Neuro-Oncology Advances
5(1), 1–11, 2023 | https://doi.org/10.1093/noajnl/vdad045 | Advance Access date 19 April 2023

Predicting methylation class from diffusely infiltrating adult gliomas using multimodality MRI data

Zahangir Alom, Quynh T. Tran, Asim K. Bag, John T. Lucas, and Brent A. Orr

Department of Pathology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA (Z.A., Q.T., B.A.O.); Department of Diagnostic Imaging, St. Jude Children's Research Hospital, Memphis, Tennessee, USA (A.K.B.); Department of Radiation Oncology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA (J.T.L.)-

Corresponding Author: Brent A. Orr MD, PhD, Department of Pathology, St. Jude Children's Research Hospital, 262 Danny Thomas Place, MS 250, Memphis, TN 38-105-3678, USA (brent.orr@stjude.org).

Abstract

Background. Radiogenomic studies of adult-type diffuse gliomas have used magnetic resonance imaging (MRI) data to infer tumor attributes, including abnormalities such as IDH-mutation status and 1p19q deletion. This approach is effective but does not generalize to tumor types that lack highly recurrent alterations. Tumors have intrinsic DNA methylation patterns and can be grouped into stable methylation classes even when lacking recurrent mutations or copy number changes. The purpose of this study was to prove the principle that a tumor's DNA-methylation class could be used as a predictive feature for radiogenomic modeling.

Methods. Using a custom DNA methylation-based classification model, molecular classes were assigned to diffuse gliomas in The Cancer Genome Atlas (TCGA) dataset. We then constructed and validated machine learning models to predict a tumor's methylation family or subclass from matched multisequence MRI data using either extracted radiomic features or directly from MRI images.

Results. For models using extracted radiomic features, we demonstrated top accuracies above 90% for predicting IDH-glioma and GBM-IDHwt methylation families, IDH-mutant tumor methylation subclasses, or GBM-IDHwt molecular subclasses. Classification models utilizing MRI images directly demonstrated average accuracies of 80.6% for predicting methylation families, compared to 87.2% and 89.0% for differentiating IDH-mutated astrocytomas from oligodendrogliomas and glioblastoma molecular subclasses, respectively.

Conclusions. These findings demonstrate that MRI-based machine learning models can effectively predict the methylation class of brain tumors. Given appropriate datasets, this approach could generalize to most brain tumor types, expanding the number and types of tumors that could be used to develop radiomic or radiogenomic models.

Key Points

- Radiomic and radiogenomic tumor classification has potential clinical benefit but is currently available only for tumors with highly recurrent genomic abnormalities.
- Using the TCGA diffuse glioma dataset, this study demonstrates the proof of principle that a tumor's methylation class can effectively be used for radiomic modeling.
- This approach has potential to generalize to most brain tumors with appropriate matched methylation and imaging data.

Importance of the Study

Classification of brain tumors increasingly relies on molecular testing of tumor tissue acquired during an invasive surgical procedure. In some instances, surgical intervention may be infeasible or medically contraindicated. In addition, the knowledge of a brain tumor's molecular type in advance may inform surgical approach or facilitate adjuvant therapy. Radiomic models could facilitate clinical decision making in the absence of surgical intervention but have largely been restricted to those that model-specific histomorphologic tumor types or tumors with highly recurrent genomic abnormalities. In this proof of principle study, we demonstrate that a tumor's methylation class can be used in image classification models derived from extracted radiomic features or directly from MRI images. This represents a generalizable approach that could be applied to most brain tumors and will significantly expand the number and types of tumors that could be used for radiomic modeling.

