

## Article

# Multicenter DSC–MRI-Based Radiomics Predict IDH Mutation in Gliomas

Georgios C. Manikis <sup>1,\*</sup>, Georgios S. Ioannidis <sup>1</sup>, Loizos Siakallis <sup>2</sup>, Katerina Nikiforaki <sup>1</sup>, Michael Iv <sup>3</sup>, Diana Vozlic <sup>4,5</sup>, Katarina Surlan-Popovic <sup>4,5</sup>, Max Wintermark <sup>3</sup>, Sotirios Bisdas <sup>2,6,†</sup> and Kostas Marias <sup>1,7,†</sup>

- <sup>1</sup> Computational BioMedicine Laboratory (CBML), Foundation for Research and Technology—Hellas (FORTH), 70013 Heraklion, Greece; geo3721@ics.forth.gr (G.S.I.); kat@ics.forth.gr (K.N.); kmarias@ics.forth.gr (K.M.)
  - <sup>2</sup> Department of Brain Repair and Rehabilitation, Queen Square Institute of Neurology, UCL, London WC1N 3BG, UK; loizos.siakallis.20@ucl.ac.uk (L.S.); s.bisdas@ucl.ac.uk (S.B.)
  - <sup>3</sup> Division of Neuroimaging and Neurointervention, Department of Radiology, Stanford University, Stanford, CA 94305, USA; miv@stanford.edu (M.I.); mwinterm@stanford.edu (M.W.)
  - <sup>4</sup> Department of Radiology, Faculty of Medicine, University of Ljubljana, 1000 Ljubljana, Slovenia; diana.vozlic@kclj.si (D.V.); katarina.surlan@kclj.si (K.S.-P.)
  - <sup>5</sup> Department of Neuroradiology, University Medical Centre, 1000 Ljubljana, Slovenia
  - <sup>6</sup> Department of Neuroradiology, The National Hospital for Neurology and Neurosurgery, University College London NHS Foundation Trust, London WC1N 3BG, UK
  - <sup>7</sup> Department of Electrical & Computer Engineering, Hellenic Mediterranean University, 71410 Heraklion, Greece
- \* Correspondence: gmanikis@ics.forth.gr; Tel.: +30-281-139-1593  
† These authors contributed equally as last authors.



**Citation:** Manikis, G.C.; Ioannidis, G.S.; Siakallis, L.; Nikiforaki, K.; Iv, M.; Vozlic, D.; Surlan-Popovic, K.; Wintermark, M.; Bisdas, S.; Marias, K. Multicenter DSC–MRI-Based Radiomics Predict IDH Mutation in Gliomas. *Cancers* **2021**, *13*, 3965. <https://doi.org/10.3390/cancers13163965>

Academic Editor: Brigitta G. Baumert

Received: 24 June 2021

Accepted: 31 July 2021

Published: 5 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Simple Summary:** Significant efforts have been put toward developing MRI-based radiogenomics for IDH status subtyping predictions; however, in the vast majority of these approaches, the external validation sets are absent. Another limitation in current studies is the lack of explainability in radiomics models, which hampers clinical trust and translation. Motivated by these challenges, we proposed a multicenter DSC–MRI-based radiomics study based on an independent exploratory set, which was externally validated on two independent cohorts, for IDH mutation status prediction. Our results demonstrated that DSC–MRI radiogenomics in gliomas, coupled with dynamic-based image standardization techniques, hold the potential to provide (a) increased predictive performance by offering models that generalize well, (b) reasoning behind the IDH mutation status predictions, and (c) interpretability of the radiomics features' impacts in model performance.

**Abstract:** To address the current lack of dynamic susceptibility contrast magnetic resonance imaging (DSC–MRI)-based radiomics to predict isocitrate dehydrogenase (IDH) mutations in gliomas, we present a multicenter study that featured an independent exploratory set for radiomics model development and external validation using two independent cohorts. The maximum performance of the IDH mutation status prediction on the validation set had an accuracy of 0.544 (Cohen's kappa: 0.145, F1-score: 0.415, area under the curve–AUC: 0.639, sensitivity: 0.733, specificity: 0.491), which significantly improved to an accuracy of 0.706 (Cohen's kappa: 0.282, F1-score: 0.474, AUC: 0.667, sensitivity: 0.6, specificity: 0.736) when dynamic-based standardization of the images was performed prior to the radiomics. Model explainability using local interpretable model-agnostic explanations (LIME) and Shapley additive explanations (SHAP) revealed potential intuitive correlations between the IDH–wildtype increased heterogeneity and the texture complexity. These results strengthened our hypothesis that DSC–MRI radiogenomics in gliomas hold the potential to provide increased predictive performance from models that generalize well and provide understandable patterns between IDH mutation status and the extracted features toward enabling the clinical translation of radiogenomics in neuro-oncology.

**Keywords:** dynamic susceptibility contrast MRI; gliomas; radiomics; IDH mutation; generalizability; explainability; external validation

## 1. Introduction

Gliomas consist of a heterogeneous group of tumors and account for the majority of malignant primary brain tumors [1]. Histological grading and classification into low- and high-grade gliomas [2] has been limited in deciphering the heterogeneous treatment response and survival of glioma patients [3], apart from predicting worse clinical outcomes in patients with glioblastoma [4]. Molecular and genetic analyses yielded distinct biochemical and genetic traits of glioma subgroups beyond histological grading [3]. Of these, isocitrate dehydrogenase (IDH) is probably the most important in terms of prognostication [5]. The identification of an IDH mutation in low- and high-grade gliomas and its role in gliomagenesis elucidate the progression of low-grade glioma (LGG) to an IDH-mutant, “secondary” glioblastoma via a common pathway [6]. While IDH-mutant tumors are favored for their improved clinical outcomes, their IDH-mutant counterparts are characterized by a dismal prognosis, irrespective of histological grading [7]. Due to its propensity to stratify glioma patients, IDH mutations were identified in the recent world health organization (WHO) classification as a distinctive feature of glioma subgroups [3].

Characterization of the IDH mutation status has therefore become imperative prior to treatment selection and patient stratification [8]. However, this still relies on invasive tissue sampling and analysis, harboring inherent limitations, including post-operative complications and sample error [9]. The non-invasive classification of gliomas based on imaging was proposed as an alternative to this approach, leading to the identification of imaging features as correlates to molecular subtypes. Structural features, including calcification and T2-FLAIR mismatch and tumor enhancement, yielded promising, though inconsistent, outcomes for the identification of IDH status [10]. This was partly attributed to interobserver variability during the visual assessment of such features [11]. To this end, machine learning (ML) enhanced the identification of high-throughput quantitative imaging features that were extracted from conventional and multiparametric magnetic resonance imaging (MRI) was successfully applied for the characterization of IDH mutation status, with promising outcomes outlined in recent radiogenomics studies [12]. Results that were found using contrast-enhanced T1-weighted, T2-weighted, and arterial spin labeling (ASL) images from 105 glioma patients showed an accuracy of 0.82 and an area under the curve (AUC) of 0.77 on internal validation using an iterated cross-validation [13]. A recent review of radiomics implications in glioma [14] produced an accuracy of 0.83 when multicenter multiparametric MRI samples were divided randomly into training and validation sets [15], and an accuracy of 0.99 was found when T2-weighted MRI-based radiomics were applied to training and validation sets from a single center [16]. An accuracy of 0.917 (sensitivity: 0.857, specificity: 1) was achieved in [17] using a three-level radiomics model based on multiparametric MRI (postcontrast T1-weighted, T2-weighted, T2-FLAIR, and diffusion-weighted imaging) obtained from the Cancer Imaging Archive (TCIA) and validated on a cohort of 12 patients. Finally, a random forest showed an accuracy of  $0.885 \pm 0.041$  and an AUC equal to  $0.931 \pm 0.036$  when tested on multiparametric MRI data using an iterated training/testing split from 126 patients [18]. Although several studies, as presented in a recent review paper [12] managed to assess the value of conventional and multiparametric MRI in predicting glioma IDH mutation status, a lack of independent sets for validating externally related radiomics signatures was evident in most of the studies, as denoted by the radiomics quality score (RQS) [19]. This limitation also concerns neuro-oncology radiomics in general, where only 3.9% (2/51) of studies were found to be validated using multicenter data [20].

Potential imaging biomarkers beyond structural features were highlighted by the long recognized tumor vascularity’s correlation with histological grade [21], and the role of dynamic susceptibility contrast (DSC) perfusion was noted by the recent identification of distinct angiogenesis transcriptome signatures of IDH-mutant gliomas linked to identifiable perfusion phenotypes [22]. However, not much research has investigated the integrative role of DSC perfusion, as opposed to standalone structural MRI, in radiogenomics toward the non-invasive characterization of IDH status based on distinct vascular profiles [22].

More specifically, to the best of our knowledge, IDH status mutation prediction is limited to a single study from [23] in which repeated cross-validations, performed on a unique cohort of six tertiary centers, reported an accuracy of 0.72 (overall specificity: 0.77, sensitivity: 0.65). This was further confounded by variations in DSC perfusion acquisition and analysis, despite current attempts for consensus guidelines [24]. Another limitation in current studies is the lack of explainability in radiomics models, which hampers clinical trust and translation to neuro-oncology clinical practice.

