HHS Public Access

Author manuscript

Semin Ultrasound CT MR. Author manuscript; available in PMC 2023 April 01.

Published in final edited form as:

Semin Ultrasound CT MR. 2022 April ; 43(2): 153–169. doi:10.1053/j.sult.2022.02.005.

Brain Tumor Imaging: Applications of Artificial Intelligence

Muhammad Afridi, BS[*], **Abhi Jain, DO**[†], **Mariam Aboian, MD PhD**[†], **Seyedmehdi Payabvash, MD**[†]

[*]Rowan University, School of Osteopathic Medicine, Stratford, NJ, United States

[†]Department of Radiology and Biomedical Imaging, Yale School of Medicine, New Haven, CT, United States

Abstract

Artificial intelligence has become a popular field of research with goals of integrating it into the clinical decision-making process. A growing number of predictive models are being employed utilizing machine learning that includes quantitative, computer-extracted imaging features known as radiomic features, and deep learning systems. This is especially true in brain-tumor imaging where artificial intelligence has been proposed to characterize, differentiate, and prognostication. We reviewed current literature regarding the potential uses of machine learning-based, and deep learning-based artificial intelligence in neuro-oncology as it pertains to brain tumor molecular classification, differentiation, and treatment response. While there is promising evidence supporting the use of artificial intelligence in neuro-oncology, there are still more investigations needed on a larger, multi-center scale along with a streamlined and standardized image processing workflow prior to its introduction in routine clinical decision-making protocol.

Introduction

Artificial intelligence (AI)-based analysis of imaging data has revolutionized the field of noninvasive biomarker discovery. It relies on using radiologic images as mineable databases with quantitative radiomic or texture features that can be learned and/or predict clinically significant output[1]. Machine learning (ML) and deep learning are subsets of AI, each with unique qualities that allow for computerized image analysis.

Radiomics

Radiomics is most currently described as the “high-throughput extraction of quantitative features that result in the conversion of images into mineable data and the subsequent analysis of these data for decision support”[2]. While the concept of data mining is not novel, and nor is it based in AI, the recent advances in ML has made possible radiomic feature extraction with subsequent image analysis. More specifically, ML can extrapolate the mined

Seyedmehdi Payabavsh, MD, Department of Radiology and Biomedical Imaging, Yale University School of Medicine, 789 Howard Ave, PO Box 208042, New Haven, CT 06519, Phone: +1 (203) 214 4650, Fax: +1 (203) 737 1241, sam.payabvash@yale.edu.

data to produce clinically significant prediction models and classifiers through computer algorithms[1]. While the scope of this article centers around the use of AI in neuro-oncology imaging, the combination of radiomics and AI is applicable to a wider range of systems and pathology.

Radiomics can be further subdivided into feature-based or deep learning-based radiomics, based on the method of radiomic feature acquisition. In feature-based radiomics, predetermined features are mathematically extracted from specific region-of-interest (ROI) and are commonly referred to as “handcrafted” or “hand-engineered” features[1]. These radiomic features are then selected based on feature selection algorithms. In contrast, deep learning-based radiomics involves training computer models from the generated data, through learning algorithms and advanced statistics, to extract pertinent radiomic features[3]. It stands to reason that feature-based radiomics is limited by finite mathematics-based relations when compared to deep learning-based radiomics, which is continuously refined with each data entry. Handcrafted features also require standardization of technique, image preprocessing and ROI selection, leaving it exposed to variations in image acquisition, data analysis and generalizability. Due to the predetermined nature of handcrafted features, they are better suited for smaller data sets, which could explain their prevalence in literature.

Deep learning-based radiomics seeks to imitate the function of the human brain by using artificially constructed neural networks. These neural architectures, such as convolutional neural networks (CNNs) find the most relevant features from the input data, which are used for pattern recognition or the classification of non-linear data. Individual neural layers with linear/nonlinear activation functions learn the representation of imaging data with various levels of abstraction, after which the layers are then stacked and connected for classification and output[4]. Each hidden layer within the network is responsible for data from one level - for example, the first level may represent edges in an image oriented in a specific direction, while the second layer could be responsible for motif detection in the observed edges, and the third could recognize objects from the ensembles of motifs[5]. The extracted features can be processed by the network itself for analysis of performance and classification, or they can undergo model generation through a similar process as feature-based radiomics by using different classifiers such as support vector machines (SVM), regression models or decision trees[3]. While feature-based radiomics requires image preprocessing, the opposite might be true for deep learning as standardization may have a negative impact by removing information. Due to the self-learning aspect of deep neural networks, it is more likely to have poor performance on smaller datasets, which is one of the reasons that most studies utilize feature-based radiomics to test their hypothesis[1].

Characterization of brain tumors

Over 150 different brain tumors have been described based on histopathological characteristics. The gold-standard for their characterization requires histopathological analysis by retrieving tumor samples from biopsy[6]. However, due to the heterogeneity of some tumors, their inaccessible location, or the patient’s clinical status, noninvasive radiological characterization of the brain tumors will be ideal. AI is a promising tool that

serves to compliment, and possibly replace, the need for invasive biopsies by combining radiomic and non-radiomic features to characterize brain tumors[1].

Brain tumor classification

Gliomas are the most common brain tumor and can be divided into grades based on recently modified WHO criteria. In 2016, the WHO introduced molecular markers in conjunction with histopathology to characterize gliomas based on potential malignancy, where a designation of grade II confers lowest risk of malignancy and grade IV confers the highest[7]. Grade II and III gliomas can be characterized as low-grade gliomas (LGG) with the most favorable outcomes, whereas grade IV gliomas are considered high-grade gliomas (HGG) and are associated with poor outcomes. The accurate and efficient classification of gliomas is paramount in planning appropriate treatment and follow up, and the introduction of molecular and genomic markers in their classification has introduced novel applications for ML.

MRI is the mainstay of brain tumor imaging. In a study by Cho and colleagues[8], the investigators sought to use handcrafted feature-based radiomics to classify glioma grades. They utilized cases from the Brain Tumor Segmentation 2017 Challenge (BraTS 2017), and analyzed each case with multi-modal MRI including T1-weighted, T1-contrast enhanced, T2-weighted and FLAIR images (Table 1). They identified a total of 468 radiomic features from three different ROIs, from which they isolated five relevant features using the minimum redundancy maximum relevance algorithm. The five narrowed features then served to build three classifier models including logistics, SVM, and random forest classifiers. The results suggested that tumor morphological property features, including spherical disproportion and compactness, along with grey level co-occurrence matrix (GLCM) features, which represent texture, were most effective at discriminating LGG from HGG. On average, the classifiers graded gliomas with an accuracy of 93%, sensitivity of 98%, specificity of 79% and receiver operating characteristics (ROC) area under the curve (AUC) of 94%. Sun and colleagues[9] demonstrated the use of least absolute shrinkage and selection operator (LASSO) to select the most predictive radiomics features for glioma grading. They calculated a radiomics score (Rad-score) and built a logistic regression model to investigate correlation between glioma grade and Rad-score. They performed retrospective analysis on 146 glioma patients using 5 radiomic features selected by LASSO, with AUC for glioma grading to be 0.919 (Table 1). These studies demonstrate the use of multiple classifiers in accurately characterizing glioma grades.

Sudre and colleagues[10] evaluated the role of dynamic susceptibility contrast (DSC)-MRI-based radiomics in classifying gliomas across their WHO grades II-IV and their isocitrate dehydrogenase (IDH) mutation status. DSC-MRI data from 333 patients from 6 different tertiary centers was processed for normalized leakage-corrected relative cerebral blood volume (rCBV) maps. A random forest algorithm was used to predict glioma grades and mutation status using extracted and selected features. Their results showed that shape, distribution, and texture features were significantly different across mutation status. WHO grade II vs. III differentiation was driven primarily by shape features whereas grade III vs. IV was mainly differentiated with texture and intensity features. In their study, 71% of

the cases were correctly predicted based on mutation status, and 53% of the cases were correctly stratified based on WHO grades (Table 1). Tian and colleagues[11] compared the utility of single sequence MRI to multiparametric MRI, as well as comparing the efficacy of histogram parameters and texture features in glioma grading.

MRI sequences used included pre and post contrast T1-weighted, T2-weighted, multi-b-value diffusion-weighted, and 3D arterial spin labeling sequences. SVM-based recursive feature was used to isolate optimal features, which was then used to establish classifiers. They were able to differentiate LGG from HGG with 96.8% accuracy, and grade III from grade IV glioma with 98.1% accuracy. Moreover, their results suggested that texture features were more effective at grading gliomas than histogram parameters in terms of accuracy, sensitivity, specificity and AUC. The results also suggested that multiparametric MRI was superior to single sequence MRI, with T1-weighted contrast enhanced (89.2% accurate), DSC (86.9% accurate) and T2-weighted (86.5% accurate) being the most accurate sequences (Table 1). Huang and colleagues[12] supported these results as they investigated the role of different MRI sequences in grading gliomas using radiomics. Their results demonstrated that radiomics analysis based on multiparametric MRI can accurately grade gliomas, with T1-weighted contrast enhanced images being the most effective in isolation but improved when combined with clinical features. These studies did not utilize separate validation cohorts, however, and due to the imbalance and limited sample size, significant variance in the models' performance in a separate validation cohort cannot be excluded. Hsieh and colleagues[13] sought to mitigate variations in scanning and image acquisition by converting texture features in MR imaging to intensity-invariant ones. They created a computer-aided detection (CAD) model using intensity-invariant MR images to differentiate between LGG and GBM. They collected MRI datasets from the cancer genome atlas (TCGA) and the cancer imaging archive (TCIA)[14], and transformed ROI texture features into a local binary pattern (LBP), which transformed local textures in MR imaging to intensity-invariant ones. From LBP, they could extract and combine histogram moments and texture into a logistic regression model classifier used for predicting glioma grade. The CAD performance showed an accuracy of 93%, sensitivity of 97%, and a NPV of 99%, compared to conventional texture features, which showed an accuracy of 84%, sensitivity of 76% and a NPV of 89%.