Diffusely infiltrating gliomas are the most common brain tumors in adults.[1] Significant advancement in the molecular classification of infiltrating gliomas has been made to improve both diagnosis and risk stratification.[2] For example, the *WHO Classification of the Tumours of Central Nervous System* in 2016 was updated with IDH-mutant and IDH-wildtype tumor classes[3] based on the presence or absence of mutations in the isocitrate dehydrogenase genes (IDH1 or IDH2), making them an important predictor of clinical risk.[4–6] Among IDH-mutant tumors, specific histopathologic subtypes are also defined based on molecular abnormalities in *TP53/ATRX* (astrocytomas) or co-deletion of 1p and 19q and *TERT* promoter (oligodendrogliomas).[2,6]

As molecular classification of tumors has improved, there has also been significant interest in predicting molecular features of human tumors from medical imaging data including magnetic resonance imaging (MRI) images.[7,8] Successful implementation of these approaches has potential to improve preoperative risk stratification and to guide therapeutic intervention when biopsy or tumor resection is medically contraindicated or impractical. Using extracted radiomic features or structural imaging, models have successfully predicted highly recurrent and class-specific genomic features, including IDH mutation status or 1p19q co-deletion status in diffuse gliomas.[8–10] Although promising, the previous approaches do not generalize outside specific use cases because many tumor types do not have defined abnormalities or are characterized by molecular heterogeneity that makes prediction based on individual genomic abnormalities infeasible. Approaches that incorporate molecular class information rather than specific mutations or copy number abnormalities could expand the tumor types available for radiomic modeling and could represent a more generalizable approach to molecular class prediction from radiology image data.

Similar to molecular classification of tumors by their genetic abnormalities, brain tumors and other solid tumors can also be characterized epigenetically by highly specific genome-wide DNA methylation signatures. These DNA methylation signatures are believed to represent a combination of the methylation pattern derived from a tumor's cell of origin and those derived from specific tumor driving abnormalities.[11] Tumor-specific methylation signatures have been successfully exploited to resolve novel brain tumor entities and to uncover molecular subtypes within established histopathologically defined tumor types.[12–14] More recently, supervised classification models have been deployed in the clinical diagnostic laboratory, extending the utility of methylation profiling.[15–18] The approach has some advantages over traditional histopathologic classification including, objective class designation with less opportunity for interobserver variability, the ability to modulate clinical classification stringency through thresholding, and classification may be possible even in tumors lacking defining genomic abnormalities.[15,19]

Despite the adoption of methylation profiling in several clinical laboratories, only selected studies have incorporated methylation class labels into radiogenomic models.[20,21] Progress is significantly impeded by a paucity of datasets with matched methylation and imaging data. Furthermore, the few studies that have utilized methylation class in imaging models have relied on extracted radiomic features within specific histomorphologic tumor types such as glioblastoma.

Herein, we used a custom deep neural network (DNN) model to assign methylation class to the TCGA adult-type diffuse glioma datasets, expanding the public datasets with class labels matched to MRI images. Using the molecular class labels, we evaluated whether machine learning models could be constructed to predict methylation class in our cohort from either extracted radiomic features or directly from multisequence MRI data. Using explainable methods, we evaluated the regional and MRI sequence specific importance in our predictive models.

Methods

Ethics Statement

The data utilized in this study are publicly available through The Cancer Genome Atlas and The Cancer Imaging Archive and exempt from institutional review board oversight.

Processing of Methylation Data

All methylation data were processed from raw IDAT files. Illumina Infinium 450K methylation data from the TCGA Lower grade glioma (LGG) and glioblastoma (GBM) datasets were downloaded from https://portal.gdc.cancer.gov/.[22]

The brain tumor reference methylation cohort (GSE90496 [15]) was downloaded from Gene Expression Omnibus (GEO) database. Raw idat files were processed in R (http://www.r-project.org, version 4.0.2), using several packages from Bioconductor and other repositories. Specifically, array data were preprocessed using the *minfi* package (v.1.36.0).[23] Background correction with dye-bias normalization was performed for all samples using noob (normal-exponential out-of-band) with "single" dye method[24] with preprocessFunNorm. Probe filtering was performed after normalization. Specifically, probes located on sex chromosomes containing a nucleotide polymorphism (dbSNP132 Common) within five base pairs of and including the targeted CpG-site, or mapping to multiple sites on hg19 (allowing for one mismatch), and cross-reactive probes were removed from the analysis. To make the DNA methylation model generalizable to both 450K and 850K/EPIC arrays, only overlapping probes between the two arrays were used. After the filtering process, 408,862 probes remained.