In an effort to address all the concerns stated above, this study aimed to explore the potential of DSC–MRI-based radiomics to predict IDH mutation status in gliomas. The proposed radiogenomics analysis was built using an independent exploratory set, a dynamic-based standardization of the MR images, and an external validation from two independent cohorts. To interpret the IDH mutation status predictions, model explainability was performed using local interpretable model-agnostic explanations (LIME) and Shapley additive explanations (SHAP). To the best of our knowledge, this is the first attempt toward validating a DSC–MRI radiogenomics study on independent external cohorts and interpreting the corresponding IDH mutation status prediction results for gliomas.

## 2. Materials and Methods

### 2.1. Patient Population

The studied population included 160 patients (age:  $58.4 \pm 15.9$  (mean  $\pm$  SD), 70 female) from 3 tertiary centers, each with a histopathological diagnosis of primary glioma (WHO grades 2–4), molecular characterization of IDH mutation status (IDH-mutant = 41, IDH-wildtype = 119) and DSC–MRI data prior to any treatment. Specifically, cohort A consisted of 92 patients (66 out of 92 IDH-mutant), cohort B included 50 patients (39 out of 50 IDH-mutant), and cohort C contained 14 out of 18 patients with an IDH-mutant status. We excluded patients without a histologically confirmed diagnosis of glioma, incomplete molecular characterization of IDH status, non-enhancing anaplastic grade III gliomas, or patients having received any treatment prior to image acquisition. The study was approved by each local ethics committee and all patients signed informed consent forms prior to examination.

### 2.2. MRI Protocol

Patients underwent DSC–MRI from different scanners and acquisition protocols. The imaging parameters for each cohort are presented in Table 1.

**Table 1.** Imaging protocol parameters for the patient population.

Scanner Type	Cohort A		Cohort B		Cohort C
	3T Discovery MR750 (GE Healthcare, Chicago, Illinois, United States)	3T Siemens Skyra (Siemens Healthineers, Erlangen, Germany)	1.5T Magnetom Avanto (Siemens Healthineers, Erlangen, Germany)	1.5T Philips Achieva (Philips Medical Systems, Eindhoven, Netherlands)	
Acquisition type	2D Echo-Planar Imaging (EPI) with fat suppression (FS)				
Magnetic field strength	3T	3T	1.5T	1.5T	
Repetition time (ms)	1800	1870	1850	1525	
Echo time (ms)	40	30	30	40	
Echo train length	1	63	1	47	
Flip angle (deg)	60	90	90	75	
In-plane resolution (mm <sup>2</sup> )	1.718 $\times$ 1.718	1.719 $\times$ 1.719	1.796 $\times$ 1.796	1.75 $\times$ 1.75	
Number of averages	1	1	1	1	
Image slice thickness (mm)	5	5	5	5	
Image slice spacing (mm)	5	5	5	5	
Temporal resolution	60 $\times$ 1.87 s	60 $\times$ 2.07 s	40 $\times$ 1.53 s	60 $\times$ 1.8 s	
Matrix size	128 $\times$ 128	128 $\times$ 128	128 $\times$ 128	128 $\times$ 128	

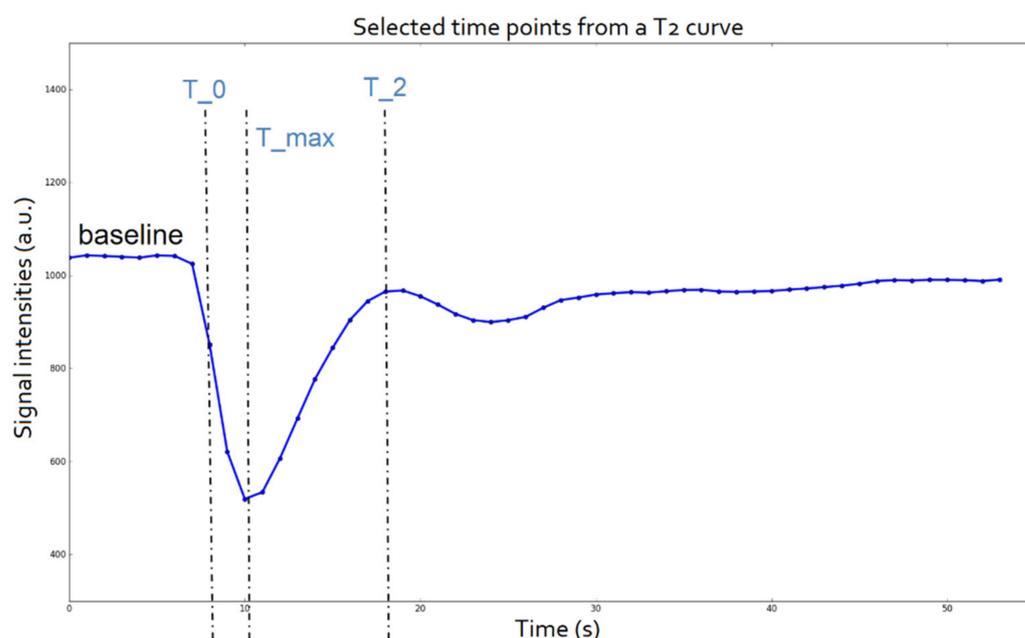
### 2.3. Tumor Delineation

Lesion voxels of interest (VOIs) were automatically generated using Bratumia [25] software (<https://www.nitrc.org/projects/bratumia>, accessed on 24 June 2021), which uses four different contrasts (T2, T2 FLAIR, T1 before and after contrast) as input in order to

correctly identify tumor enhancement, edema, and necrosis from normal brain tissue. Two expert neuroradiologists with 7 and 4 years of experience, respectively, visually inspected the generated tumor voxels of interest (VOIs) and provided corrections when necessary on a consensus basis.

#### 2.4. Time Point Selection

After the tumor delineation, the mean signal intensity from all voxels assigned as a lesion were used to create the signal intensity curve of each patient over time. DSC-MRI data were inspected for an adequate number of time points to establish the baseline (minimum 5) before the maximum signal drop resulting from the arrival of the contrast agent to the voxels in question. Three characteristic points of interest ( $T_0$ ,  $T_{max}$ , and  $T_2$ ), defined as the time of the first measurement after contrast detection, time of maximum signal drop, and time of return to baseline, respectively, were selected for each specific patient from their corresponding signal intensity curve over time (Figure 1).



**Figure 1.** Time points of interest of a T2 mean signal intensity curve.

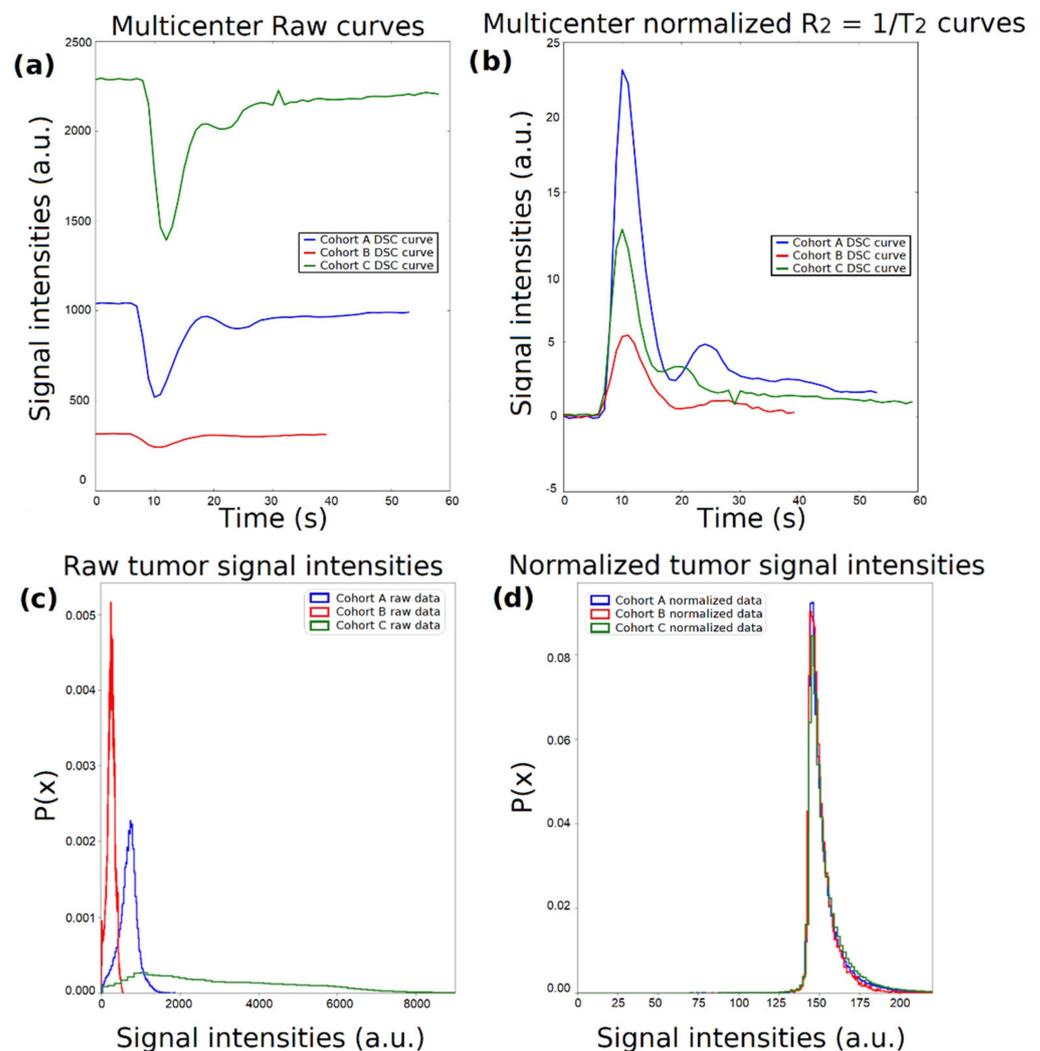
#### 2.5. Dynamic-Based Standardization of the MRI Data