While these studies demonstrate the utility of standard, handcrafted radiomic features for glioma grade prediction, Gutta and colleagues[15] investigated whether the use of deep convolutional neural networks (CNN) would significantly improve glioma classification. They retrospectively analyzed 237 patients with gliomas using multiparametric MRI, after which the images were resampled, registered, skull-stripped, and segmented to extract the tumors using automatic segmentation via a cascade of CNNs proposed by Wang et al[16]. The learned features from the trained CNN were then used to predict glioma grade, and its performance was compared with standard ML approaches including SVM, random forests and gradient boosting trained with radiomic features. Their results demonstrated an average accuracy of 87% in predicting glioma grade when utilizing learned features extracted from CNN, compared to an accuracy of 64% when using the top-performing ML model. These findings are in accordance with previous studies that used CNNs to classify glioma grades with accuracies ranging from 71% to 96%[17–19].

Novel methods to distinguish between LGG and HGG based on conventional MR images using CNNs have been proposed. Zhuge and colleagues[20] proposed two such methods for glioma grading. Both methods rely on 3D tumor segmentation using a modification of the U-Net model and tumor classification based on segmented brain tumor, however the first method uses the mask R-CNN[21] model for tumor grading while the second method uses a 3D volumetric CNN (called 3DConvNet) on ROIs for segmented tumor grading. These methods were subsequently tested on the TCIA and BraTS datasets. The R-CNN resulted in a sensitivity of 93.5%, specificity of 97.2% and accuracy of 96.3% while the 3DConvNet showed a sensitivity, specificity, and accuracy of 94.7%, 96.8%, and 97.1%, respectively. Ozcan and colleagues[22] also trained a fully automatic custom CNN from scratch and compared its performance in glioma grade prediction with pretrained models including AlexNet, GoogLeNet, and SqueezeNet. Their results suggest a comparable or even enhanced performance compared to pretrained models based on five-fold cross-validation of 104 pathology-proven cases. These studies advocate for the use of CNNs for glioma grading in conjunction with, and in certain circumstances, instead of, surgical biopsies.

Zhang and colleagues[23] demonstrated the utility of Diffusion Tensor Imaging (DTI) in extracting radiomic features pertinent to glioma grading. This retrospective study utilized pre-trained CNNs as well as traditional radiomic features to extract features from manually selected tumor regions in DTI images. When differentiating LGG vs. HGG using a combination of FA and MD, they found accuracy, sensitivity, and specificity of 94%, 98% and 86%, respectively. When differentiating grade III from IV using the same combination, they achieved an accuracy, sensitivity, and specificity of 98%, 98% and 100%, respectively. They also suggested that deep radiomic features derived from CNN exhibited superior prediction of glioma grade than handcrafted features. Takahashi and colleagues[24] also support the use of DTI in glioma grading as they created an accurate ML model using 6 features extracted from ADC and mean kurtosis (MK) images using SVM that had accuracies of 91% and 93% respectively.

Pyka and colleagues[25] evaluated the utility of amino acid positron emission tomography (PET) with [18F]-fluoroethyl-L-tyrosine (FET) tracer in differentiating between WHO grade III and IV gliomas. The FET PET-based metabolic tumor volume combined with textural features derived from GLCM were used to result in a diagnostic accuracy of 85%. Other groups investigating the role of nuclear medicine in grading gliomas had similar results[26]. Studies investigating the role of AI in predicting glioma grade are summarized in Table 1.

Predicting 1p/19q co-deletion status and IDH genotype in gliomas

The introduction of molecular biomarkers and genotypic parameters in the grading of gliomas has added a layer of objectivity to diagnosis in hopes of increased homogeneity and narrower definitions of glioma classification. The two molecular genetic features to highlight in the classification of gliomas are the IDH genotype, and loss of heterozygosity of the 1p/19q chromosome arms[7]. Specifically, IDH mutant gliomas, usually astrocytomas without 1p/19q co-deletions or oligodendrogliomas harboring 1p/19q co-deletions, have a significantly better prognosis in comparison with IDH wildtype gliomas, or GBM[3]. To minimize invasive procedures in gathering tissue samples for histological evaluation, the

role of radiomics has been evaluated to predict these molecular biomarkers in patients with gliomas.

Jian and colleagues[27] conducted a systematic review and meta-analysis to investigate the use of ML in predicting molecular markers for glioma grading. They identified 512 studies until April 2020, of which 44 met inclusion criteria. Of the 44 studies, 32 studies extracted radiomics features such as texture, intensity, and tumor shape, seven studies utilized deep learning, and 5 studies exclusively used quantitative parameters such as MR spectroscopy or Visually Accessible Rembrandt Imaging (VASARI) features. Random forest and SVM were the most common classifiers utilized. They applied 18 studies on training datasets and found that the pooled sensitivity and specificity of predicting IDH mutation was 88% and 86%, respectively with an AUC of 0.92. The pooled sensitivity and specificity of the 12 studies applied on the validation sets were 85% and 83%, respectively with an AUC of 0.90. Six studies investigating 1p/19q codeletion reported training results with a pooled sensitivity and specificity of 83% and 76%, respectively with AUC of 0.83. Validation performance across five studies yielded a pooled sensitivity of 70%, specificity of 72% and AUC of 0.75. Bhandari and colleagues[28] conducted a similar systematic review where they investigated the use of MRI radiomics in predicting IDH and 1p/19q status of LGG. They selected 14 journal articles out of 532 based on inclusion criteria. Their results suggested that optimal classification of 1p/19q status occurred with texture-based radiomics and had a 90% sensitivity and 96% specificity. The most accurate classifier for predicting IDH status used conventional radiomics in combination with CNN derived features as this exhibited a 94.4% sensitivity and 86.7% specificity. However, examining deep features exclusively was found to be superior in predicting other genotypic mutations[29]. The stark limitation in both these systematic review meta-analyses is the relatively high heterogeneity in both studies with Bhandari and colleagues noting Higgins I2 heterogeneity of 88.55% and 86.19% in predicting IDH and 1p/19q status, respectively. This can be explained by the variation in radiomic pipelines, and manual segmentation.

Other groups have also investigated the role of radiomics in predicting genotypes of gliomas. Shofty and colleagues[30] tested the utility of different classifiers in predicting 1p/19q codeletion status in LGG. Their results suggested that the Ensemble Bagged Trees classifier has the most accurate prediction with sensitivity, specificity, and accuracy of 92%, 83% and 87% respectively. Lu and colleagues[31] proposed a three-level ML model based on multimodal MR radiomics to classify IDH mutations and 1p/19q codeletions into 5 subtypes: LGG with IDH mutation and 1p/19q codeletion; LGG with IDH mutation and 1p/19q non-codeletion; LGG with wild-type IDH; GBM with IDH mutation; and GBM with wild-type IDH. Using 4 binary classifiers, their results ranged in accuracy between 87% and 96% on the training cohort, and 80% to 92% on the validation cohort (Table 2). Han and colleagues[32] investigated the utility of combining pertinent clinical factors with the radiomics signature via logistic regression algorithm in differentiating 1p/19q codeletion genotypes. The random forest classifier was used and the results showed an AUC of 0.887 and 0.760 on training and validation cohorts using only the radiomic signature, respectively. The combination of clinical features to radiomic signature did not significantly improve performance and yielded an AUC of 0.885 and 0.753, respectively. Zhou and colleagues[33] investigated a similar concept where they extracted histogram, shape and texture features

from multimodal MRIs and combined it with patient age using a random forest algorithm to generate a model predictive of IDH mutation status and 1p19q codeletion in LGG and HGG. They suggest that age offered the highest predictive value, followed by shape features. The overall accuracy for prediction of IDH-wild type, IDH-mutant and 1p19q codeletion, and IDH-mutant and no 1p19q codeletion was 78.2% (Table 2).

Groups have also investigated the use of deep learning in predicting molecular markers in gliomas. Chang and colleagues[34] sought to train a CNN to predict underlying molecular genetic mutation status in gliomas and identify the most predictive imaging features in each mutation. The CNN algorithm was then used on 259 patients from The Cancer Institute Archive[14] with LGG and HGG. It predicted IDH mutation with an accuracy of 94% and AUC of 0.91 and predicted 1p/19q codeletions with an accuracy of 92% and AUC of 0.88. The principal component analysis of the final CNN revealed that for IDH mutations, the most predictive features were absent or minimal areas of enhancement, central areas of cysts with low T1 and FLAIR suppression, and well-defined tumor margins. The same analysis revealed that for 1p/19q codeletion, the most predictive features were left frontal lobe location, ill-defined tumor margins and larger portion of enhancement. Li and colleagues[35] directly compared the accuracy of deep learning CNNs to standard radiomics in predicting IDH mutations in LGG. They used a modified CNN structure with 6 convolutional layers and obtained image features by normalizing the information of the last convolutional layers of the CNN. Using the same dataset in the prediction of IDH mutations, the normal radiomics method had an AUC of 0.86 whereas the deep learning-based radiomics had an AUC of 92%, which was further improved to 0.95 when based on multimodal MR images. Yan and colleagues[36] used Bayesian-regularization neural networks to predict IDH mutation and compare performance of different MR parameters. They found that an image fusion model incorporating radiomic signatures based on contrast-enhanced T1-weighted imaging and apparent diffusion coefficient, had the most accurate prediction of IDH mutations with an AUC of 0.884. Whereas the contrast-enhanced T1-weighted images had the most favorable performance in predicting 1p/19q codeletion status with an AUC of 0.815. Eichinger and colleagues[37] evaluated the utility of DTI features to predict IDH genotype in LGG. They used a single hidden layer neural network trained on texture features generated from preoperative B0 and fractional anisotropy (FA) to predict IDH status. Their results showed prediction accuracy of 92% in training data and 95% in the validation cohort. The ten most important features for prediction comprised tumor size and both B0 and FA texture information.

