, placebo-controlled, multicenter phase II study of Onartuzumab plus bevacizumab versus placebo plus bevacizumab in patients with recurrent glioblastoma: efficacy, safety, and hepatocyte growth factor and O(6)-methylguanine-DNA methyltransferase biomarker analyses, *J. Clin. Oncol.* 35 (3) (2017) 343–351.
- [3] J. Duerinck, S. Du Four, F. Bouthens, C. Andre, V. Verschaeve, F. Van Fraeyenhove, C. Chaskis, N. D'Haene, M. Le Mercier, A. Rogiers, A. Michotte, I. Salmon, B. Neyns, Randomized phase II trial comparing axitinib with the combination of axitinib and lomustine in patients with recurrent glioblastoma, *J. Neuro Oncol.* 136 (1) (2018) 115–125.
- [4] A. Forster, J. Bohme, M.E. Maros, S. Brehmer, M. Seiz-Rosenhagen, D. Hanggi, F. Wenz, C. Groden, W.B. Pope, F.A. Giordano, Longitudinal MRI findings in patients with newly diagnosed glioblastoma after intraoperative radiotherapy, *J. Neuroradiol.* 47 (2) (2020) 166–173.
- [5] H. Elinzano, M. Glantz, M. Mrugala, S. Kesari, D.E. Piccioni, L. Kim, E. Pan, S. Yunus, T. Coyle, K. Timothy, D. Evans, K. Mantripragada, J. Boxerman, T. DiPetrillo, J.E. Donahue, N. Hebda, K.M. Mitchell, K.L. Rosati, H. Safran, PPX and concurrent radiation for newly diagnosed glioblastoma without MGMT methylation: a randomized phase II study: BrUOG 244, *Am. J. Clin. Oncol.* 41 (2) (2018) 159–162.
- [6] T.A. Lawrie, D. Gillespie, T. Dowswell, J. Evans, S. Erridge, L. Vale, A. Kernohan, R. Grant, Long-term neurocognitive and other side effects of radiotherapy, with or without chemotherapy, for glioma, *Cochrane Database Syst. Rev.* 8 (2019) CD013047.
- [7] Q.T. Ostrom, H. Gittleman, P. Liao, T. Vecchione-Koval, Y. Wolinsky, C. Kruchko, J.S. Barnholtz-Sloan, CBTRUS Statistical Report: primary brain and other central nervous system tumors diagnosed in the United States in 2010–2014, *Neuro Oncol.* 19 (2017) v1–v88 (suppl\_5).
- [8] T.C. Booth, A. Luis, L. Brazil, G. Thompson, R.A. Daniel, H. Shuaib, K. Ashkan, A. Pandey, Glioblastoma post-operative imaging in neuro-oncology: current UK practice (GIN CUP study), *Eur. Radiol.* 31 (5) (2021) 2933–2943.
- [9] F.S. Lieberman, M. Wang, H.I. Robins, C.I. Tsien, W.J. Curran Jr., M. Werner-Wasik, R.P. Smith, C. Schultz, A.C. Hartford, P. Zhang, M.P. Mehta, Phase 2 study of radiation therapy plus low-dose temozolomide followed by temozolomide and irinotecan for glioblastoma: NRG oncology RTOG trial 0420, *Int. J. Radiat. Oncol. Biol. Phys.* 103 (4) (2019) 878–886.
- [10] C.G. Brahm, M.W. den Hollander, R.H. Enting, J.C. de Groot, A.M. Solouki, W.F.A. den Dunnen, M. Heesters, M. Wagemakers, H.M.W. Verheul, E.G.E. de Vries, J. Pruim, A.M.E. Walenkamp, Serial FLT PET imaging to discriminate between true progression and pseudoprogression in patients with newly diagnosed glioblastoma: a long-term follow-up study, *Eur. J. Nucl. Med. Mol. Imag.* 45 (13) (2018) 2404–2412.