Considering that the MR signal is measured in arbitrary units, it is challenging to perform joint analyses from data originating from different clinical centers since the image intensities exhibit inherent variability and different ranges of values. Indicative perfusion curves from each patient cohort are presented in Figure 2a, where the differences in the range of the exported MR image intensity values are evident. Supposing a DSC perfusion curve over time  $T_2(t)$  or the relaxation rate curve  $R_2(t) = 1/T_2(t)$ , the first step was the baseline correction, which was achieved by subtracting the last signal of the baseline ( $T_2(0)$ , contrast agent enters the vasculature system) from each curve as follows:  $\Delta T_2(t) = T_2(t) - T_2(0)$  (Figure 2b). As a second step, due to the variability between different protocols with respect to echo time ( $TE$ ), each curve was modified to account for the exponential signal decay in time by using Equation (1):

$$\Delta R_2(t) = \frac{1}{TE} \ln \frac{\overline{T_2(t)}_{baseline}}{\Delta T_2(t)} \quad (1)$$

where  $\overline{T_2(t)}_{baseline}$  is the mean baseline of the raw signal  $T_2(t)$  [26]. Moreover, signal intensity probability density functions for all voxels annotated as a tumor per imaging

group of patients for the maximum enhancement time point are presented for the raw (Figure 2c) and the normalized (Figure 2d)  $T_2(t)$  data.



**Figure 2.** Schematic representation of the normalization process. (a) Dynamic susceptibility contrast (DSC) perfusion curves from the 3 tertiary centers. (b)  $\Delta R_2(t)$  curves after normalization using equation (1). (c) Center-specific probability density functions of the raw signal intensities within the voxels of interest (VOIs) at time point  $T_{max}$ . (d) Center-specific probability density functions of the normalized signal intensities within the VOIs at time point  $T_{max}$ .

## 2.6. Radiomics Analysis Workflow

The proposed analysis workflow is illustrated in Figure 3 and presented in detail below.

### 2.6.1. Image Postprocessing

Initially, tumor delineations were automatically generated as reported in Section 2.3 (step A, Figure 4) and corresponding 3D images at the specific time points  $T_0$ ,  $T_{max}$ , and  $T_2$  were extracted from the dynamic acquisition data (Section 2.4 and step B from Figure 4). Subsequently, 3D images were spatially resampled and interpolated (step C, Figure 4) to compensate for multicenter effects (e.g., differences in the acquisition protocol and the reconstruction settings), enabling textural feature extraction in the 3D domain since rotation-invariant 3D images are recommended. Although there is no clear suggestion of whether to use a downsampling or upsampling interpolation, volumes in this study were resampled to isotropic voxels of length 1 mm using the cubic B-spline interpolation method.

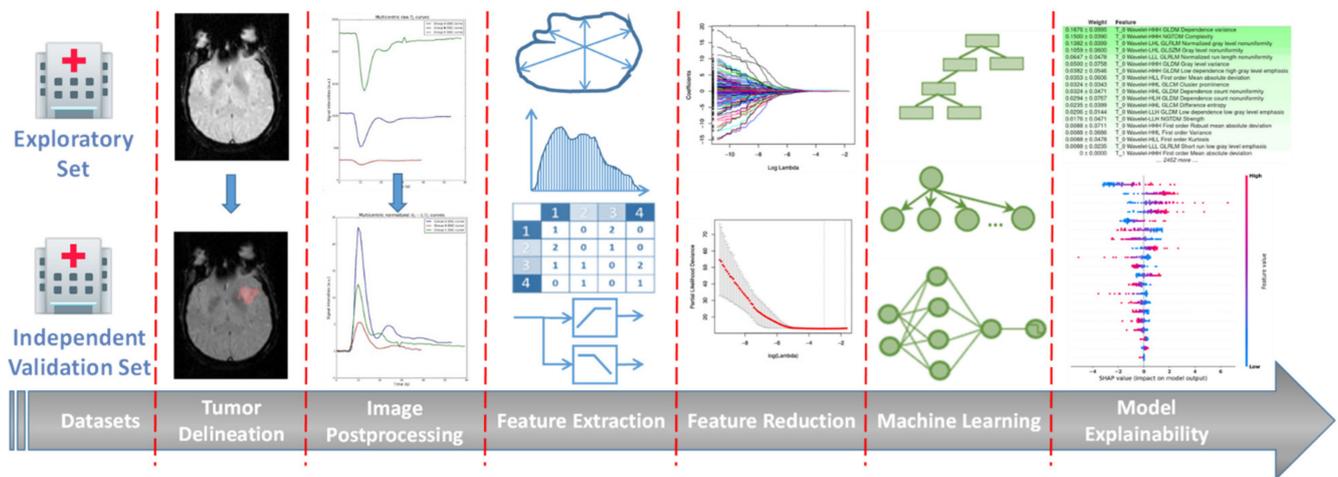


Figure 3. An illustrative representation of the proposed radiomics analysis workflow.

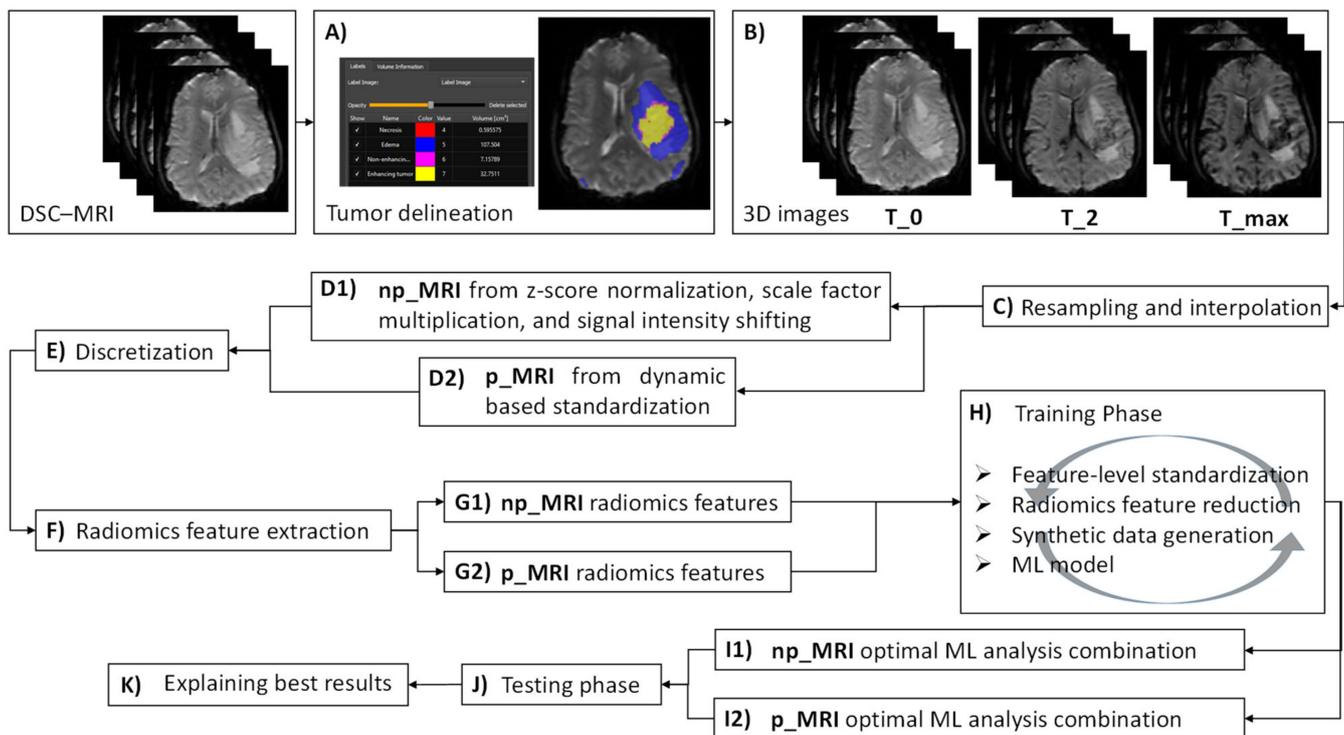


Figure 4. A schematic diagram summarizing the proposed radiomics analysis workflow steps.