Literature also advocates for the use of nuclear medicine in predicting molecular genotype. Lohmann and colleagues[38] investigated the potential of O-(2-[18F]fluoroethyl)-L-tyrosine (FET) PET radiomics based on textural features in conjunction with static and dynamic parameters of FET uptake for prediction of the IDH genotype. A total of 84 patients were scanned using either a standard scanner or high-resolution hybrid PET/MR scanner. Independent of scanner type, their results suggested significantly improved diagnostic accuracy in predicting IDH genotype when combining PET parameters with textural features, compared to textural features alone, with the highest diagnostic accuracy being 93% while using the hybrid PET/MR scanner. Additionally, Zaragori and colleagues[39] investigated the utility of 18F-FDOPA PET imaging in conjunction with MRI to predict

IDH mutation and 1p/19q status. They extracted a set of 114 features, which included conventional static features, dynamic features and other radiomic features, from ML models used to predict IDH and 1p/19q codeletion status. The most accurate models were able to predict IDH mutation and 1p/19q codeletion status with an AUC of 0.83 and 0.72, respectively. Feature importance, assessed using SHapley Additive exPlanation (SHAP) values, suggested that dynamic features were the most important features in the model to predict IDH mutations while other radiomics features were the most important in predicting 1p/19q codeletion status. Table 2 summarizes the studies that investigated the utility of AI in prediction IDH and 1p19q codeletions status.

Predicting MGMT promoter methylation status in GBM

The epigenetic silencing of the O6-methylguanine-DNA methyltransferase (MGMT) DNA-repair gene via promoter methylation decreases DNA repair. Multiple studies have shown this silencing to be associated with significantly longer survival in patients with GBM who are being treated with alkylating agents[40,41]. The following section reviews literature pertinent to noninvasively predicting MGMT promoter methylation status in patients with GBM.

Huang and colleagues[42] aimed to build a radiological model based on standard MR sequences to detect MGMT methylation status in gliomas using texture analysis. They generated a combined model using the top five most effective texture features (selected from a total of 396 features) in each MR sequence to predict MGMT methylation status in a GBM dataset and an overall glioma dataset. Their model predicted MGMT methylation status with a 90.5% sensitivity and a 72.7% sensitivity (AUC=0.818) in the GBM dataset, and a 70.2% sensitivity and a 90.6% specificity (AUC=0.833) for the glioma dataset. Li and colleagues[43] sought to build a reliable radiomics model from conventional MRI for the prediction of MGMT promoter methylation sequence in GBM patients. They retrospectively extracted 1,705 multiregional radiomics features, and isolated six features using ML-based algorithm, Boruta, to build a random forest classification model that predicted MGMT status, which they tested on a primary cohort of 133 patients, and a validation cohort of 60 patients. Their model predicted MGMT promoter methylation status with an accuracy of 80% (AUC=0.88). Combining clinical features with radiomics features did not improve prediction performance. Xi and colleagues[44] investigated a similar hypothesis but utilized LASSO to isolate 36 radiomics features that were based on conventional MRI. Twenty GBM patients were in the validation cohort, and their results suggest that the best classification system for predicting MGMT promoter methylation status combined T1, T2, and contrast-enhanced T1 weighted imaging features, which had an accuracy of 86.6% in the validation cohort, and 80% in the test dataset. Vils and colleagues[45] utilized data from the DIRECTOR trial[46] to investigate the role of radiomics in predicting MGMT status for patients with recurrent GBM. Contrast-enhanced T1-weighted images were used to extract 180 features, after which principal component analysis was used to perform radiomic feature selection. 69 patients enrolled into the DIRECTOR trial served as the training cohort and 49 independent patients served as the external validation cohort. Their model predicted MGMT status with an AUC of 0.67 on the training dataset, and an AUC of 0.673 for the validation cohort. Recently, Le and colleagues[47] hoped to improve accuracy of radiomics-based models in

predicting MGMT status by investigating a radiomics-based eXtreme Gradient Boosting (XGBoost) model in IDH1 wildtype mutation GBM patients. XGBoost is widely used in competitions due to its potential of controlling overfitting. They extracted radiomic features from multimodality MRI and tested with F-score analysis to identify important features to improve the model. They tested MGMT status prediction of their model on 53 patients, and the results identified nine radiomic features with and AUC of 0.896. Crisi and colleagues[48] evaluated whether radiomic features from DSC-MRI would have sufficient strength to predict MGMT methylation status in GBM patients. Their results found 14 radiomic quantitative imaging features that helped differentiate between non-methylated and methylated MGMT sequences, which they used to build a perceptron deep learning model to classify MGMT status into 3 groups: unmethylated MGMT promoter sequence (< 10% methylated), intermediate-methylated sequence (between 10% and 30% methylated), and methylated MGMT promoter sequence (>29% methylated). Their model classified MGMT status into these groups with an AUC, sensitivity, and specificity of 0.84, 75% and 85%, respectively.

The use of deep learning to predict MGMT promoter methylation status has also been evaluated. Chang and colleagues[34] sought to train a CNN that could independently predict MGMT promoter methylation status in gliomas. They retrospectively obtained MRI data from The Cancer Imaging Archive[14] for 259 patients with LGG and HGG. Their feature analysis found that for MGMT status, the most predictive features were a heterogenous, nodular enhancement; the presence of an eccentric cyst; mass-like edema with cortical involvement and slight frontal and superficial temporal predominance. Their CNN model predicted MGMT status with an accuracy of 83%.

Korfiatis and colleagues[49] compared three different residual deep neural network (ResNet) architectures in their ability to predict MGMT status in GBM patients without the need for a distinct tumor segmentation step, eliminating extensive image preprocessing. The three ResNet architectures consisted of 18 layers (ResNet18), 34 layers (ResNet30), and 50 layers (ResNet50). Accuracy was based on the model's ability to classify each slice as no tumor, methylated MGMT, non-methylated. Their results showed that ResNet50 was the most predictive of MGMT status with an accuracy of 95% during the validation phase, and an accuracy of 97% during the test phase. Lu and colleagues[50] found the optimal cutoff of MGMT promoter methylation status to be 12.75%, based on prediction of overall survival. They used top radiomic features based on MRI, Visually Accessible Rembrandt Images (VASARI) features and clinical features to build multiple ML models that predict MGMT status. Their models had accuracies ranging from 45% to 67%.

Current literature also evaluates the use of radiomics based on nuclear medicine images in predicting MGMT methylation status. Qian and colleagues[51] investigated the use of radiomic features derived from 18F-DOPA PET imaging in predicting MGMT promoter methylation status. Using features extracted from HGG contour based on a tumor-to-normal hemispheric ratio >2.0 with a random forest model, they achieved an accuracy of 80% for predicting MGMT status. Kong and colleagues[52] evaluated the use of radiomic features extracted based on 18F-fluorodeoxyglucose (FDG) PET images in predicting MGMT promoter methylation status. They used a 3D ROI and extracted 1561 radiomics features, of

which five features were selected for the radiomics signature. The radiomics signature was evaluated independently, and in combination with clinical features referred to as a fusion signature. Their results show that the radiomics signature alone produced the most accurate prediction of MGMT promoter methylation status with the AUC reaching 0.94 and 0.86 in the primary and validation cohorts, respectively. Table 3 summarizes the studies that investigated the role of AI in predicting MGMT methylation status in patients with glioma.

Differentiation of different tumor types

The high soft-tissue contrast seen in MRI allows it to be used as the primary imaging modality in differentiating brain tumors. However, multiple tumor types have similar appearance on MRI. GBM and metastases are the two most common brain tumors and are treated differently with maximal tumor resection followed by radiotherapy and temozolamide, and stereotactic radiosurgery, respectively. Unfortunately, both brain lesions present similarly on conventional brain MRI making clinical differentiation difficult. Furthermore, advanced MRI features have shown utility in differentiating GBM and metastases, no individual finding has enough evidence to drive clinical decision-making[53]. Multiple studies have demonstrated the use of machine-based learning to isolate pertinent radiomic features and classifiers, and evaluate brain lesions to differentiate between GBM and metastases[1,3,54]. The same approach can be taken to further differentiate the subtypes of metastatic brain lesions[55]. Some studies also compared practicing neuroradiologists to the best-performing ML classifiers in characterizing tumor type, and the results showed significantly better performance by the ML classifiers[51,53]. Table 4 summarizes the utility of MRI in ML models to differentiate between various brain tumors[53–62].

The role of AI in Digital Pathology Images

While much of this article has focused on the noninvasive applications of AI in neuro-oncologic imaging, it is important to note the utility of deep learning-based radiomics for the digital analysis of histopathology slides. Pei and colleagues[63] used a deep neural network-based classification method that fuses molecular and cellular features to grade gliomas in 549 patients from TCGA dataset. Their model had an accuracy of 93.8% in differentiating between HGG and LGG, and an accuracy of 74% in differentiating grade II vs. grade III gliomas (the 74% accuracy outperforms current state-of-the-art methods in classifying grade II vs. grade III gliomas). While Pei and colleagues segmented slides to analyze specific ROIs, Im and colleagues[64] used deep-learning to analyze whole-slide images and classify glioma grades and subtypes. Their model had an accuracy of 87.3% of diffuse glioma subtype classification. These studies highlight the utility of deep learning in analyzing histopathology. Further work needs to be done investigating the use of AI in digital pathology images in conjunction with noninvasive techniques discussed previously to maximize accuracy in glioma grading.

Prognostication

We have discussed the utility of AI in predicting molecular biomarkers such as IDH mutation, 1p/19q codeletion, and MGMT promoter methylation status, and their effects on patient prognosis. In this section, we highlight studies conducted by groups that evaluate the

use of other prognostic markers in ML to determine prognosis of patients with brain tumors, particularly GBM, which is the most common and most aggressive primary brain tumor with a median survival between 12 and 15 months[65].