Construction of Supervised Classification Model for DNA Methylation Data

The DNN for DNA methylation-based tumor classification model was constructed in Python using the scikit-learn package,[25,26] using a public reference dataset (GSE90496) for training. A detailed description of the training and validation of the model can be found in Supplementary Material.

Assignment of Methylation Class for TCGA Samples

To obtain methylation class labels from the diffuse gliomas cohorts of the TCGA, consisting of tumors in the lower-grade glioma (LGG) and glioblastoma (GBM) datasets, the preprocessed methylation data were subset to include only probes selected for the methylation-based classification model. For each tumor, a class label was assigned representing the top output score from the DNA methylation-based DNN model. Family labels were assigned by addition of the respective subclass scores, as previously described.[15] Tumors assigned to an adult-type diffuse glioma subtypes with subclass classification scores above 0.8 were deemed high-confidence and carried forward for subsequent analysis and MRI modeling. Validation of the class assignment for the high-confidence samples was performed by comparing the class specific summarized metadata downloaded from the TCGA portal.[2]

Classification Models from Defined Radiomic Features

Radiomic features from TCGA MRI data were downloaded from cancerimagingarchive.net and matched to tumors with high-confidence methylation class labels.[27] The 724 radiomic features were normalized in python using the min-max normalization method in scikit-learn. Three machine learning model types were deployed, including a DNN model implemented in Tensorflow and a random forest (RF) classifier and support vector machine (SVM) implemented using scikit-learn. Models were trained to predict the IDH-mutant versus glioblastoma, IDH-wildtype (GBM-IDHwt) methylation families, the astrocytoma, IDH-mutant (AIDH) versus oligodendroglioma, IDH-mutant (OIDH) methylation subclasses, and the glioblastoma, IDHwt, subclass Receptor Tyrosine Kinase I/II (GBM-RTK) versus glioblastoma, IDHwt, subclass mesenchymal (GBM-MES) groups. Each DNN model had an input dimension of 724 and an output dimension of 2, with seven total layers (724, 500, 300, 200, 100, 50, 2 nodes at each layer, respectively). Each layer was fully connected with a ReLU activation function followed by a batch normalization and dropout layers (0.3) except the final layer which utilized a SoftMax layer for classification. The RF and SVM models were trained using scikit-learn in python, with default parameters.[25] Each model was evaluated using the average accuracy of 10 bootstraps following five-fold cross validation.

MRI Preprocessing Pipeline

For direct MRI-based classification models, preoperative multisequencing MRI data and tumor segmentation masks were downloaded from the cancer imaging archive (https://www.cancerimagingarchive.net).[27,28] T2-weighted (T2WI), T2-weighted-fluid attenuated inversion recovery (FLAIR), T1 with contrast enhancement (T1GD), and pre-contrast T1 weighted MRI images (T1WI) from 78 patients meeting

Table 1. Demographic Features of TCGA MRI Cohort

Methylation family	Methylation subclass	# Included	Gender (M:F)	Mean Age (range)
IDH-glioma	OIDH	13	(5:8)	54.363 (28-74)
IDH-glioma	AIDH	38	(19:19)	38.66 (21-67)
GBM-IDHwt	GBM-RTK	13	(7:6)	60.2(50-72)
GBM-IDHwt	GBM-MES	14	(8:6)	54.38(46-74)

IDH-glioma, methylation family IDH glioma; GBM-IDHwt, methylation family glioblastoma, IDH-wildtype; AIDH, methylation subclass astrocytoma; OIDH, methylation subclass 1p19q co-deleted oligodendroglioma; GBM-RTK, methylation subclass receptor tyrosine kinase I/II; GBM-MES, methylation subclass mesenchymal.

our inclusion criteria were utilized for modeling. The demographic information describing the TCGA MRI cohort is provided in Table 1. For each tumor, the MRI sequences were registered and up to seven axial slices (slice with greatest area of the segmentation mask, and three slices superior and inferior) were used for model training. The MRI images were normalized using *Z*-score normalization prior to model training, as previously described.[29]