To assess the potential impact of the proposed standardization technique (Section 2.5) on the IDH status prediction, two different versions of the extracted 3D images were generated. In the first version (step D1, Figure 4), the signal intensities from the 3D images (hereafter denoted as np\_MRI) were z-score normalized using the mean and standard deviation from all voxels within the VOIs drawn on the overall DSC-MRI data (VOIs drawn across all time points to keep the dynamic changes of the tissue over time). Next, normalized signals from np\_MRI data were multiplied by a scaling factor that was set to 100 and shifted by the minimum signal intensity that was derived from np\_MRI histograms to ensure that all voxels within the VOIs had positive values. In the second version (step D2, Figure 4), the 3D images (hereafter denoted as p\_MRI) were dynamically standardized (Section 2.5). At the next step, both np\_MRI and p\_MRI data underwent discretization (step E, Figure 4) to account for differences in the voxel intensity ranges using Equation (2),

where  $FBS$  is the fixed bin size,  $FBN$  is the fixed bin number, and  $mean_{Range}$  is the derived mean intensity range from the VOIs histogram of all patients within the examined cohort. In the current analysis,  $FBN$  was set to 32 and 64 since a low number of bins is recommended in brain MRI radiomics [27].

$$FBS = \frac{1}{FBN} \times mean_{Range} \quad (2)$$

### 2.6.2. Radiomics Feature Extraction

Radiomics feature extraction (step F, Figure 4) was carried out in compliance with the image biomarker standardization initiative (IBSI) guidelines. A total of 833 features were automatically extracted from the segmented regions of every image (both np\_MRI and p\_MRI) using pyradiomics [28] comprising features: (a) without any filter applied before the extraction and (b) after images were decomposed using a wavelet transform (Coiflet 1) of level 1. Among the non-wavelet features (105 features, summarized in Supplementary Table S1): 16 first-order statistics were calculated from the histogram of the VOIs (group 1); 14 morphological features contained information about the shape and size of the lesion (group 2); 24 s-order statistics (group 3) computed from the grey level co-occurrence matrix (GLCM); and 14, 16, 16, and 5 third-order statistics on the gray level dependence matrix (GLDM), gray level size zone matrix (GLSZM), gray level run length matrix (GLRLM), and neighboring gray tone difference matrix (NGTDM), respectively (group 4). Additionally, to address the multi-scale nature of texture, 728 features from groups 1, 3, and 4 were derived using a set of eight sub-bands of the decomposed images (HHH, HHL, HLH, HLL, LHH, LHL, LLH, and LLL) with different resolutions corresponding to the low- (L) and high-pass (H) filters applied along the X-, Y, and Z-axis (91 features  $\times$  8 sub-bands). In total, 2499 (833  $\times$  3, where three corresponds to the three time points) np\_MRI (step G1, Figure 4) and 2499 p\_MRI related radiomics (step G2, Figure 4) features were integrated into a single feature vector, respectively.

### 2.6.3. Machine Learning Analysis Framework

A machine learning analysis pipeline was designed comprising: (a) feature-level standardization, (b) radiomics feature reduction, (c) synthetic data generation, and (d) the application of the ML model. Concerning (a), a univariate analysis was initially conducted using the Shapiro–Wilk test to assess the normality of the radiomics features. Most of the features achieved a  $p$ -value less than 5%, indicating a non-normal distribution. Therefore, centering and scaling of all exported radiomics features was performed with “RobustScaler” [29] since it is less prone to outliers and can operate on non-normal distributed data. Several feature reduction techniques performed during (b) were utilized to decrease the high-dimensional radiomics feature vector by considering their low computational complexity and their effectiveness, as reported in similar radiomics studies. These included univariate filtering techniques (Gini Index, F-ratio, information gain, and Spearman correlation) and a multivariate feature scoring criterion using: (a) the minimum-redundancy maximum-relevance (mRMR), (b) mutual information maximization (MIM), and (c) the statistical inference relief (STIR); all of these were tested with a pre-selected number of features varying from 5 to 50. Additionally, the least absolute shrinkage and selection operator (LASSO) regression was also examined to identify a subset of informative radiomics features (features with nonzero coefficients) using the L1 norm penalty and a logistic regression with elastic net regularization, combining the LASSO and ridge regression with a ratio of 0.5 between the L1 and L2 penalties.

Considering the highly unbalanced ratio between the two examined classes, different synthetic data generation techniques based on data oversampling, undersampling, and a combination of over- and undersampling were applied during the model training to manage the class unbalancing. To this end, five different approaches were examined: synthetic minority oversampling technique (SMOTE), adaptive synthetic (ADASYN), SMOTE followed by cleaning using Tomek links (SMOTE-Tomek), random undersampling (RUS), and undersampling using the Nearmiss method (NearMiss). Additionally, a no-resampling

choice was added as an option to the analysis pipeline. For the application of the ML model, the current study utilized a support vector machine (SVM) with a linear kernel; a random forest (RF); a K-nearest neighbor (KNN); a logistic regression with L2, elastic net, and L1 norm penalties; and adaptive boosting (AdaBoost) using a decision tree as the base classifier. Additional ensemble learning techniques were initiated, comprising a RUSBoost (an AdaBoost coupled with a random undersampling) with a decision tree as the base classifier, a gradient boost (XGBoost), and bagged classification trees (BAG) integrated with resampling of the majority class (IDH-mutant). When RUSBoost and BAG were operated, no resampling was applied. Feature-level optimization was implemented using scikit-learn [29]; radiomics feature reduction was developed using scikit-learn [29] and ITMO FS ([https://github.com/ctlab/ITMO\\_FS](https://github.com/ctlab/ITMO_FS), accessed on 24 June 2021); imbalanced-learn [30] was utilized for synthetic data generation; and packages xgboost (<https://github.com/dmlc/xgboost>, accessed on 24 June 2021), scikit-learn [29], and imbalanced-learn [30] contributed in the development of the ML models.

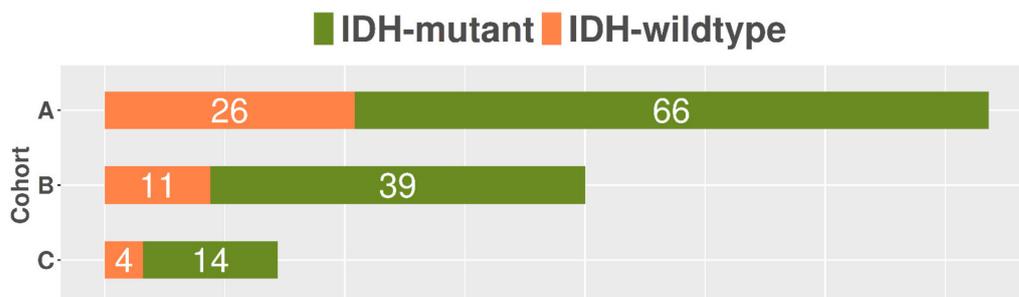
#### 2.6.4. Performance Validation and Explainability of the Predictions

The ML analysis framework (Section 2.6.3) was initially applied to both the np\_MRI and p\_MRI data from cohort A (exploratory set) and subsequently validated externally using np\_MRI and p\_MRI data from cohorts B and C (validation set). An internal validation according to a stratified five-fold repeated cross-validation with two hundred repeats was conducted during the training phase using the exploratory set (step H, Figure 4). Identical results regarding the samples of each fold were subsequently used from every possible combination of the analysis steps in both cases (np\_MRI and p\_MRI). An average estimate of Cohen's kappa metric was then obtained at the validation phase (within the internal validation) as a criterion to determine the highest performance among all combinations. Cohen's kappa was chosen as the preferable performance metric to select the optimal ML analysis combination since it considers class unbalancing in the prediction accuracy. To eliminate any bias during training, the feature-level standardization, radiomics feature reduction, and synthetic data generation from Section 2.6.3 were applied exclusively at the training phase. In all cases, the ML models defined in Section 2.6.3 were initialized using the default parameters. For both the np\_MRI and p\_MRI data, the combination that showed the highest performance with respect to the Cohen's kappa (steps I1 and I2 from Figure 4) was then tested in the independent test set (step J, Figure 4). Model performance was also assessed using the F1-score, accuracy, AUC, sensitivity, specificity, positive and negative predictive value (PPV and NPV). The IDH-wildtype was considered as the positive class. Apart from constructing generalizable ML models, the explainability and interpretability of their decisions are the key milestones of increased impact toward improving the clinical translation of radiogenomics studies. To this end, the present study developed a postprocessing phase using SHAP [31] and LIME [32] methodologies to provide reasoning behind the IDH mutation status predictions and interpret the impact of each radiomics feature in the model's performance (step K, Figure 4). According to LIME, once a model is trained on the exploratory set, feature permutation importance (reported as mean  $\pm$  SD) is then computed on the validation set by shuffling all corresponding features individually. In our analysis, this yielded a top-down radiomics feature order in decreasing importance, which was related to the reshuffling of each particular feature from the ordered list. To further illustrate the interpretability of the radiomics analysis pipeline that showed the highest predictive performance regarding the IDH status, a SHAP explainer was designed on the exploratory set and the distribution of features' impact on the predictions was then calculated based on the validation set and displayed using the SHAP summary plot.

### 3. Results

A multicenter radiogenomics study was conducted using imaging data that was acquired from different vendors, who used different magnetic strengths and acquisition

protocols. Overall, cohort A included 60 males and 32 females with a mean and standard deviation (SD) age of 57 (15.1) years, cohort B was composed of 28 males and 22 females (age in years: mean = 58, SD = 17.1), and 9 males and 9 females belonged to cohort C with a mean age of 55.7 years and a standard deviation equal to 15 years. No statistically significant changes were observed between the ages of the three investigated cohorts. A high-class imbalance of the IDH status of the patients was evident in all examined cohorts, as depicted in Figure 5.



**Figure 5.** Distribution of the isocitrate dehydrogenase (IDH) status across the three examined centers. Numbers of total IDH-mutant and IDH-wildtype statuses from each clinical center are shown in white.

Several combinations of all techniques reported in Section 2.6.3 were developed using a total of 2499 radiomics features from cohort A and validated on cohorts B and C using Cohen’s kappa metric. Combinations having a mean Cohen’s kappa equal to or lower than the baseline (i.e., a zero value in the case of random model’s accuracy) were discarded from the validation phase. In the case of the radiomics analysis on the np\_MRI, the best model training performance was obtained when the features were:

1. Normalized with “RobustScaler”;
2. Reduced to the 15 most significant features using univariate filtering according to the Gini Index;
3. Subjected to synthetic minority oversampling based on the ADASYN;
4. Providing input to an elastic net logistic regression model, yielding a mean Cohen’s kappa of 0.193 (SD: 0.096) using the exploratory set and 0.145 when validated externally.