- [11] S.B. Strauss, A. Meng, E.J. Ebani, G.C. Chiang, Imaging glioblastoma posttreatment: progression, pseudoprogression, pseudoresponse, radiation necrosis, *Radiol. Clin.* 57 (6) (2019) 1199–1216.
- [12] N. Galldiks, I. Law, W.B. Pope, J. Arbizu, K.J. Langen, The use of amino acid PET and conventional MRI for monitoring of brain tumor therapy, *Neuroimage Clin.* 13 (2017) 386–394.
- [13] P.D. Delgado-Lopez, E. Rinones-Mena, E.M. Corrales-Garcia, Treatment-related changes in glioblastoma: a review on the controversies in response assessment criteria and the concepts of true progression, pseudoprogression, pseudoresponse and radionecrosis, *Clin. Transl. Oncol.* 20 (8) (2018) 939–953.



- [14] R.G. Yoon, H.S. Kim, W. Paik, W.H. Shim, S.J. Kim, J.H. Kim, Different diagnostic values of imaging parameters to predict pseudoprogression in glioblastoma subgroups stratified by MGMT promoter methylation, *Eur. Radiol.* 27 (1) (2017) 255–266.
- [15] C. Tsakiris, T. Siempis, G.A. Alexiou, A. Zikou, C. Sioka, S. Voulgaris, M.I. Argyropoulou, Differentiation between true tumor progression of glioblastoma and pseudoprogression using diffusion-weighted imaging and perfusion-weighted imaging: systematic review and meta-analysis, *World Neurosurg.* 144 (2020) e100–e109.
- [16] B.J. Guo, Z.L. Yang, L.J. Zhang, Gadolinium deposition in brain: current scientific evidence and future perspectives, *Front. Mol. Neurosci.* 11 (2018) 335.
- [17] H. Richter, P. Buckner, L.F. Martin, C. Dunker, S. Fingerhut, A. Xia, A. Karol, M. Sperling, U. Karst, A. Radbruch, A. Jeibmann, Gadolinium tissue distribution in a large-animal model after a single dose of gadolinium-based contrast agents, *Radiology* 301 (3) (2021) 637–642.
- [18] V. Dalal, J. Carmicheal, A. Dhaliwal, M. Jain, S. Kaur, S.K. Batra, Radiomics in stratification of pancreatic cystic lesions: machine learning in action, *Cancer Lett.* 469 (2020) 228–237.
- [19] J.E. Park, S. Ham, H.S. Kim, S.Y. Park, J. Yun, H. Lee, S.H. Choi, N. Kim, Diffusion and perfusion MRI radiomics obtained from deep learning segmentation provides reproducible and comparable diagnostic model to human in post-treatment glioblastoma, *Eur. Radiol.* 31 (5) (2021) 3127–3137.
- [20] S. Ursprung, L. Beer, A. Bruining, R. Woitek, G.D. Stewart, F.A. Gallagher, E. Sala, Radiomics of computed tomography and magnetic resonance imaging in renal cell carcinoma—a systematic review and meta-analysis, *Eur. Radiol.* 30 (6) (2020) 3558–3566.
- [21] H.J. Park, B. Park, S.S. Lee, Radiomics and deep learning: hepatic applications, *Korean J. Radiol.* 21 (4) (2020) 387–401.
- [22] R. Cuocolo, M.B. Cipullo, A. Stanzione, L. Ugga, V. Romeo, L. Radice, A. Brunetti, M. Imbriaco, Machine learning applications in prostate cancer magnetic resonance imaging, *Eur. Radiol. Exp.* 3 (1) (2019) 35.
- [23] M. Kocher, M.I. Ruge, N. Galldiks, P. Lohmann, Applications of radiomics and machine learning for radiotherapy of malignant brain tumors, *Strahlenther. Onkol.* 196 (10) (2020) 856–867.
- [24] E.R. DeLong, D.M. DeLong, D.L. Clarke-Pearson, Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach, *Biometrics* 44 (3) (1988) 837–845.