Prasanna and colleagues[66] investigated the utility of radiomic features extracted from preoperative conventional MR images of the peritumoral brain zone in predicting long-term (>18 months) versus short-term (<7 months) survival in GBM patients. They obtained contrast-enhanced T1-weighted, T2 weighted, and FLAIR sequences on 65 patients from The Cancer Imaging Archive[14] and an expert reader segmented each study as enhancing, peritumoral brain zone and tumor necrosis. A minimum redundancy maximum relevance (mRMR) feature selection scheme was employed to extract 402 radiomic features, after which a random forest classifier was employed to isolate the most predictive features. From this, they addressed two questions - what is the relative role of each region within and around the tumor is predicting long-term vs. short-term GBM survival; and how does the addition of clinical features to the radiomics model affect prediction of overall survival. The results showed that peritumoral radiomic features were predictive across T2 weighted, with a concordance index (CI) of 0.637, and FLAIR sequences (CI=0.694), and radiomic features from the tumor necrosis segment were the most predictive of long-term vs. short-term survival for contrast-enhanced T1 weighted images (CI=0.69). Peritumoral radiomic features when combined across multi-parametric sequences were the best at predicting long-term vs. short-term survival for GBM patients (CI=0.70). When clinical features were combined with the peritumoral radiomic features across multi-parametric sequences, the model yielded highest predictive accuracy of GBM survival (CI=0.735). Kickingereder and colleagues[67] conducted a similar investigation on 119 GBM patients by using multi-parametric MRI based radiomic features from multiregional tumor volumes. Analysis based on 11 features allowed stratification into either high-risk or low-risk groups for progression free survival with a hazard ratio of 2.28 in the validation group and predicted overall survival with a hazard ratio of 3.45 in the validation cohort. In alignment with the previous study, they also found that prediction of patient prognosis improved when radiomic features were combined with clinical data. Park and colleagues[68] aimed to include diffusion- and perfusion-weighted MRI with conventional MRI to develop and validate a radiomics model for prognostication of patients with GBM. Radiomic features were extracted from a total of 216 patients, and feature selection via LASSO regression followed by calculation of radiomic score. A prognostic model was then developed using the radiomic score combined with clinical predictors. The radiomics model with clinical data performed best with a C-index of 0.74. External validation also showed good discrimination with a C-index of 0.70.

Tumor hypoxia is known to decrease survival in GBM patients,[69] thus, Beig and colleagues[70] investigated the use of radiomic features extracted from multi-parametric MRI to detect hypoxic changes that could stratify GBM patients into short-term (STS), mid-term (MTS) and long-term (LTS) survivors. A total of 115 different multi-parametric MR studies were segmented by 3 neuroradiologists and top 8 radiomic features were extracted to generate a hypoxia enrichment score (HES) based on 21 genes implicated in the hypoxia pathway for GBM[71], and predict patient survival. Their results on the validation set showed that there was a statistically significant separation between the Kaplan-Meier survival curves of STS vs. LTS (p=0.0032).

Another method by which prognosis can be stratified is through measuring the proliferative index of a tumor. Ki-67 is the most reliable marker of cell proliferation[72] and its expression levels have shown to confer a worse prognosis[73]. Li and colleagues[74] evaluated a radiomics-based approach in predicting expression levels of Ki-67 by extracting 431 radiomic features from 117 patients with LGG. A group of 9 radiological features were used for the final model, which predicted Ki-67 expression with an accuracy of 83.3% and 88.6% in the training and validation sets, respectively. Of the extracted features, only spherical disproportion of the tumor was found to be predictive of prognosis.

True Progression (TP) vs. Pseudoprogression (PsP)

Pseudoprogression (PsP) refers to treatment-related changes that mimic the true progression (TP) of post-treatment GBM. This occurs primarily within the first six months after completion of treatment, which includes surgical excision and chemoradiation with temozolomide. Accurate differentiation between TP and PsP is essential for assessing response to treatment and patient prognosis. This section reviews the role of ML in differentiating TP from PsP[75].

Many groups have investigated the role of feature-based radiomics in differentiating TP from PsP. Zhang and colleagues[76] used conventional MRI sequences to extract 285 radiomic features that were selected through concordance correlation coefficients to construct a model that would differentiate TP from PsP. Using five selected radiomics features, their model had an overall accuracy of 73.2% in predicting TP or PsP. Kim and colleagues[77] further incorporated diffusion- and perfusion-weighted MRI on top of conventional MR images to extract 6472 radiomic features from the enlarging contrast-enhancing portions of 61 GBM patients to predict TP vs. PsP. They used LASSO to select 12 significant radiomics features to build their model. This multiparametric radiomics model showed a robust performance in both external validation (AUC=0.85) and internal validation (AUC=0.96) cohorts for differentiating TP and PsP. Peng and colleagues[78] directly compared the performance of a radiomics-based model to a neuroradiologist in differentiating between TP and PsP. Their radiomics-based model extracted features from T1-weighted and T2-FLAIR sequences and top features were entered into a hybrid feature selection/classification model – i.e., IsoSVM. Images from 66 patients were used for performance evaluation and the model differentiated between TP and PsP with a sensitivity, specificity, and AUC of 65.4%, 86.67% and 0.81, respectively, on the validation cohort. In comparison, the neuroradiologist was only able to classify 73% of the cases with a sensitivity and specificity of 97% and 19%, respectively.

The role of nuclear medicine-based radiomics in differentiating TP from PsP has also been evaluated in literature. Lohmann and colleagues[79] investigated the potential of FET PET radiomics to discriminate between TP and PsP. Their study used data from 35 GBM patients who underwent a dynamic FET PET scan. Their final model utilized random forest regression for feature selection and the number of parameters was limited to three. They found that the diagnostic accuracy of the best single FET PET parameter (TBR_{max}) was 75% in differentiating TP from PsP. The highest accuracy was achieved by the three-parameter model, combining the dynamic parameter time-to-peak (TTP) with two radiomic features: 92% on the test cohort, and 86% on the validation cohort. In another study, Lohmann

and colleagues[80] compared the performance of contrast-enhanced MRI and FET PET in differentiating TP from PsP. They built radiomics-based models on both CE-MRI and FET PET and tested them on a cohort of 52 patients. Their results showed a diagnostic accuracy of 81% when using textural features extracted from contrast-enhanced MRI to differentiate TP from PsP. The accuracy of the FET PET model was slightly higher at 83%. The highest accuracy was achieved by combining contrast-enhanced MRI and FET PET features, which was 89%, with a sensitivity and specificity of 85% and 96%, respectively.

Groups have also investigated the role of deep learning and CNNs in differentiating between TP and PsP. Jang and colleagues[81] used MR images from 52 GBM patients to build three CNN models based on a CNN-LTSM structure: model 1 combined MRI data with clinical features, model 2 only included MRI data and model 3 was a random forest model with clinical features only. Model 1 had the best performance in differentiating TP and PsP with an AUC of 0.83.

Limitations and Future Considerations

Early evidence for the use of ML in clinical practice shows great promise, however there are limitations that prevent it from becoming a routine part of clinical work up. One of the factors limiting the routine use of ML is the burdensome process of image segmentation. There is no reliable and automated tumor-segmentation algorithm currently used, and few studies have significant validation for their attempts at automation of tumor segmentation. Future studies should look to develop such algorithms as an additional benefit of standardizing tumor segmentation would be the quantification of tumor volumes, which aids in evaluating treatment response.

Studies investigating ML are also impacted by a lack of reproducibility of their results, which likely stems from poor standardization in image acquisition and radiomics analysis workflow. A systematic review suggested that the repeatability and reproducibility of radiomic features are sensitive to processing details at various degrees[82]. The complexity behind image processing, feature extraction and the prediction algorithms add to difficulty in standardization and implementation of radiomic pipelines. Future studies should focus on the standardization of radiomics analysis, including image acquisition and tumor segmentation, before validating findings on large-scale, multi-centered patient cohorts that will require data sharing and collaboration. The nature of AI is such that it continuously refines algorithms based on the availability of data, and by providing access to varied, complete data, generalizable algorithms can be conceived, leading to the use of ML as a routine clinical tool in patient diagnosis.

Acknowledgments

Dr. Payabvash is supported by NIH/NINDS K23NS118056, Doris Duke Charitable Foundation (2020097), and Foundation of American Society of Neuroradiology.

Abbreviations

AI	artificial intelligence

ML	machine learning
CNN	convoluted neural network
SVM	support vector machine
LGG	low-grade glioma
HGG	high-grade glioma
LASSO	least absolute shrinkage and selection operator
GLCM	grey level co-occurrence matrix
LBP	local binary pattern
CAD	computer-aided detection
TCGA	the Cancer Genome Atlas
TCIA	the Cancer Imaging Archive
GBM	glioblastoma multiforme
IDH	isocitrate dehydrogenase
MGMT	O6-methylguanine-DNA methyltransferase
VASARI	Visually Accessible Rembrandt Imaging
XGBoost	eXtreme Gradient Boosting
mRMR	minimum redundancy maximum relevance
TP	true progression
PsP	pseudoprogression

References

1. Forghani R Precision Digital Oncology: Emerging Role of Radiomics-based Biomarkers and Artificial Intelligence for Advanced Imaging and Characterization of Brain Tumors. Radiol Imaging Cancer 2020;2(4):e190047. doi:10.1148/rycan.2020190047 [PubMed: 33778721]
2. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016;278(2):563–577. doi:10.1148/radiol.2015151169 [PubMed: 26579733]
3. Lohmann P, Galldiks N, Kocher M, et al. Radiomics in neuro-oncology: Basics, workflow, and applications. Methods 2021;188:112–121. doi:10.1016/j.ymeth.2020.06.003 [PubMed: 32522530]
4. Avanzo M, Wei L, Stancanello J, et al. Machine and deep learning methods for radiomics. Med Phys 2020;47(5). doi:10.1002/mp.13678
5. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–444. doi:10.1038/nature14539 [PubMed: 26017442]
6. Urbańska K, Sokołowska J, Szmidt M, Sysa P. Glioblastoma multiforme – an overview. Contemp Oncol 2014;18(5):307–312. doi:10.5114/wo.2014.40559
7. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol (Berl) 2016;131(6):803–820. doi:10.1007/s00401-016-1545-1 [PubMed: 27157931]