The p\_MRI radiomics analysis performance was found to be significantly higher, yielding a Cohen’s kappa for the exploratory and the validation set equal to  $0.303 \pm 0.125$  (mean  $\pm$  SD) and 0.282, respectively. Among all possible combinations, the optimal p\_MRI radiomics analysis pipeline was constructed from “RobustScaler”, which is a multivariate feature scoring criterion that is based on MIM; an ADASYN synthetic minority oversampling; and a logistic regression classifier with an L2 penalty. Increased performance was also evident on any calculated performance metric, except from the sensitivity (Table 2 and Figure 6).

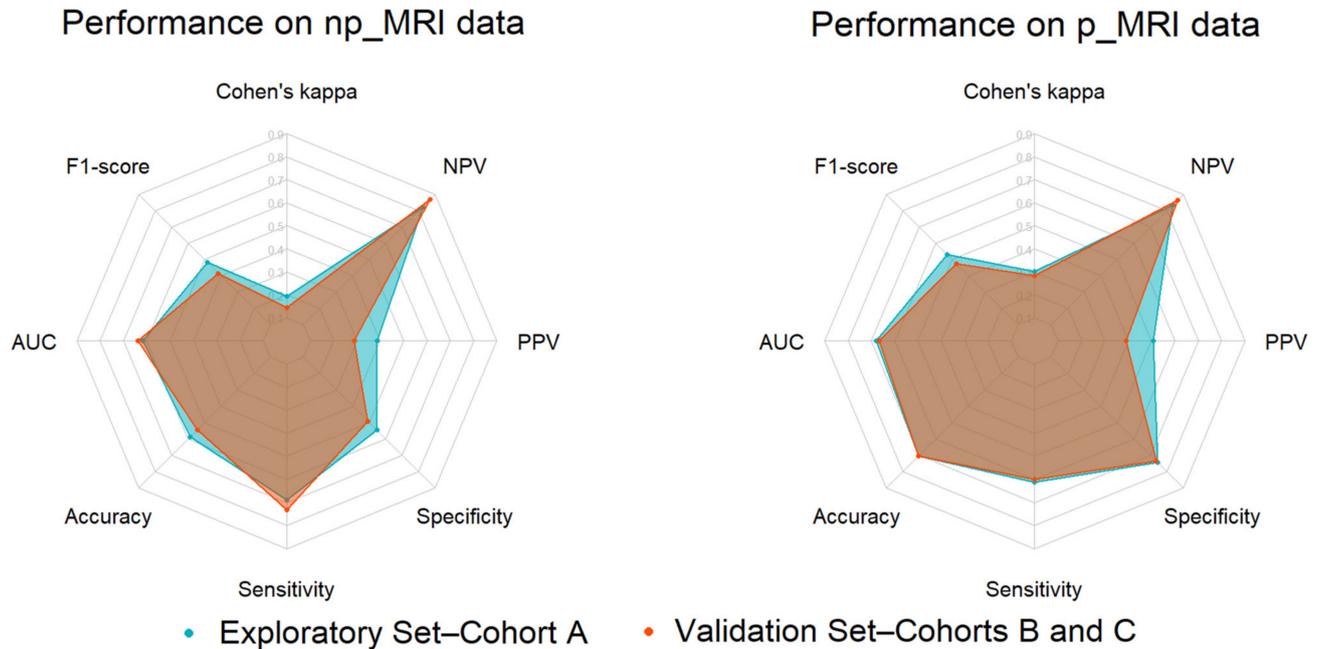
**Table 2.** Performance of np\_MRI and p\_MRI radiomics in isocitrate dehydrogenase (IDH) mutation status prediction.

Performance Metrics	Exploratory Set		Validation Set	
	np_MRI	p_MRI	np_MRI	p_MRI
Cohen’s kappa	0.193 $\pm$ 0.096	0.303 $\pm$ 0.125	0.145	0.282
F1-score	0.482 $\pm$ 0.163	0.533 $\pm$ 0.164	0.415	0.474
AUC	0.618 $\pm$ 0.148	0.678 $\pm$ 0.103	0.639	0.667
Accuracy	0.587 $\pm$ 0.144	0.707 $\pm$ 0.108	0.544	0.706
Sensitivity	0.689 $\pm$ 0.208	0.611 $\pm$ 0.183	0.733	0.600
Specificity	0.547 $\pm$ 0.189	0.745 $\pm$ 0.146	0.491	0.736
PPV	0.388 $\pm$ 0.150	0.508 $\pm$ 0.196	0.290	0.391
NPV	0.823 $\pm$ 0.158	0.834 $\pm$ 0.095	0.868	0.868

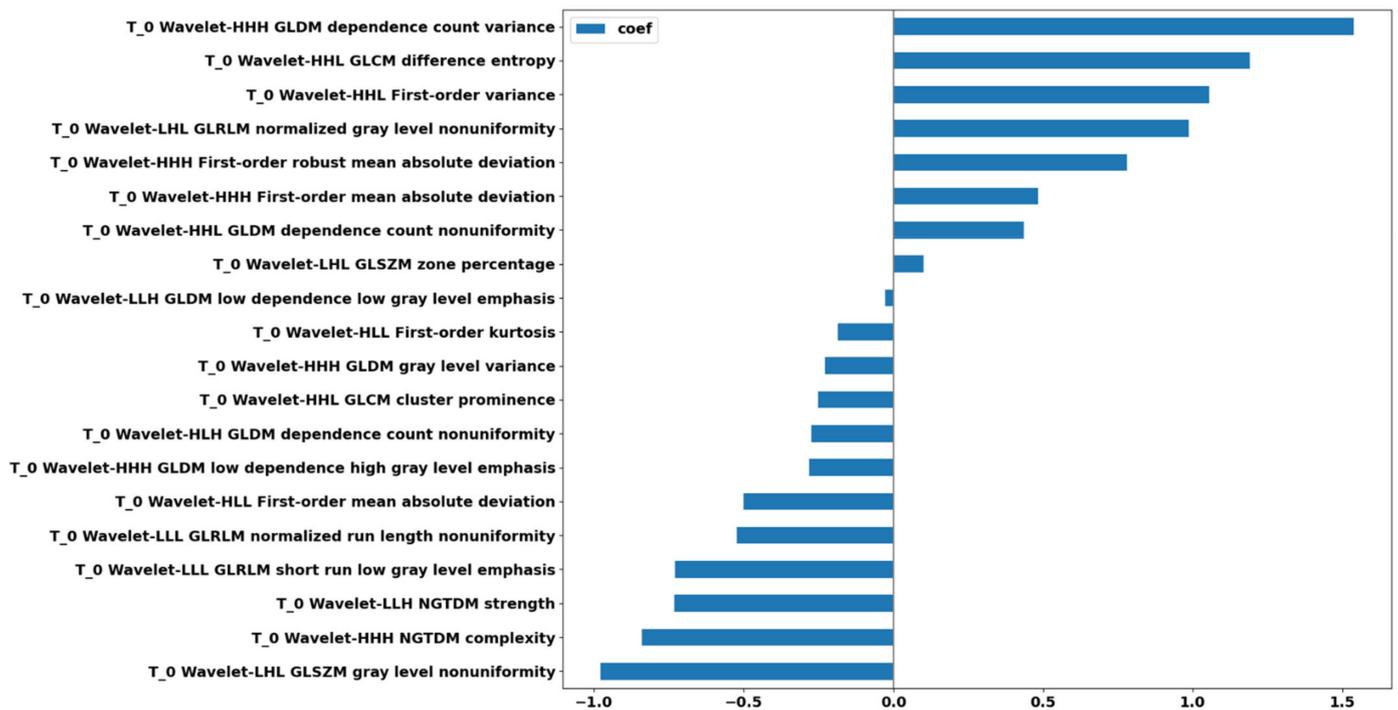
AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value.

Additionally, only subtle changes were presented between the exploratory and validation performances, as illustrated by a notable overlap on their polygon areas drawn using their corresponding radar plots (Figure 6).

Of the 2499 p\_MRI-based radiomics features, 20 features were found to be significant; all of them contributed to the prediction performance according to the L2 regularization of the logistic regression (Figure 7).



**Figure 6.** From left to right: Radar plots illustrating the IDH mutation status prediction performance during the model development (internal validation using the exploratory set) and validation (external validation set) phases from np\_MRI and p\_MRI data.



**Figure 7.** Feature ranking of the 20 selected radiomics features after feature reduction based on the regression coefficient profiles.

The descriptive statistics of the 20 selected features (Figure 8) depict the contribution of solely wavelet-based radiomics features extracted from images at time point T\_0.