- [25] M. Weller, M. van den Bent, M. Preusser, E. Le Rhun, J.C. Tonn, G. Minniti, M. Bendszus, C. Balana, O. Chinot, L. Dirven, P. French, M.E. Hegi, A.S. Jakola, M. Platten, P. Roth, R. Ruda, S. Short, M. Smits, M.J.B. Taphoorn, A. von Deimling, M. Westphal, R. Soffietti, G. Reifenberger, W. Wick, Author Correction: EANO guidelines on the diagnosis and treatment of diffuse gliomas of adulthood, *Nat. Rev. Clin. Oncol.* 19 (5) (2022) 357–358.
- [26] O.M. Henriksen, M. Del Mar Alvarez-Torres, P. Figueiredo, G. Hangel, V.C. Keil, R.E. Nechifor, F. Riemer, K.M. Schmainda, E.A.H. Warnert, E.C. Wiegers, T.C. Booth, High-grade glioma treatment response monitoring biomarkers: a position statement on the evidence supporting the use of advanced MRI techniques in the clinic, and the latest bench-to bedside developments. Part 1: perfusion and diffusion techniques, *Front. Oncol.* 12 (2022), 810263.
- [27] G. Singh, S. Manjila, N. Sakla, A. True, A.H. Wardeh, N. Beig, A. Vaysberg, J. Matthews, P. Prasanna, V. Spektor, Radiomics and radiogenomics in gliomas: a contemporary update, *Br. J. Cancer* 125 (5) (2021) 641–657.
- [28] B.S. Jang, S.H. Jeon, I.H. Kim, I.A. Kim, Prediction of pseudoprogression versus progression using machine learning algorithm in glioblastoma, *Sci. Rep.* 8 (1) (2018), 12516.
- [29] B.S. Jang, A.J. Park, S.H. Jeon, I.H. Kim, D.H. Lim, S.H. Park, J.H. Lee, J.H. Chang, K.H. Cho, J.H. Kim, L. Sunwoo, S.H. Choi, I.A. Kim, Machine learning model to predict pseudoprogression versus progression in glioblastoma using MRI: a multi-institutional study (KROG 18-07), *Cancers (Basel)* 12 (9) (2020).
- [30] Y.Z. Sun, L.F. Yan, Y. Han, H.Y. Nan, G. Xiao, Q. Tian, W.H. Pu, Z.Y. Li, X.C. Wei, W. Wang, G.B. Cui, Differentiation of pseudoprogression from true progression in glioblastoma patients after standard treatment: a machine learning strategy combined with radiomics features from T1-weighted contrast-enhanced imaging, *BMC Med. Imag.* 21 (1) (2021) 17.
- [31] M. Patel, J. Zhan, K. Natarajan, R. Flintham, N. Davies, P. Sanghera, J. Grist, V. Duddalwar, A. Peet, V. Sawlani, Machine learning-based radiomic evaluation of treatment response prediction in glioblastoma, *Clin. Radiol.* 76 (8) (2021) 628 e17–628 e27.
- [32] B.M. Ellingson, P.Y. Wen, T.F. Cloughesy, Modified criteria for radiographic response assessment in glioblastoma clinical trials, *Neurotherapeutics* 14 (2) (2017) 307–320.
- [33] P. Kickingereder, U. Neuberger, D. Bonekamp, P.L. Piechotta, M. Gotz, A. Wick, M. Sill, A. Kratz, R.T. Shinohara, D.T.W. Jones, A. Radbruch, J. Muschelli, A. Unterberg, J. Debus, H.P. Schlemmer, C. Herold-Mende, S. Pfister, A. von Deimling, W. Wick, D. Capper, K.H. Maier-Hein, M. Bendszus, Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma, *Neuro Oncol.* 20 (6) (2018) 848–857.
- [34] Z. Qian, L. Zhang, J. Hu, S. Chen, H. Chen, H. Shen, F. Zheng, Y. Zang, X. Chen, Machine learning-based analysis of magnetic resonance radiomics for the classification of gliosarcoma and glioblastoma, *Front. Oncol.* 11 (2021), 699789.