8. Cho H, Lee S, Kim J, Park H. Classification of the glioma grading using radiomics analysis. PeerJ 2018;6:e5982. doi:10.7717/peerj.5982 [PubMed: 30498643]
9. Sun X, Liao W, Cao D, et al. A logistic regression model for prediction of glioma grading based on radiomics. Zhong Nan Da Xue Xue Bao Yi Xue Ban 2021;46(4):385–392. doi:10.11817/j.issn.1672-7347.2021.200074 [PubMed: 33967085]
10. Sudre CH, Panovska-Griffiths J, Sanverdi E, 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. doi:10.1186/s12911-020-01163-5 [PubMed: 32631306]
11. Tian Q, Yan L-F, Zhang X, et al. Radiomics strategy for glioma grading using texture features from multiparametric MRI: Radiomics Approach for Glioma Grading. J Magn Reson Imaging 2018;48(6):1518–1528. doi:10.1002/jmri.26010 [PubMed: 29573085]
12. Huang W-Y, Wen L-H, Wu G, et al. Comparison of Radiomics Analyses Based on Different Magnetic Resonance Imaging Sequences in Grading and Molecular Genomic Typing of Glioma. J Comput Assist Tomogr 2021;45(1):110–120. doi:10.1097/RCT.0000000000001114 [PubMed: 33475317]
13. Li-Chun Hsieh K, Chen C-Y, Lo C-M. Quantitative glioma grading using transformed gray-scale invariant textures of MRI. Comput Biol Med 2017;83:102–108. doi:10.1016/j.compbiomed.2017.02.012 [PubMed: 28254615]
14. Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. J Digit Imaging 2013;26(6):1045–1057. doi:10.1007/s10278-013-9622-7 [PubMed: 23884657]
15. Gutta S, Acharya J, Shiroishi MS, Hwang D, Nayak KS. Improved Glioma Grading Using Deep Convolutional Neural Networks. AJNR Am J Neuroradiol 2021;42(2):233–239. doi:10.3174/ajnr.A6882 [PubMed: 33303522]
16. Wang G, Li W, Ourselin S, Vercauteren T. Automatic Brain Tumor Segmentation Based on Cascaded Convolutional Neural Networks With Uncertainty Estimation. Front Comput Neurosci 2019;13:56. doi:10.3389/fncom.2019.00056 [PubMed: 31456678]
17. Kabir Anaraki A, Ayati M, Kazemi F. Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. Biocybern Biomed Eng 2019;39(1):63–74. doi:10.1016/j.bbe.2018.10.004
18. Yang Y, Yan L-F, Zhang X, et al. Glioma Grading on Conventional MR Images: A Deep Learning Study With Transfer Learning. Front Neurosci 2018;12:804. doi:10.3389/fnins.2018.00804 [PubMed: 30498429]
19. Ertosun MG, Rubin DL. Automated Grading of Gliomas using Deep Learning in Digital Pathology Images: A modular approach with ensemble of convolutional neural networks. AMIA Annu Symp Proc 2015;2015:1899–1908. [PubMed: 26958289]
20. Zhuge Y, Ning H, Mathen P, et al. Automated glioma grading on conventional MRI images using deep convolutional neural networks. Med Phys 2020;47(7):3044–3053. doi:10.1002/mp.14168 [PubMed: 32277478]
21. He K, Gkioxari G, Dollar P, Girshick R. Mask R-CNN In: ; 2017:2961–2969. Accessed September 28, 2021. https://openaccess.thecvf.com/content_iccv_2017/html/He_Mask_R-CNN_ICCV_2017_paper.html
22. Özcan H, Emiroğlu BG, Sabuncuoğlu H, et al. A comparative study for glioma classification using deep convolutional neural networks. Math Biosci Eng 2021;18(2):1550–1572. doi:10.3934/mbe.2021080 [PubMed: 33757198]
23. Zhang Z, Xiao J, Wu S, et al. Deep Convolutional Radiomic Features on Diffusion Tensor Images for Classification of Glioma Grades. J Digit Imaging 2020;33(4):826–837. doi:10.1007/s10278-020-00322-4 [PubMed: 32040669]
24. Takahashi S, Takahashi W, Tanaka S, et al. Radiomics Analysis for Glioma Malignancy Evaluation Using Diffusion Kurtosis and Tensor Imaging. Int J Radiat Oncol Biol Phys 2019;105(4):784–791. doi:10.1016/j.ijrobp.2019.07.011 [PubMed: 31344432]
25. Pyka T, Gempt J, Hiob D, et al. Textural analysis of pre-therapeutic [18F]-FET-PET and its correlation with tumor grade and patient survival in high-grade gliomas. Eur J Nucl Med Mol Imaging 2016;43(1):133–141. doi:10.1007/s00259-015-3140-4 [PubMed: 26219871]

26. Haubold J, Demircioglu A, Gratz M, et al. Non-invasive tumor decoding and phenotyping of cerebral gliomas utilizing multiparametric 18F-FET PET-MRI and MR Fingerprinting. Eur J Nucl Med Mol Imaging 2020;47(6):1435–1445. doi:10.1007/s00259-019-04602-2 [PubMed: 31811342]
27. Jian A, Jang K, Manuguerra M, Liu S, Magnussen J, Di Ieva A. Machine Learning for the Prediction of Molecular Markers in Glioma on Magnetic Resonance Imaging: A Systematic Review and Meta-Analysis. Neurosurgery 2021;89(1):31–44. doi:10.1093/neuros/nyab103 [PubMed: 33826716]
28. Bhandari AP, Liong R, Koppen J, Murthy SV, Lasocki A. Noninvasive Determination of IDH and 1p19q Status of Lower-grade Gliomas Using MRI Radiomics: A Systematic Review. Am J Neuroradiol 2021;42(1):94–101. doi:10.3174/ajnr.A6875 [PubMed: 33243896]
29. Fukuma R, Yanagisawa T, Kinoshita M, et al. Prediction of IDH and TERT promoter mutations in low-grade glioma from magnetic resonance images using a convolutional neural network. Sci Rep 2019;9(1):20311. doi:10.1038/s41598-019-56767-3 [PubMed: 31889117]
30. Shofty B, Artzi M, Ben Bashat D, et al. MRI radiomics analysis of molecular alterations in low-grade gliomas. Int J Comput Assist Radiol Surg 2018;13(4):563–571. doi:10.1007/s11548-017-1691-5 [PubMed: 29270916]
31. Lu C-F, Hsu F-T, Hsieh KL-C, et al. Machine Learning–Based Radiomics for Molecular Subtyping of Gliomas. Clin Cancer Res 2018;24(18):4429–4436. doi:10.1158/1078-0432.CCR-17-3445 [PubMed: 29789422]
32. Han Y, Xie Z, Zang Y, et al. Non-invasive genotype prediction of chromosome 1p/19q co-deletion by development and validation of an MRI-based radiomics signature in lower-grade gliomas. J Neurooncol 2018;140(2):297–306. doi:10.1007/s11060-018-2953-y [PubMed: 30097822]
33. Zhou H, Chang K, Bai HX, et al. Machine learning reveals multimodal MRI patterns predictive of isocitrate dehydrogenase and 1p/19q status in diffuse low- and high-grade gliomas. J Neurooncol 2019;142(2):299–307. doi:10.1007/s11060-019-03096-0 [PubMed: 30661193]
34. Chang P, Grinband J, Weinberg BD, et al. Deep-Learning Convolutional Neural Networks Accurately Classify Genetic Mutations in Gliomas. Am J Neuroradiol 2018;39(7):1201–1207. doi:10.3174/ajnr.A5667 [PubMed: 29748206]
35. Li Z, Wang Y, Yu J, Guo Y, Cao W. Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma. Sci Rep 2017;7(1):5467. doi:10.1038/s41598-017-05848-2 [PubMed: 28710497]
36. Yan J, Zhang B, Zhang S, et al. Quantitative MRI-based radiomics for noninvasively predicting molecular subtypes and survival in glioma patients. NPJ Precis Oncol 2021;5(1):72. doi:10.1038/s41698-021-00205-z [PubMed: 34312469]
37. Eichinger P, Alberts E, Delbridge C, et al. Diffusion tensor image features predict IDH genotype in newly diagnosed WHO grade II/III gliomas. Sci Rep 2017;7(1):13396. doi:10.1038/s41598-017-13679-4 [PubMed: 29042619]
38. Lohmann P, Lerche C, Bauer EK, et al. Predicting IDH genotype in gliomas using FET PET radiomics. Sci Rep 2018;8(1):13328. doi:10.1038/s41598-018-31806-7 [PubMed: 30190592]
39. Zaragori T, Oster J, Roch V, et al. 18F-FDOPA PET for the non-invasive prediction of glioma molecular parameters: a radiomics study. J Nucl Med Off Publ Soc Nucl Med Published online May 20, 2021:jnumed.120.261545. doi:10.2967/jnumed.120.261545
40. Hegi ME, Diserens A-C, Gorlia T, et al. MGMT Gene Silencing and Benefit from Temozolomide in Glioblastoma. N Engl J Med 2005;352(10):997–1003. doi:10.1056/NEJMoa043331 [PubMed: 15758010]
41. Hegi ME, Liu L, Herman JG, et al. Correlation of O6-methylguanine methyltransferase (MGMT) promoter methylation with clinical outcomes in glioblastoma and clinical strategies to modulate MGMT activity. J Clin Oncol 2008;26(25):4189–4199. doi:10.1200/JCO.2007.11.5964 [PubMed: 18757334]
42. Huang W, Wen L, Wu G, et al. Radiological model based on the standard magnetic resonance sequences for detecting methylguanine methyltransferase methylation in glioma using texture analysis. Cancer Sci 2021;112(7):2835–2844. doi:10.1111/cas.14918 [PubMed: 33932065]