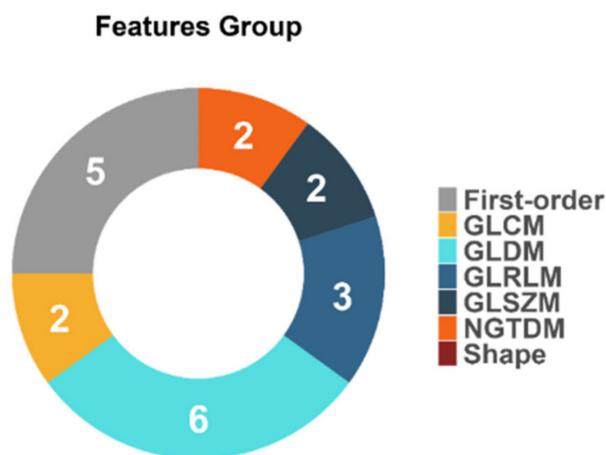


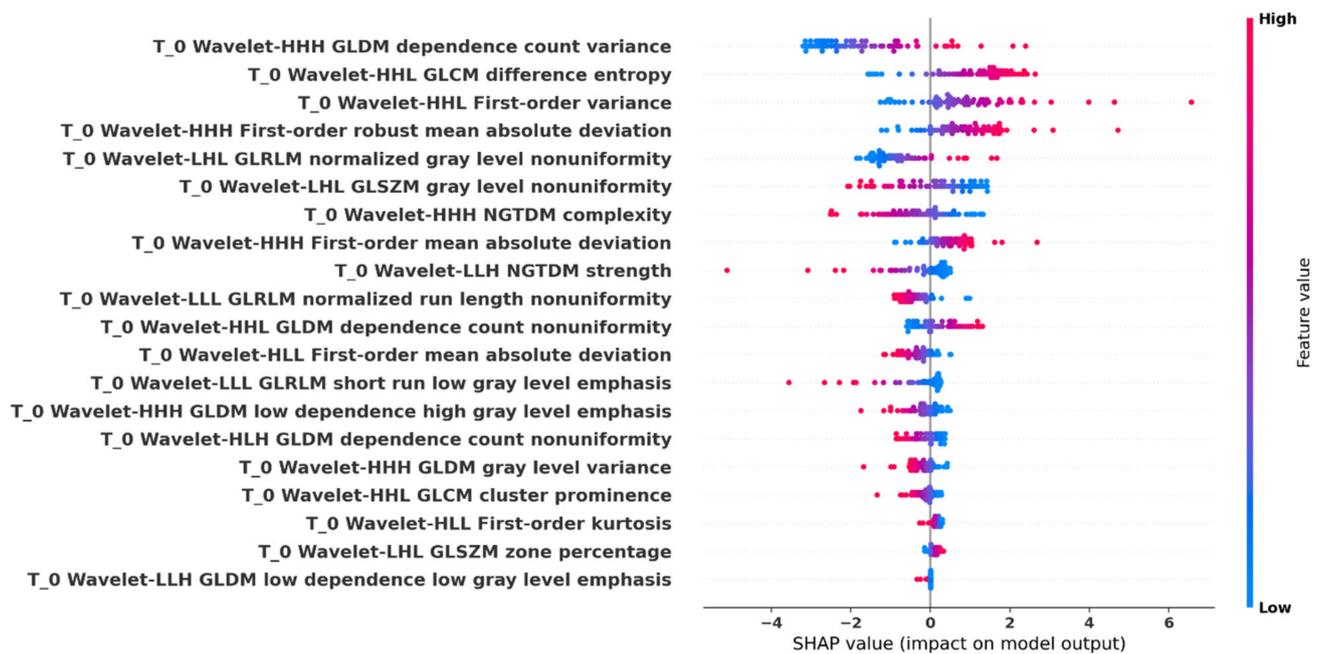
Figure 8. Feature distributions according to the group of the calculated radiomics features.

Global model interpretability of the optimal p\_MRI-based radiomics analysis pipeline with regard to the prediction performance was obtained using LIME (Figure 9). Summary plot of the corresponding SHAP values is shown in Figure 10.

Weight	Feature
0.1676 ± 0.0995	T_0 Wavelet-HHH GLDM dependence count variance
0.1500 ± 0.0390	T_0 Wavelet-HHH NGTDM complexity
0.1382 ± 0.0399	T_0 Wavelet-LHL GLRLM normalized gray level nonuniformity
0.1059 ± 0.0600	T_0 Wavelet-LHL GLSZM gray level nonuniformity
0.0647 ± 0.0478	T_0 Wavelet-LLL GLRLM normalized run length nonuniformity
0.0500 ± 0.0758	T_0 Wavelet-HHH GLDM gray level variance
0.0382 ± 0.0546	T_0 Wavelet-HHH GLDM low dependence high gray level emphasis
0.0353 ± 0.0606	T_0 Wavelet-HLL First-order mean absolute deviation
0.0324 ± 0.0343	T_0 Wavelet-HHL GLCM cluster prominence
0.0324 ± 0.0471	T_0 Wavelet-HHL GLDM dependence count nonuniformity
0.0294 ± 0.0767	T_0 Wavelet-HLH GLDM dependence count nonuniformity
0.0235 ± 0.0399	T_0 Wavelet-HHL GLCM difference entropy
0.0206 ± 0.0144	T_0 Wavelet-LLH GLDM low dependence low gray level emphasis
0.0176 ± 0.0471	T_0 Wavelet-LLH NGTDM strength
0.0088 ± 0.0711	T_0 Wavelet-HHH First-order robust mean absolute deviation
0.0088 ± 0.0686	T_0 Wavelet-HHL First-order variance
0.0088 ± 0.0478	T_0 Wavelet-HLL First-order kurtosis
0.0088 ± 0.0235	T_0 Wavelet-LLL GLRLM short run low gray level emphasis
0 ± 0.0000	T_1 Wavelet-HHH First-order mean absolute deviation
	... 2452 more ...

Figure 9. A top-down radiomics feature order in decreasing importance using local interpretable model-agnostic explanations (LIME). All radiomics features were shuffled individually and the permutation importance (reported as mean ± SD) was computed on the external validation set.

Assessment of the methodological quality of the presented radiomics study was conducted using the radiomics quality score (RQS), reaching a total of 14 out of 36 points (38.9%). Details are given in Table S2.



**Figure 10.** Shapley additive explanations (SHAP) summary plot. All important radiomics feature values are displayed as dots using a pseudocoloring (blue to red), low or zero contributors are near a SHAP value of zero, a long distance from zero denotes a higher influence of a specific feature in the prediction performance, where decreased or increased values favored the negative (IDH-mutant) or positive (IDH-wildtype) classes, respectively.

#### 4. Discussion

This study investigated the prognostic value of DSC–MRI radiogenomics in gliomas regarding IDH mutation status prediction. Data from three different clinical centers (cohorts A, B, and C) were involved in the study; ML models were developed using cohort A and validated externally to demonstrate their generalizability on the independent cohorts B and C.

##### 4.1. The Role of DSC–MRI Standardization

Radiogenomics IDH mutation status prediction is a promising but also challenging task that needs careful consideration when designing all the required steps of the analysis workflow. When initiating a multicenter radiomics study, the lack of a direct or modeled relationship between tissue and the MR signal intensities can yield a radiomics feature extraction from images that varies between different vendors and acquisition protocols. Published studies on phantoms have reported intensity-related bias that was measured to approach 100% in quantitative (pseudo-perfusion data) phantom studies due to the inability to correct for all image acquisition factors using postprocessing tools [33]. To address this issue, a dynamic-based MRI standardization was performed on perfusion data (p\_MRI) obtained from T<sub>0</sub>, T<sub>max</sub>, and T<sub>2</sub> prior to radiomics and compared to the commonly used image postprocessing techniques (np\_MRI). The performance results, outlined in the following text, showed that the radiomics from the proposed standardization method outperformed that of commonly-used techniques.

##### 4.2. Performance of IDH Status Prediction and Informative Radiomics Features

A total of 2499 np\_MRI and p\_MRI based radiomics features, in compliance with IBSI guidelines, were calculated and subjected to the proposed analysis pipeline, revealing increased performance when using the proposed image standardization technique. Specifically, by applying a logistic regression with an L2 norm to the selected p\_MRI-based radiomics features of the validation set, apart from the sensitivity in which a decrease of 18.14% was evident (np\_MRI: 0.733, p\_MRI: 0.600), we achieved a percentage change

increase of 94.48% in the Cohen's kappa metric (np\_MRI: 0.145, p\_MRI: 0.282), 14.22% in the F1-score (np\_MRI: 0.415, p\_MRI: 0.474), 4.38% in the AUC (np\_MRI: 0.639, p\_MRI: 0.667), 29.78% in the accuracy (np\_MRI: 0.544, p\_MRI: 0.706), and 49.90% in the specificity (np\_MRI: 0.491, p\_MRI: 0.736). Especially in the case of specificity, the np\_MRI-based radiomics performance in the validation phase was lower than the baseline of 50%.

Feature selection on p\_MRI data revealed a vital role of time point T<sub>0</sub> when making IDH status predictions since all 20 significant prognostic factors were calculated from images at that time point of the perfusion curve. This observation is in accordance with recent studies highlighting the importance of the initial components of the signal time-intensity curve for IDH status determination [34]. Signal deviation in these regions correlates to predictable vascularity phenotypes of IDH-mutant and IDH-wildtype tumors [35], as determined by IDH-specific vascular gene signatures [36]. Specifically, Choi et al. [34] demonstrated that the curve components between T<sub>0</sub> and T<sub>max</sub> performed best for the identification of IDH-mutant versus IDH-wildtype gliomas by applying an explainable recurrent neural network to attribute attention weights to specific parts of the T2\* signal time-intensity curve. IDH-wildtype tumors were characterized by a steeper downslope and larger signal drop in contrast to their IDH-mutant counterparts, which was compatible with the increased tumor angiogenesis in this group [22]. In the same study, the upslope of the signal drop (between T<sub>max</sub> and T<sub>2</sub>) and post-bolus plateau (beyond T<sub>2</sub>) revealed a less steep and attenuated signal recovery in IDH-wildtype tumors, which likely reflected the increased permeability due to immature and leaky tumor vessels [22,34]. The latter observation was less apparent in our analysis, potentially due to the higher percentage of IDH-wildtype tumors in our patient population, which commonly confounds relevant studies and is in accordance with tumor epidemiology [1]. Such minuscule time-dependent differentiating traits between the two tumor groups regarding MR perfusion are potentially enhanced by MR data decomposition based on wavelet transform.