End-to-end MRI Molecular Class Prediction

For direct MRI-based prediction models, processed MRI images from T2 alone or multimodality (MM) MRI sequences were used to train a modified version of VGG-net, ResNET50, and denseNET deep convolutional neural network (DCNN) models. For the VGG-Net model architecture, we used a modified version of VGG-Net with six convolutional neural networks (CNN) units. Each CNN unit consists of a convolutional layer with Exponential Linear Unit (ELU) activation function, batch normalization layer, and dropout layer. The Global Average Pooling (GAP) layer was used instead of fully connected layers to reduce the number of network parameters. Finally, a SoftMax layer was used at the end of the classification model. We used the default architecture of ResNet50 for the input size of 240 × 240 pixels[30] and DenseNet[31] with a number of dense blocks of 4, a growth rate of 64, number of filters of 128, depth of 7, and a dropout rate of 0.2, respectively.

Separate models were constructed to predict either methylation family, IDH-glioma subclasses, or GBM-IDHwt subclasses. For the T2 specific models, we evaluated the performance of the models with only T2 modality with input dimension of $N \times 240 \times 240 \times 7$, where N stand for number of cases. Eventually, the seven slices from each modality were merged for implementing the multimodality-based learning system. Hence, input dimension of the multimodal system was $N \times 240 \times 240 \times 28$.

For all CNN models, the following training features were utilized: Adam optimization with default parameters, categorical cross entropy loss function, batch size of 4, 150 epochs, patience of 15 used to model performance with respect to validation loss. Ten-fold bootstrapping was performed on five-fold cross validation, maintaining representation of each appropriate tumor classes in each split. Accuracy was reported as the mean bootstrapped accuracy. The deep learning features were extracted from the bottleneck layers of all the models. The dimension of the feature representation of the bottleneck layers were (7 × 7 × 1024), (8 × 8 × 512), and 8 × 8 × 512 for VGG-Net, ResNet50, and DenseNet models, respectively.

Interpretability Analysis for End-to-end Family and Subclass Models

For interpretability models, an output heatmap was generated from the feature representation of the bottleneck layer of the VGG-net model and then the cumulative importance was represented as an importance per area of tissue mask for the tumor and nontumor brain regions. The steps for generating the final outputs heatmaps and the calculation of the importance scores for each region are represented as pseudo-code in Supplementary Methods and described in Procedure S1.

Statistical Analysis

Statistical analysis was performed using the SciPy and bioinfokity packages in python.[32] For continuous variables, one-way analysis of variance (ANOVA) was performed followed by Tukey honest significant difference post hoc analysis.

Code Availability

The code for this manuscript will be made available upon request to the corresponding author.

Results

Development of a Methylation-based Classification Model for Brain Tumor Classification

To establish a predictive model for methylation classes using radiology images, we first had had to establish the methylation classes from a brain tumor cohort with matched MRI data (see Figure 1). Combining the TCGA LGG and GBM adult-type diffuse gliomas datasets, we identified 665 tumors with Illumina Infinium 450K methylation data. To assign class labels to the TCGA diffuse glioma dataset, we developed a custom supervised deep neural network (DNNmeth) using a previously published comprehensive methylation cohort for training (Supplementary Figure 1).[15] The DNA methylation-based classification model produced an overall accuracy of 98.56% for methylation class, and average precision of 99% (Figure 2A and 2B). The accuracy in distinguishing the IDH-glioma family from the GBM-IDHwt methylation family was 100% on hold out testing. The subclass-specific accuracies on hold out test samples for adult-type diffuse glioma classes expected to be represented in the TCGA GBM and LGG datasets was between 80% and 100% (Figure 2C).