43. Li Z-C, Bai H, Sun Q, et al. Multiregional radiomics features from multiparametric MRI for prediction of MGMT methylation status in glioblastoma multiforme: A multicentre study. Eur Radiol 2018;28(9):3640–3650. doi:10.1007/s00330-017-5302-1 [PubMed: 29564594]
44. Xi Y, Guo F, Xu Z, et al. Radiomics signature: A potential biomarker for the prediction of MGMT promoter methylation in glioblastoma. J Magn Reson Imaging 2018;47(5):1380–1387. doi:10.1002/jmri.25860 [PubMed: 28926163]
45. Vils A, Bogowicz M, Tanadini-Lang S, et al. Radiomic Analysis to Predict Outcome in Recurrent Glioblastoma Based on Multi-Center MR Imaging From the Prospective DIRECTOR Trial. Front Oncol 2021;11:636672. doi:10.3389/fonc.2021.636672 [PubMed: 33937035]
46. Weller M, Tabatabai G, Kästner B, et al. MGMT Promoter Methylation Is a Strong Prognostic Biomarker for Benefit from Dose-Intensified Temozolomide Rechallenge in Progressive Glioblastoma: The DIRECTOR Trial. Clin Cancer Res 2015;21(9):2057–2064. doi:10.1158/1078-0432.CCR-14-2737 [PubMed: 25655102]
47. Le NQK, Do DT, Chiu F-Y, Yapp EKY, Yeh H-Y, Chen C-Y. XGBoost Improves Classification of MGMT Promoter Methylation Status in IDH1 Wildtype Glioblastoma. J Pers Med 2020;10(3):128. doi:10.3390/jpm10030128
48. Crisi G, Filice S. Predicting MGMT Promoter Methylation of Glioblastoma from Dynamic Susceptibility Contrast Perfusion: A Radiomic Approach. J Neuroimaging 2020;30(4):458–462. doi:10.1111/jon.12724 [PubMed: 32374045]
49. Korfiatis P, Kline TL, Lachance DH, Parney IF, Buckner JC, Erickson BJ. Residual Deep Convolutional Neural Network Predicts MGMT Methylation Status. J Digit Imaging 2017;30(5):622–628. doi:10.1007/s10278-017-0009-z [PubMed: 28785873]
50. Lu Y, Patel M, Natarajan K, et al. Machine learning-based radiomic, clinical and semantic feature analysis for predicting overall survival and MGMT promoter methylation status in patients with glioblastoma. Magn Reson Imaging 2020;74:161–170. doi:10.1016/j.mri.2020.09.017 [PubMed: 32980505]
51. Qian J, Herman MG, Brinkmann DH, et al. Prediction of MGMT Status for Glioblastoma Patients Using Radiomics Feature Extraction From 18F-DOPA-PET Imaging. Int J Radiat Oncol Biol Phys 2020;108(5):1339–1346. doi:10.1016/j.ijrobp.2020.06.073 [PubMed: 32634544]
52. Kong Z, Lin Y, Jiang C, et al. 18F-FDG-PET-based Radiomics signature predicts MGMT promoter methylation status in primary diffuse glioma. Cancer Imaging 2019;19(1):58. doi:10.1186/s40644-019-0246-0 [PubMed: 31426864]
53. Qian Z, Li Y, Wang Y, et al. Differentiation of glioblastoma from solitary brain metastases using radiomic machine-learning classifiers. Cancer Lett 2019;451:128–135. doi:10.1016/j.canlet.2019.02.054 [PubMed: 30878526]
54. Artzi M, Bressler I, Ben Bashat D. Differentiation between glioblastoma, brain metastasis and subtypes using radiomics analysis. J Magn Reson Imaging 2019;50(2):519–528. doi:10.1002/jmri.26643 [PubMed: 30635952]
55. Kniep HC, Madesta F, Schneider T, et al. Radiomics of Brain MRI: Utility in Prediction of Metastatic Tumor Type. Radiology 2019;290(2):479–487. doi:10.1148/radiol.2018180946 [PubMed: 30526358]
56. Kim Y, Cho H-H, Kim ST, Park H, Nam D, Kong D-S. Radiomics features to distinguish glioblastoma from primary central nervous system lymphoma on multi-parametric MRI. Neuroradiology. 2018;60(12):1297–1305. doi:10.1007/s00234-018-2091-4 [PubMed: 30232517]
57. Shrot S, Salhov M, Dvorski N, Konen E, Averbuch A, Hoffmann C. Application of MR morphologic, diffusion tensor, and perfusion imaging in the classification of brain tumors using machine learning scheme. Neuroradiology 2019;61(7):757–765. doi:10.1007/s00234-019-02195-z [PubMed: 30949746]
58. Niu L, Zhou X, Duan C, et al. Differentiation Researches on the Meningioma Subtypes by Radiomics from Contrast-Enhanced Magnetic Resonance Imaging: A Preliminary Study. World Neurosurg 2019;126:e646–e652. doi:10.1016/j.wneu.2019.02.109 [PubMed: 30831287]
59. Nakagawa M, Nakaura T, Namimoto T, et al. Machine learning based on multi-parametric magnetic resonance imaging to differentiate glioblastoma multiforme from primary cerebral

nervous system lymphoma. Eur J Radiol 2018;108:147–154. doi:10.1016/j.ejrad.2018.09.017 [PubMed: 30396648]

60. Dong F, Li Q, Xu D, et al. Differentiation between pilocytic astrocytoma and glioblastoma: a decision tree model using contrast-enhanced magnetic resonance imaging-derived quantitative radiomic features. Eur Radiol 2019;29(8):3968–3975. doi:10.1007/s00330-018-5706-6 [PubMed: 30421019]
61. Zhang S, Song G, Zang Y, et al. Non-invasive radiomics approach potentially predicts non-functioning pituitary adenomas subtypes before surgery. Eur Radiol 2018;28(9):3692–3701. doi:10.1007/s00330-017-5180-6 [PubMed: 29572634]
62. Chakrabarty S, Sotiras A, Milchenko M, LaMontagne P, Hileman M, Marcus D. MRI-based Identification and Classification of Major Intracranial Tumor Types by Using a 3D Convolutional Neural Network: A Retrospective Multi-institutional Analysis. Radiol Artif Intell 2021;3(5):e200301. doi:10.1148/ryai.2021200301 [PubMed: 34617029]
63. Pei L, Jones KA, Shboul ZA, Chen JY, Iftekharuddin KM. Deep Neural Network Analysis of Pathology Images With Integrated Molecular Data for Enhanced Glioma Classification and Grading. Front Oncol 2021;11:668694. doi:10.3389/fonc.2021.668694 [PubMed: 34277415]
64. Im S, Hyeon J, Rha E, et al. Classification of Diffuse Glioma Subtype from Clinical-Grade Pathological Images Using Deep Transfer Learning. Sensors 2021;21(10):3500. doi:10.3390/s21103500 [PubMed: 34067934]
65. Yang I, Aghi MK. New advances that enable identification of glioblastoma recurrence. Nat Rev Clin Oncol 2009;6(11):648–657. doi:10.1038/nrclinonc.2009.150 [PubMed: 19806145]
66. Prasanna P, Patel J, Partovi S, Madabhushi A, Tiwari P. Radiomic features from the peritumoral brain parenchyma on treatment-naïve multi-parametric MR imaging predict long versus short-term survival in glioblastoma multiforme: Preliminary findings. Eur Radiol 2017;27(10):4188–4197. doi:10.1007/s00330-016-4637-3 [PubMed: 27778090]
67. Kickingereder P, Burth S, Wick A, et al. Radiomic Profiling of Glioblastoma: Identifying an Imaging Predictor of Patient Survival with Improved Performance over Established Clinical and Radiologic Risk Models. Radiology 2016;280(3):880–889. doi:10.1148/radiol.2016160845 [PubMed: 27326665]
68. Park JE, Kim HS, Jo Y, et al. Radiomics prognostication model in glioblastoma using diffusion- and perfusion-weighted MRI. Sci Rep 2020;10:4250. doi:10.1038/s41598-020-61178-w [PubMed: 32144360]
69. Gilbert MR, Dignam JJ, Armstrong TS, et al. A Randomized Trial of Bevacizumab for Newly Diagnosed Glioblastoma. N Engl J Med 2014;370(8):699–708. doi:10.1056/NEJMoa1308573 [PubMed: 24552317]
70. Beig N, Patel J, Prasanna P, et al. Radiogenomic analysis of hypoxia pathway is predictive of overall survival in Glioblastoma. Sci Rep 2018;8(1):7. doi:10.1038/s41598-017-18310-0 [PubMed: 29311558]
71. Diehn M, Nardini C, Wang DS, et al. Identification of noninvasive imaging surrogates for brain tumor gene-expression modules. Proc Natl Acad Sci 2008;105(13):5213–5218. doi:10.1073/pnas.0801279105 [PubMed: 18362333]
72. Duregon E, Bertero L, Pittaro A, et al. Ki-67 proliferation index but not mitotic thresholds integrates the molecular prognostic stratification of lower grade gliomas. Oncotarget 2016;7(16):21190–21198. doi:10.18632/oncotarget.8498 [PubMed: 27049832]
73. Zeng A, Hu Q, Liu Y, et al. IDH1/2 mutation status combined with Ki-67 labeling index defines distinct prognostic groups in glioma. Oncotarget 2015;6(30):30232–30238. doi:10.18632/oncotarget.4920 [PubMed: 26338964]
74. Li Y, Qian Z, Xu K, et al. Radiomic features predict Ki-67 expression level and survival in lower grade gliomas. J Neurooncol 2017;135(2):317–324. doi:10.1007/s11060-017-2576-8 [PubMed: 28900812]
75. Brandsma D, van den Bent MJ. Pseudoprogression and pseudoresponse in the treatment of gliomas. Curr Opin Neurol 2009;22(6):633–638. doi:10.1097/WCO.0b013e328332363e [PubMed: 19770760]

76. Zhang Z, Yang J, Ho A, et al. A predictive model for distinguishing radiation necrosis from tumour progression after gamma knife radiosurgery based on radiomic features from MR images. Eur Radiol 2018;28(6):2255–2263. doi:10.1007/s00330-017-5154-8 [PubMed: 29178031]

77. Kim JY, Park JE, Jo Y, et al. Incorporating diffusion- and perfusion-weighted MRI into a radiomics model improves diagnostic performance for pseudoprogression in glioblastoma patients. Neuro-Oncol 2019;21(3):404–414. doi:10.1093/neuonc/noy133 [PubMed: 30107606]