Another important observation was that all 20 significant predictors were calculated from decomposed MRI data after the wavelet transform. The inclusion of multi-resolution radiomics introduces band-pass frequency filtering encompassing smooth and detailed parts of the images. This result indicates that wavelet image decomposition enhanced the discriminatory power of radiomics, possibly due to the noise suppression and identification of frequency bands that better capture the actual imaging phenotype variability within the heterogeneous dataset. Our results confirmed the scale-dependent nature of texture and highlight its importance in multicenter studies, even if image standardization is included. Of the 20 selected wavelet-based features, 5 were from the first-order group and 15 were texture features (2 GLCM, 6 GLDM, 3 GLRLM, 2 GLSZM, and 2 NGTDM). Apparently, no shape descriptors were found to be important in the prediction. These results are in accordance with recent literature, highlighting the strong impact of the texture-based features in glioma IDH mutation status prediction [12]. According to the L2 logistic regression coefficient profiles, the feature ranking of these 20 radiomics features during the model development identified a significant predictive role of texture features that are sensitive to gray level variability, texture complexity, and non-uniformity. The most highly ranked radiomics features were found to be "dependence count variance" from GLDM, "difference entropy" from GLCM, and first-order "variance".

#### 4.3. Radiomics-Based ML Models' Explainability

In light of the above observation, we applied model explainability mechanisms on the external validation phase using LIME and SHAP to evaluate and interpret the influence of the selected radiomics features in the radiomics analysis result. LIME verified the aforementioned results with a slight change in the features' importance ranking, again with "dependence count variance" as the radiomics feature that most affected the predictive performance. Strongly associated with LIME, SHAP summary plot results indicated, among others, "dependence count variance" from GLDM, "difference entropy" from GLCM, and first-order "variance" as the strongest indicators of IDH status mutation prediction. The

texture feature “dependence count variance” is sensitive to variability in voxel group sizes of similar intensity, implying increased complexity of the texture as its value increases, “difference entropy” is a texture disorder parameter that is analogous to the increase in the VOI’s signal intensity distribution differences, and “variance” reflects how gray levels are spread out about the mean. The first-order radiomic feature “robust mean absolute deviation” was also an important factor; it is associated with an increased contrast of high and low gray values between the 10th and 90th percentiles of the glioma histogram when the value increases. These characteristics of the dominant predictive radiomics can be explained using the more heterogeneous profile/nature of the IDH-wildtype glioma cases [22,34]. An analogous pattern was illustrated in the SHAP summary plot, where increased values of the top four features favored the IDH-wildtype mutation status (positive class) against the IDH-mutant status (negative class).

#### 4.4. Limitations and Future Work

Several limitations were identified in the study, including the lack of radiomics features robustness assessment using a test–retest study. This is an important part of the radiomics quality score (RQS) questionnaire; however, this was not feasible due to the dynamic nature of the acquired data and the retrospective collection from all centers. Another important issue is the limited number of time points that were used to select data from the perfusion curve. Extending our current study, future work will focus on investigating the added value that might arise when changes in radiomics features between specific time points (i.e., similar to delta-radiomics) will be included in the analysis pipeline. Additionally, the next step will be to involve both radiomics features extracted from raw DSC–MRI data and the derived parametric maps (e.g., cerebral blood flow). A comparative study between the proposed dynamic-based standardization technique and efficient feature-based methods (i.e., radiomics features harmonization using ComBat [37]) will be an important task in our future work. Despite the promising results, the high imbalance favoring the IDH status mutation in the examined cohorts, and hence a small number of IDH-wildtype patients, both for the developing and validating the proposed radiomics analysis pipeline, was a major limitation of this study, calling for extended studies with more participating centers.

## 5. Conclusions

In conclusion, the proposed study highlighted the impact of DSC–MRI radiogenomics in gliomas for IDH status subtyping. Increased accuracy was observed when dynamic-based standardization of the acquired imaging data was performed prior to radiomics and the generalizability of the proposed radiomics analysis was confirmed using two independent validation sets. Model explainability mechanisms discovered intuitive patterns between the IDH mutation status and radiomics features, which enhances the potential clinical value of the proposed DSC–MRI-based radiogenomics.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/cancers13163965/s1>, Table S1: List of the calculated features using Pyradiomics, Table S2: Methodological quality assessment of the proposed study using the radiomics quality score (RQS).

**Author Contributions:** G.C.M.: machine learning and radiomics analysis, writing—review and editing, and data visualization; G.S.I.: data normalization and data curation, writing original draft preparation, and data visualization; K.N.: data curation and writing—original draft preparation; L.S.: literature review and writing—original draft preparation, data curation, and tumor annotations; M.I., D.V., K.S.-P., and M.W.: data curation and writing—review and editing; S.B.: tumor annotations, supervision, and writing—review and editing; K.M.: supervision and writing—review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by: (a) the National Institute for Health Research to UCLH Biomedical Research Center (grant BRC399/NS/RB/101410) and (b) the Stavros Niarchos Foundation within the framework of the project ARCHERS (“Advancing Young Researchers’ Human Capital in Cutting Edge Technologies in the Preservation of Cultural Heritage and the Tackling of Societal Challenges”).

**Institutional Review Board Statement:** This multicenter study was based on retrospectively identified or prospectively acquired neuro-oncology MRI data collection in the participating centers, according to the local institutional review boards guidelines and the respective local ethical committees' approvals. Data sharing agreements between the collaborators and the central reading and post-processing site were also in place. Informed consent was deemed unnecessary due to the retrospective nature of the study and according to the national ethics regulations for medical research. The ethics approval protocol was submitted to the University College London/University College London Hospitals Joint Research Office (reference number: 213920) and the assigned North West–Liverpool Central Research Ethics Committee approved the study (reference number: 18/NW/0395). In addition, bilateral data transfer agreements were reviewed by the University College London Hospitals Joint Research Office and the counterpart research committees were put in place with each contributing institution. All individual data was anonymized and only the collated data on perfusion parameters was seen by statisticians.

**Informed Consent Statement:** Informed consent was deemed unnecessary due to the retrospective nature of the study and according to the national ethics regulations for medical research.

**Data Availability Statement:** The datasets used and analyzed during the current study, and the code for the machine learning and the radiomics analysis are available from the corresponding author on reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Ostrom, Q.T.; Cioffi, G.; Gittleman, H.; Patil, N.; Waite, K.; Kruchko, C.; Barnholtz-Sloan, J.S. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012–2016. *Neuro. Oncol.* **2019**, *21*, v1–v100. [\[CrossRef\]](#)
- The Cancer Genome Atlas Research Network. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *New Engl. J. Med.* **2015**, *372*, 2481–2498. [\[CrossRef\]](#) [\[PubMed\]](#)
- Louis, D.N.; Perry, A.; Reifenberger, G.; von Deimling, A.; Figarella-Branger, D.; Cavenee, W.K.; Ohgaki, H.; Wiestler, O.D.; Kleihues, P.; Ellison, D.W. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: A summary. *Acta Neuropathol.* **2016**, *131*, 803–820. [\[CrossRef\]](#)
- Stupp, R.; Mason, W.P.; van den Bent, M.J.; Weller, M.; Fisher, B.; Taphoorn, M.J.B.; Belanger, K.; Brandes, A.A.; Marosi, C.; Bogdahn, U.; et al. Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. *New Engl. J. Med.* **2005**, *352*, 987–996. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lu, V.M.; O'Connor, K.P.; Shah, A.H.; Eichberg, D.G.; Luther, E.M.; Komotar, R.J.; Ivan, M.E. The prognostic significance of CDKN2A homozygous deletion in IDH-mutant lower-grade glioma and glioblastoma: A systematic review of the contemporary literature. *J. Neurooncol.* **2020**, *148*, 221–229. [\[CrossRef\]](#) [\[PubMed\]](#)
- Leeper, H.E.; Caron, A.A.; Decker, P.A.; Jenkins, R.B.; Lachance, D.H.; Giannini, C. IDH mutation, 1p19q codeletion and ATRX loss in WHO grade II gliomas. *Oncotarget* **2015**, *6*, 30295–30305. [\[CrossRef\]](#)
- Yan, H.; Parsons, D.W.; Jin, G.; McLendon, R.; Rasheed, B.A.; Yuan, W.; Kos, I.; Batinic-Haberle, I.; Jones, S.; Riggins, G.J.; et al. IDH1 and IDH2 Mutations in Gliomas. *New Engl. J. Med.* **2009**, *360*, 765–773. [\[CrossRef\]](#) [\[PubMed\]](#)
- Weller, M.; van den Bent, M.; Preusser, M.; Le Rhun, E.; Tonn, J.C.; Minniti, G.; Bendszus, M.; Balana, C.; Chinot, O.; Dirven, L.; et al. EANO guidelines on the diagnosis and treatment of diffuse gliomas of adulthood. *Nat. Rev. Clin. Oncol.* **2021**, *18*, 170–186. [\[CrossRef\]](#)
- Taweksomboonyat, C.; Tunthanathip, T.; Sae-Heng, S.; Oearsakul, T. Diagnostic Yield and Complication of Frameless Stereotactic Brain Biopsy. *J. Neurosci. Rural Pract.* **2019**, *10*, 78–84. [\[CrossRef\]](#)
- Park, S.I.; Suh, C.H.; Guenette, J.P.; Huang, R.Y.; Kim, H.S. The T2-FLAIR mismatch sign as a predictor of IDH-mutant, 1p/19q-noncodeleted lower-grade gliomas: A systematic review and diagnostic meta-analysis. *Eur. Radiol.* **2021**, *31*, 5289–5299. [\[CrossRef\]](#)
- Patel, S.H.; Poisson, L.M.; Brat, D.J.; Zhou, Y.; Cooper, L.; Snuderl, M.; Thomas, C.; Franceschi, A.M.; Griffith, B.; Flanders, A.E.; et al. T2-FLAIR Mismatch, an Imaging Biomarker for IDH and 1p/19q Status in Lower-grade Gliomas: A TCGA/TCIA Project. *Clin. Cancer Res.* **2017**, *23*, 6078–6085. [\[CrossRef\]](#)
- Bhandari, A.P.; Liong, R.; Koppen, J.; Murthy, S.V.; Lasocki, A. Noninvasive determination of IDH and 1p19q status of lower-grade gliomas using MRI radiomics: A systematic review. *Am. J. Neuroradiol.* **2021**, *42*, 94–101. [\[CrossRef\]](#) [\[PubMed\]](#)
- Peng, H.; Huo, J.; Li, B.; Cui, Y.; Zhang, H.; Zhang, L.; Ma, L. Predicting Isocitrate Dehydrogenase (IDH) Mutation Status in Gliomas Using Multiparameter MRI Radiomics Features. *J. Magn. Reson. Imaging* **2021**, *53*, 1399–1407. [\[CrossRef\]](#) [\[PubMed\]](#)
- Habib, A.; Jovanovich, N.; Hoppe, M.; Ak, M.; Mamindla, P.; Colen, R.R.; Zinn, P.O. MRI-Based Radiomics and Radiogenomics in the Management of Low-Grade Gliomas: Evaluating the Evidence for a Paradigm Shift. *J. Clin. Med.* **2021**, *10*, 1411. [\[CrossRef\]](#) [\[PubMed\]](#)