Methylation-based Classification of TCGA Glioma Data

Next, we used our supervised DNNmeth model to assign methylation family and class labels to the TCGA adult-type gliomas cohort. To increase certainty in classification criteria, we restricted subsequent analysis to unique tumors with a supervised classification score of 0.8 or higher (619/689) (Supplementary Figure 3). The family and class assignments from the methylation-based model are provided in Supplementary Table 1. The majority of tumors fell into two methylation families, IDH-glioma (410 tumors) and GBM-IDHwt (164 tumors). The remaining samples fell into methylation classes representing a control group, a pediatric-type low- or high-grade glioma, or selected other rare entities.

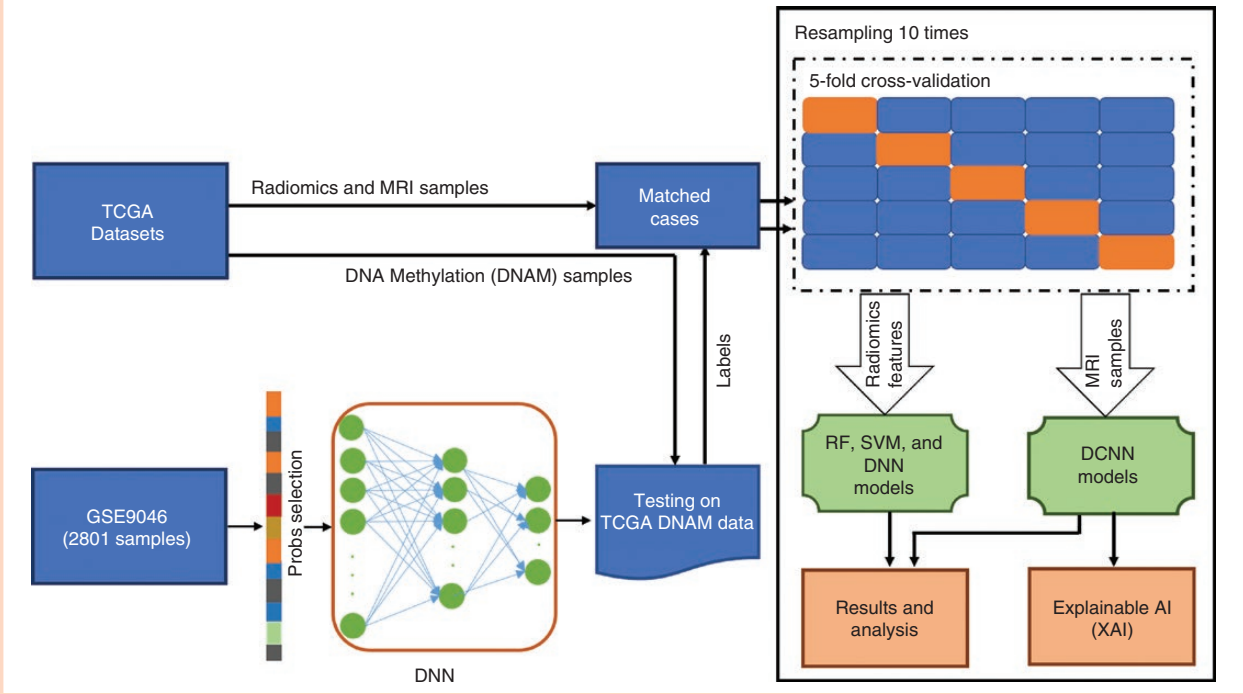

Figure 1. Schematic of modeling approach. Using a public reference cohort (GSE9046) as training data, a deep neural network (DNN) classification model was constructed to assign methylation class labels to tumors from the TCGA diffuse gliomas cohorts. Next, the radiomic features from TCGA tumors with matched MRI data were used as an input to random forest (RF), support vector machine (SVM), or DNN models to predict the methylation class labels. To test whether images could be used directly to predict methylation class, multisequence MRI images were utilized as input to an end-to-end convolutional neural networks (CNN) to predict the methylation class labels. Finally, explainable AI models were utilized in conjunction with the highest performing end-to-end classification model to determine the specific spatial locations and MRI sequences being utilized for class prediction.