78. Peng L, Parekh V, Huang P, et al. Distinguishing True Progression From Radionecrosis After Stereotactic Radiation Therapy for Brain Metastases With Machine Learning and Radiomics. Int J Radiat Oncol 2018;102(4):1236–1243. doi:10.1016/j.ijrobp.2018.05.041

79. Lohmann P, Elahmadawy MA, Werner J, et al. OS9.6 Diagnosis of pseudoprogression using FET PET radiomics. Neuro-Oncol 2019;21(Suppl 3):iii19. doi:10.1093/neuonc/noz126.064

80. Lohmann P, Kocher M, Ceccon G, et al. Combined FET PET/MRI radiomics differentiates radiation injury from recurrent brain metastasis. NeuroImage Clin 2018;20:537–542. doi:10.1016/j.nicl.2018.08.024 [PubMed: 30175040]

81. Jang B-S, Jeon SH, Kim IH, Kim IA. Prediction of Pseudoprogression versus Progression using Machine Learning Algorithm in Glioblastoma. Sci Rep 2018;8(1):12516. doi:10.1038/s41598-018-31007-2 [PubMed: 30131513]

82. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and Reproducibility of Radiomic Features: A Systematic Review. Int J Radiat Oncol Biol Phys 2018;102(4):1143–1158. doi:10.1016/j.ijrobp.2018.05.053 [PubMed: 30170872]

Table 1

Studies Investigating the Role of AI in Grading Gliomas

Study	Purpose	Number of Patients	Findings
Cho et al.[8] (2018)	Grading gliomas (HGG vs LGG) using multimodal MRI-based radiomics	285 (HGG n = 210; LGG n = 75), BraTS challenge 2017	• Three classifiers showed average AUC = 0.94 in training set, and AUC = 0.90 in test cohort • Tumor morphological features + GLCM were most effective at discriminating HGG vs LGG
Hsieh et al.[13] (2017)	Grading gliomas by developing CAD system based on intensity-invariant MRI (achieved by converting MR images to local binary pattern)	107 (n = 34 GBM; n = 73 LGG)	• CAD system based on LBP features had accuracy = 93%, sensitivity = 97%, NPV = 99% and AUC = 0.94 • Conventional texture features had accuracy = 84%, sensitivity = 76%, NPV = 89% and AUC = 0.89
Tian et al.[11] (2018)	Distinguishing HGG vs LGG, and Grade III vs Grade IV gliomas using multiparametric MRI and evaluate the grading potential of different MRI sequences	153 (n = 42 Grade II; n = 33 Grade III; n = 78 Grade IV)	• SVM models established using 30 and 28 optimal features for HGG vs LGG and Grade III vs IV gliomas, respectively • Differentiating HGG vs LGG accuracy = 96.8%, AUC = 0.987 • Differentiating Grade III vs IV gliomas accuracy = 98.1%, AUC = 0.992 • Multiparametric MRI was more useful than histogram parameters or single sequence MRI
Pyka et al.[25] (2015)	Using textural FET-PET features for grading and prognostication of patients with HGG	113	• All FET-PET parameter differentiated between grade III and IV tumors (AUC = 0.775) • Combination of texture and metabolic tumor volume graded HGG with an accuracy of 85% (AUC = 0.830)
Yang et al.[18] (2018)	Differentiating HGG vs LGG by training CNN (AlexNet & GoogLeNet) on MR images	113	• GoogLeNet performance: validation accuracy = 0.87, test accuracy = 0.91, test AUC = 0.94 • Performances improved with transfer learning and fine tuning of both AlexNet and GoogLeNet (validation accuracy = 0.87 and 0.87; test accuracy 0.93 and 0.95; test AUC = 0.97 and 0.97, respectively)
Gutta et al.[15] (2020)	Comparing performance of features learned from CNN with standard radiomic features for glioma grade prediction	237	• CNN-learned features predicted glioma grade with an average accuracy of 87% • Top performing ML model (gradient boosting) predicted with average accuracy of 64%
Takahashi et al.[24] (2019)	Grading gliomas (GBM vs LGG) using ML based on multiparametric MRI including DTI	54 (n = 14 grade II glioma; n = 12 grade III glioma; n = 29 GBM)	• Most accurate ML model was created using 6 features extracted from ADC and MK images with test data accuracy = 0.91 and AUC 0.90
Zhang et al.[23] (2020)	Grading gliomas (LGG vs HGG and Grade III vs IV gliomas) using both deep neural networks and standard radiomics based on DTI	108 (n = 43 LGG; n = 65 HGG)	• Combining FA+MD had the best performance with accuracy = 94% and AUC = 0.93 in differentiating LGG from HGG; accuracy = 98% and AUC 0.99 in classifying grade III vs IV gliomas • Deep learning features are more predictive of glioma grades than conventional texture and morphological features
Haubold et al.[26] (2020)	Grading gliomas and predicting mutational status using FET-PET MRI-based radiomics	42	• Differentiating LGG vs HGG AUC = 0.85 • Predicting ATRX mutation AUC = 0.85 • Predicting MGMT mutation AUC = 0.75

Study	Purpose	Number of Patients	Findings
			• Predicting IDH1 mutation AUC = 0.89 • Predicting 1p19q mutation AUC = 0.98
Zhuge et al.[20] (2020)	Automatic differentiation of LGG vs HGG on conventional MRI images by using CNNs	(TCIA LGG data, BraTS Benchmark 2018)	• 2D Mask R-CNN based method had accuracy = 96.3%, sensitivity = 93.5% and specificity = 97.2% • 3DConvNet method had accuracy = 97.1%, sensitivity = 94.7% and specificity = 96.8%
Ozcan et al.[22] (2021)	Compare the performance of custom trained CNN against pretrained models in predicting LGG vs HGG grade of gliomas	104	• Custom model predicted HGG vs LGG with an accuracy = 97.1%, AUC = 0.99, sensitivity = 98% and specificity = 96.3% • GoogLeNet had the best performance of pretrained models with accuracy = 93.3%, AUC = 0.99, sensitivity = 98% and specificity = 89%
Huang et al.[12] (2021)	Distinguishing LGG vs HGG, IDH1 mutation status and MGMT mutation status using MRI-based radiomics and comparing the utility of each sequence	59	• CE-T1WI sequence performed best compared to other sequences alone in predicting tumor grade and IDH1 status of glioma • T2WI sequence performed best in predicting MGMT methylation status of glioma
Sun et al.[9] (2021)	Evaluating the role of logistic regression model based on radiomics to predict glioma grade	146	• 5 imaging features selected by LASSO were used for the logistic regression model and had an AUC = 0.92 for grading gliomas • Hosmer-Lemeshow test was used to measure accuracy and it showed no significant difference between the calibration and ideal curve ($P = 0.808$) indicating high predictive accuracy of the model
Sudre et al.[10] (2020)	Using DSC-MRI-based radiomics to differentiate WHO grades of gliomas and IDH1 mutation status and the utility of each feature	333	• Shape, distribution and texture features were best at predicting IDH1 mutation status • Grade II vs III differentiation was best achieved through shape features • Grade III vs IV differentiation was best achieved through intensity and texture features • IDH1 mutation prediction accuracy = 71% • Glioma grade prediction accuracy = 53% (87% of cases received grade classification with distance less than or equal to 1)

Table 2:

Studies investigating the role of AI predicting IDH mutation and 1p19q codeletion status

Study	Purpose	Number of Patients	Findings
Shofty et al.[30] (2017)	Evaluating radiomics classifiers and different MR contrasts in predicting 1p19q codeletion status in LGG patients	47	• Best classification occurred via the Ensemble Bagged Trees classifier with accuracy = 87%, AUC = 0.87, sensitivity = 92%, and specificity = 83%
Han et al.[32] (2018)	Predicting 1p19q codeletion status using MRI-based radiomic features in LGG	277	• Radiomics signature generated via random forest algorithm • Radiomics signature independently had best performance in predicting 1p19q codeletion status with AUC of 0.89 and 0.76 in training and validation cohorts, respectively • Clinical model had AUC of 0.58 and 0.63 in training and validation cohorts, respectively • Combined model had AUC of 0.89 and 0.75 in training and validation cohort, respectively
Zhou et al.[33] (2019)	Predicting IDH mutation and 1p19q codeletion status based on MRI-based radiomics features in glioma patients	538 (3 separate institutions)	• Model predicted IDH mutation with an AUC of 0.92 and 0.92 in training and validation cohorts, respectively • Overall accuracy of 3 group prediction (IDH-wild type, IDH mutant + 1p19q codeletion, IDH mutant + 1p19q non-deletion) was 78.2%
Lu et al.[31] (2018)	Predicting IDH mutation and 1p19q codeletion status based on multiparametric MRI-based radiomics features in glioma patients	456 from TCIA	• Binary classification of IDH and 1p19q status of gliomas was predicted with AUCs between 0.92 and 0.98, and accuracies between 87.7% and 96.1% on the training set • Accuracies ranged between 80.0% and 91.7% on the validation dataset
Lohmann et al.[38] (2018)	Predicting IDH genotype in gliomas using FET-PET based radiomics and in combination with textural features	84	• Prediction accuracy by combining conventional FET-PET parameters with textural features = 93% (sensitivity = 91%, specificity = 94%) • Accuracy based on FET PET standard parameters = 79% (AUC = 0.84) • Accuracy based on FET PET textural features = 79% (AUC = 0.84)
Eichinger et al.[37] (2017)	Predicting IDH genotype in LGG using DTI-based ML	79	• Single hidden layer neural network was trained on texture features and predicted IDH status with accuracy of 92% (AUC = 0.92) in training set and accuracy of 95% (AUC = 0.95) in validation set
Chang et al.[34] (2018)	Training a CNN using MRI to predict IDH1, 1p19q and MGMT mutation status in gliomas	259 from TCIA	• Classifying IDH1 mutation status had accuracy of 94% (AUC = 0.91) • Classifying 1p19q codeletion status had accuracy of 92% (AUC = 0.88) • Classifying MGMT methylation status had accuracy of 83% (AUC = 0.81)
Li et al.[35] (2017)	Predicting IDH1 status in LGG patients using deep learning-based radiomics and comparing performance to conventional radiomics	151	• AUC of IDH1 prediction using conventional radiomics = 0.86 • AUC of IDH1 prediction using deep learning-based radiomics = 0.92 • AUC of IDH1 prediction using deep learning based on multiple-modality MR images = 0.95