15. Arita, H.; Kinoshita, M.; Kawaguchi, A.; Takahashi, M.; Narita, Y.; Terakawa, Y.; Tsuyuguchi, N.; Okita, Y.; Nonaka, M.; Moriuchi, S.; et al. Lesion location implemented magnetic resonance imaging radiomics for predicting IDH and TERT promoter mutations in grade II/III gliomas. *Sci. Rep.* **2018**, *8*, 11773. [[CrossRef](#)] [[PubMed](#)]
16. Liu, X.; Li, Y.; Li, S.; Fan, X.; Sun, Z.; Yang, Z.; Wang, K.; Zhang, Z.; Jiang, T.; Liu, Y.; et al. IDH mutation-specific radiomic signature in lower-grade gliomas. *Aging* **2019**, *11*, 673–696. [[CrossRef](#)] [[PubMed](#)]
17. Lu, C.-F.; Hsu, F.-T.; Hsieh, K.L.-C.; Kao, Y.-C.J.; Cheng, S.-J.; Hsu, J.B.-K.; Tsai, P.-H.; Chen, R.-J.; Huang, C.-C.; Yen, Y.; et al. Machine Learning–Based Radiomics for Molecular Subtyping of Gliomas. *Clin. Cancer Res.* **2018**, *24*, 4429–4436. [[CrossRef](#)]
18. Wu, S.; Meng, J.; Yu, Q.; Li, P.; Fu, S. Radiomics-based machine learning methods for isocitrate dehydrogenase genotype prediction of diffuse gliomas. *J. Cancer Res. Clin. Oncol.* **2019**, *145*, 543–550. [[CrossRef](#)]
19. Lambin, P.; Leijenaar, R.T.H.; Deist, T.M.; Peerlings, J.; De Jong, E.E.C.; Van Timmeren, J.; Sanduleanu, S.; Larue, R.T.H.M.; Even, A.J.G.; Jochems, A.; et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 749–762. [[CrossRef](#)]
20. Park, J.E.; Kim, H.S.; Kim, D.; Park, S.Y.; Kim, J.Y.; Cho, S.J.; Kim, J.H. A systematic review reporting quality of radiomics research in neuro-oncology: Toward clinical utility and quality improvement using high-dimensional imaging features. *BMC Cancer* **2020**, *20*, 1–11. [[CrossRef](#)] [[PubMed](#)]
21. Wesseling, P.; Ruiter, D.J.; Burger, P.C. Angiogenesis in brain tumors; pathobiological and clinical aspects. *J. Neurooncol.* **1997**, *32*, 253–265. [[CrossRef](#)] [[PubMed](#)]
22. Kickingereder, P.; Sahm, F.; Radbruch, A.; Wick, W.; Heiland, S.; Von Deimling, A.; Bendszus, M.; Wiestler, B. IDH mutation status is associated with a distinct hypoxia/angiogenesis transcriptome signature which is non-invasively predictable with rCBV imaging in human glioma. *Sci. Rep.* **2015**, *5*. [[CrossRef](#)]
23. Sudre, C.H.; Panovska-Griffiths, J.; Sanverdi, E.; Brandner, S.; Katsaros, V.K.; Stranjalis, G.; Pizzini, F.B.; Ghimenton, C.; Surlan-Popovic, K.; Avsenik, J.; et al. Machine learning assisted DSC-MRI radiomics as a tool for glioma classification by grade and mutation status. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 149. [[CrossRef](#)]
24. Conte, G.M.; Castellano, A.; Altabella, L.; Iadanza, A.; Cadioli, M.; Falini, A.; Anzalone, N. Reproducibility of dynamic contrast-enhanced MRI and dynamic susceptibility contrast MRI in the study of brain gliomas: A comparison of data obtained using different commercial software. *Radiol. Med.* **2017**, *122*, 294–302. [[CrossRef](#)]
25. Porz, N.; Bauer, S.; Pica, A.; Schucht, P.; Beck, J.; Verma, R.K.; Slotboom, J.; Reyes, M.; Wiest, R. Multi-Modal Glioblastoma Segmentation: Man versus Machine. *PLoS ONE* **2014**, *9*, e96873. [[CrossRef](#)] [[PubMed](#)]
26. Barker, P.B.; Golay, X.; Zaharchuk, G. (Eds.) *Clinical Perfusion MRI*; Cambridge University Press: Cambridge, UK, 2013; ISBN 9781139004053.
27. Carré, A.; Klausner, G.; Edjlali, M.; Lerousseau, M.; Briand-Diop, J.; Sun, R.; Ammari, S.; Reuzé, S.; Alvarez Andres, E.; Estienne, T.; et al. Standardization of brain MR images across machines and protocols: Bridging the gap for MRI-based radiomics. *Sci. Rep.* **2020**, *10*, 12340. [[CrossRef](#)] [[PubMed](#)]
28. Van Griethuysen, J.J.M.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.H.; Fillion-Robin, J.C.; Pieper, S.; Aerts, H.J.W.L. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* **2017**, *77*, e104–e107. [[CrossRef](#)] [[PubMed](#)]
29. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Müller, A.; Nothman, J.; Louppe, G.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2012**, *12*, 2825–2830.
30. Lemaître, G.; Nogueira, F.; Aridas, C.K. Imbalanced-Learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning. *J. Mach. Learn. Res.* **2017**, *18*, 559–563.
31. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In *Proceedings of the Advances in Neural Information Processing Systems*; Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc.: Long Beach, CA, USA, 2017; Volume 30.
32. Ribeiro, M.T.; Singh, S.; Guestrin, C. “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. In *Proceedings of the 22nd {ACM} {SIGKDD} International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 13–17 August 2016; pp. 1135–1144.
33. Keenan, K.E.; Ainslie, M.; Barker, A.J.; Boss, M.A.; Cecil, K.M.; Charles, C.; Chenevert, T.L.; Clarke, L.; Evelhoch, J.L.; Finn, P.; et al. Quantitative magnetic resonance imaging phantoms: A review and the need for a system phantom. *Magn. Reson. Med.* **2018**, *79*, 48–61. [[CrossRef](#)]
34. Choi, K.S.; Choi, S.H.; Jeong, B. Prediction of IDH genotype in gliomas with dynamic susceptibility contrast perfusion MR imaging using an explainable recurrent neural network. *Neuro. Oncol.* **2019**, *21*, 1197–1209. [[CrossRef](#)] [[PubMed](#)]
35. Li, Z.; Zhao, W.; He, B.; Koh, T.S.; Li, Y.; Zeng, Y.; Zhang, Z.; Zhang, J.; Hou, Z. Application of distributed parameter model to assessment of glioma IDH mutation status by dynamic contrast-enhanced magnetic resonance imaging. *Contrast Media Mol. Imaging* **2020**, *2020*, 8843084. [[CrossRef](#)] [[PubMed](#)]
36. Zhang, L.; He, L.; Lugano, R.; Roodakker, K.; Bergqvist, M.; Smits, A.; Dimberg, A. IDH mutation status is associated with distinct vascular gene expression signatures in lower-grade gliomas. *Neuro. Oncol.* **2018**, *20*, 1505–1516. [[CrossRef](#)]
37. Da-ano, R.; Masson, I.; Lucia, F.; Doré, M.; Robin, P.; Alfieri, J.; Rousseau, C.; Mervoyer, A.; Reinhold, C.; Castelli, J.; et al. Performance comparison of modified ComBat for harmonization of radiomic features for multicenter studies. *Sci. Rep.* **2020**, *10*, 10248. [[CrossRef](#)] [[PubMed](#)]