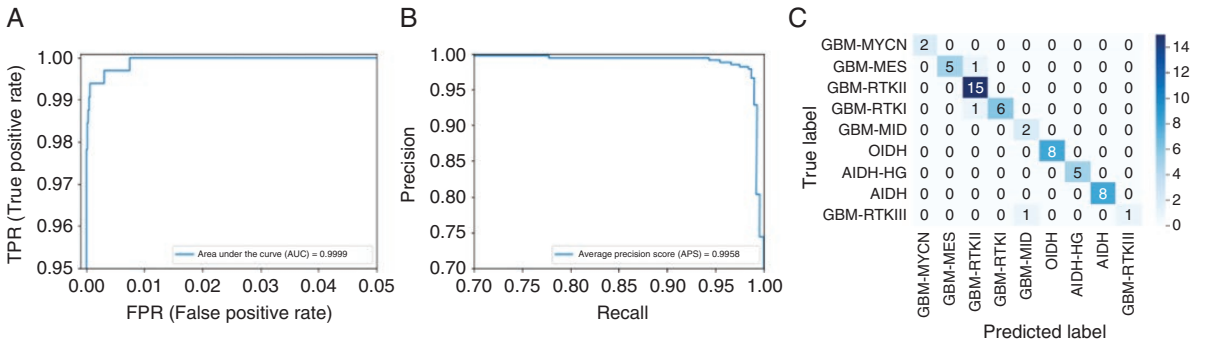

Figure 2. Construction and validation of comprehensive DNN classification model. Using a brain tumor reference dataset, a DNN was trained to predict 75 methylation family and 91 subclasses of brain or normal control tissue types. The model demonstrated good overall classification performance. The receiver operator characteristics (ROC) curve for the methylation-based classification model is presented in panel (A), the precision-recall curve in panel (B), and the confusion matrix for adult-type diffuse gliomas following cross validation in panel (C). The continuous color scale represents the proportion of tumors in each methylation class positive for specific genomic features from 0 to 1, with 0 represented by white and 1 represented by black.

Restricting the high-confidence class calls to only adult-type diffuse gliomas yielded a cohort of 574 tumors representing 8 methylation classes (Figure 3A and Supplementary Figure 3). Unsupervised projection of the of these high-confidence tumors with the reference dataset by t-SNE demonstrated that the TCGA tumors grouped with reference groups matching their respective class calls (Figure 3B). To further validate the class labels, we compared the assigned class calls to the known molecular tumor attributes. None of the tumors in the GBM-IDHwt