Study	Purpose	Number of Patients	Findings
Zaragori et al.[39] (2021)	Predicting IDH and 1p19q mutation status in glioma patients using 18F-FDOPA PET-based radiomics	72	• Best models predicted IDH mutation and 1p19q codeletion with an AUC of 0.83 and 0.72, respectively • Dynamic features were the most important in predicting IDH mutation (TTP = 35.5%) • Other radiomic features were the most important in predicting 1p19q codeletion status (up to 14.5% of importance for the small zone low grey level emphasis)
Yan et al.[36] (2021)	Classifying gliomas into molecular groups based on IDH mutation, 1p19q codeletion and TERT promoter mutation status using advanced MRI-based radiomics	357	• Image fusion model incorporating radiomic signatures from CE-T1WI and ADC achieved AUC of 0.88 and 0.67 for predicting IDH and TERT status, respectively • CE-T1WI-based radiomic signature alone had best performance in predicting 1p19q codeletion status with AUC = 0.82
Fukuma et al.[29] (2019)	Comparing MRI-based pretrained CNN and conventional radiomics in predicting IDH and TERT mutations for patients with LGG	164	• Prediction of IDH mutation was best using combination of CNN + radiomics + patient age (accuracy = 73.1%) • Characterization of LGG into 3 molecular subtypes based on IDH and TERT status was best using combination of CNN + radiomics + patient age (accuracy = 63.1%), however was not significantly different from using CNN alone (accuracy = 62.1%) • Prediction of TERT promoter mutation was best using CNN-features exclusively (accuracy = 84%)

Table 3:

Studies investigating the role of AI in predicting MGMT promoter methylation status of glioma patients

Study	Purpose	Number of Patients	Findings
Li et al.[43] (2018)	To build a radiomics model from multiregional and multiparametric MRI to predict MGMT promoter methylation status in GBM patients	193 (multicenter)	• Radiomics model with minimal set of 6 all-relevant features predicted MGMT status with accuracy of 80% (AUC = 0.88) • Radiomics model with 8 univariately-predictive and non-redundant features predicted MGMT status with accuracy of 70% (AUC = 0.76) • Combining clinical features with radiomic features did not significantly improve performance
Xi et al.[44] (2018)	To analyze utility of MRI-based radiomics features in predicting MGMT promoter methylation status in GBM patients	98 (n = 48 methylated; n = 50 unmethylated)	• Best performance for predicting MGMT status was achieved by combining T1WI, T2WI and CE-T1WI (accuracy = 86.6%) • Radiomic features of T1WI had accuracy of 67.6% • Radiomic features of CE-T1WI had accuracy of 82% • Radiomic features of T2WI had accuracy of 69.3%
Qian et al.[51] (2020)	Using 18F-DOPA PET-based radiomics to predict MGMT status in GBM patients	86	• Radiomics signature to predict MGMT methylation status using features extracted from GBM contour alone had accuracy of 80% • Prediction accuracy was not improved with additional input features
Kong et al.[52] (2019)	Using 18F-FDG PET-based radiomics to predict MGMT status in diffuse glioma patients	107	• Radiomics signature had the best performance with accuracy of 91.3% and 77.8% (AUC of 0.94 and 0.86) in the primary and validation cohorts, respectively • Clinical model had accuracy 64.8% and 66.4% in the primary and validation cohort, respectively • Fusion model had accuracy of 64.8% and 72.7% in the primary and validation cohort, respectively
Huang et al.[42] (2021)	Predicting MGMT methylation status in gliomas using MR-based radiomics with textural features	53	• Combined radiomics model using multiparametric MRI predicted MGMT methylation status with AUC, sensitivity, and specificity of 0.82, 90.5% and 72.7%, respectively in the GBM dataset • AUC, sensitivity, and specificity of 0.83, 70.2% and 90.6% in the overall glioma dataset
Vils et al.[45] (2021)	Predicting MGMT methylation status using multi-center MRI-based radiomics in recurrent GBM patients	69 (DIRECTOR trial)	• CE-T1W MRI-based radiomic model to predict MGMT status was established using linear intensity interpolation and had AUC of 0.67 in both training and validation cohorts
Korfiatis et al.[49] (2017)	Comparing three different ResNet architectures in predicting MGMT methylation status without distinct tumor segmentation step	155 (n = 66 methylated; n = 89 unmethylated tumors)	• ResNet50 (50 layers) was the best performing model with prediction accuracy of 94.9% on test set • ResNet34 (34 layers) achieved an accuracy of 80.7% • ResNet18 (18 layers) achieved an accuracy of 76.8%
Le et al.[47] (2020)	Evaluating a novel radiomics-based XGBoost model to identify MGMT methylation status in IDH1 wildtype GBM patients	53	• 9 radiomics features were extracted from multimodality MRI for model construction

Study	Purpose	Number of Patients	Findings
			• XGBoost classifier predicted MGMT status with accuracy of 88.7%, AUC of 0.896, sensitivity of 88% and specificity of 89%
Crisi & Filice[48] (2020)	Stratification of MGMT methylation status in GBM patients using DSC-MRI-based radiomics features	59	• Used 14 radiomics features to build a multilayer deep learning model that classified MGMT methylation status into 3 groups • Their model had AUC, sensitivity, and specificity of 0.84, 75% and 85%, respectively
Lu et al.[50] (2020)	Combining MRI based-radiomic, semantic and clinical features to improve prediction of MGMT methylation status in GBM patients	181 MRI studies	• Optimal cut-off value for MGMT promoter methylation index was 12.75% • Their model combined radiomic, VASARI and clinical features to predict MGMT status and had an accuracy that varied between 45% and 67%

Table 4:

Studies investigating the use of ML to differentiate between types of brain tumors

Study	Purpose	Number of Patients	Findings
Kim et al.[56] (2018)	Differentiate GBM vs. primary central nervous system lymphoma (PCNSL) using multiparametric MRI-based radiomics	143 patients (n = 86 training; n = 57 validation)	• 15 features used in final model • AUC validation = 0.956 • AUC training = 0.979
Shrot et al.[57] (2019)	Differentiating different brain tumors using basic and advanced MRI-based radiomics	141 patients (41 GBM, 38 METS, 50 meningioma & 12 PCNSL)	• Classification used morphologic MRI, perfusion MRI & DTI metrics • Feature subset selection via SVMs • Binary SVM classification accuracy ranged from 81.6 to 97.0
Niu et al.[58] (2019)	Differentiating between different meningioma subtypes using basic MRI-based radiomics	241 patients (n = 80 meningiothelial meningioma, n = 80 fibrous meningioma, n = 81 transitional meningioma)	• Fisher discriminant analysis model for binary differentiation between meningioma types had accuracies between 98.8% and 100% • Leave one out cross validation had accuracies between 91.3% and 100%
Nakagawa et al.[59] (2018)	Differentiating GBM vs PCNSL using ML method based on texture features in multiparametric MRI	70 patients	• Prediction model developed using univariate logistic regression and XGBoost • rCBV offered highest AUC of 0.86 (rCBV AUC = 0.83; skewness of CE-T1WI AUC = 0.78) • AUC of XGBoost was significantly higher than that of two radiologists (0.98 vs 0.84).
Dong et al.[60] (2019)	Differentiating between pilocytic astrocytoma (PA) and GBM using MRI quantitative radiomic features by a decision tree model	66 patients (PA n = 31; GBM n = 35)	• Subset of 12 features selected by feature stability and Boruta algorithm to build decision tree model • Training set: accuracy = 87%; sensitivity = 90%; specificity = 83% • Validation set: accuracy = 86%, sensitivity = 80%; specificity = 91%
Zhang et al.[61] (2018)	Using MRI-based radiomics to differentiate between non-functioning pituitary adenoma subtypes	112 patients (training set n = 75; test set n = 37)	• T1WI had AUC of 0.83 and 0.80 in training and test sets, respectively • CE-T1WI features added no additional value to model
Chakrabarty et al.[62] (2021)	Train a CNN to differentiate between tumor types (HGG, LGG, metastases, meningioma, pituitary adenoma, acoustic neuroma & healthy tissue)	1373 (BraTs, TCGA, LGG-1p19q dataset, internal and external dataset)	• Internal data set: sensitivities, PPVs, AUCs and area under the precision-recall curves (AUPRCs) ranged from 87%−100%, 85% to 100%, 0.98 to 1.0, and 0.91 to 1.0, respectively • External data set: sensitivities, PPVs, AUCs and AUPRCs ranged from 91% to 97%, 73% to 99%, 0.97 to 0.98, and 0.9 to 1.0, respectively
Qian et al.[53] (2019)	To identify the optimal radiomic ML classifier for differentiating GBM vs METS	412	• SVM + LASSO classifier had highest prediction efficacy (AUC = 0.90, accuracy = 82.7%, sensitivity = 79.8%, specificity = 87.3%, PPV = 90%, and NPV = 72.9%)
Artzi et al.[54] (2019)	To differentiate between GBM and METS using CE-T1WI MRI-based radiomics	439	• Best results for differentiating GBM vs. METS were obtained using SVM classifier which had a mean accuracy, AUC, sensitivity, and specificity of 85%, 0.96, 86% and 85%, respectively

Study	Purpose	Number of Patients	Findings
			• Optimal differentiation of GBM and METS subtypes achieved using SVM classifier with accuracy, AUC, sensitivity, and specificities ranging between 75%−90%, 0.57–0.98, 11% −100%, and 76%−99% respectively,
Kniep et al.[55] (2018)	Using multiparametric MRI-based radiomics to predict tumor type in brain metastasis (SCLC, BC, MM, GC and NSCLC)	189	• AUC for predicting type of brain metastasis ranged between 0.64 (NSCLC) and 0.82 (MM) • Prediction performance of classifier was superior to radiologists' readings • MM had highest increase in sensitivity (17%) using classifier compared to radiologists' readings