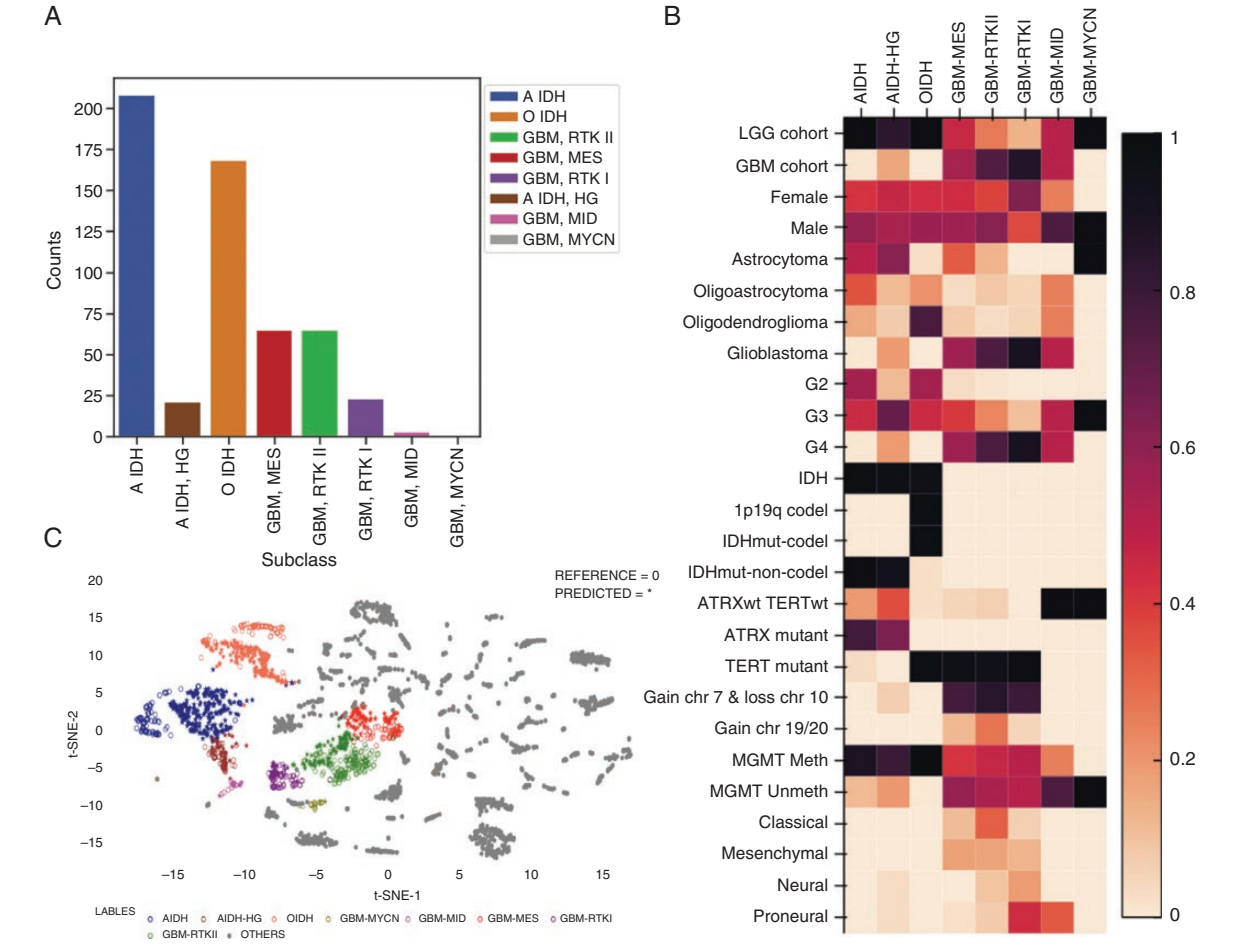

Figure 3. Labeling TCGA diffuse glioma dataset. Labeling of the adult-type diffuse glioma cohorts in the TCGA was performed by applying the DNNmeth model to 689 Illumina 450K profiles. The high-confidence diffuse gliomas consisted of two methylation families, IDH-mutant glioma and GBM-IDHwt, containing seven methylation classes (A). A heatmap representation of molecular attributes aligned with the class labels of high-confidence samples, demonstrated good concordance with the known molecular features of the tumor class assignments (B). The tSNE of the deep learned features plotted for 91 reference classes and predicted samples collected from TCGA is shown (C).

family had IDH-mutations, compared to 99.5% of tumors assigned to the IDH-glioma family (Figure 3C). One hundred percent of tumors in O-IDH methylation subclass had an IDH-mutation while 97.5% of O-IDH tumors had 1p19q deletion and *TERT* promoter mutations. In contrast, the A-IDH and A-IDH-HG groups were characterized by a high proportion of *ATRX* mutations (74.4%) (Figure 3C). Tumors assigned to GBM-IDHwt group were enriched for TERT mutations, gain of chromosome 7, and loss of chromosome 10; a pattern considered a molecular surrogate of GBM-IDHwt.[33] Our findings support the fidelity of the class assignments of our DNNmeth model on the TCGA diffuse gliomas cohorts.

Interestingly, 55 total samples in the LGG dataset exhibited a high confidence class score for an adult-type HGG group. Survival analysis demonstrated that tumors coming from the TCGA LGG cohort that classified in a GBM-IDHwt group showed a significant worse survival compared to tumors in the LGG cohort falling into an IDH-glioma subtype and overall aligned well with the other GBM-IDHwt tumors (Supplementary Figure 4). These findings suggest that traditional histopathologic classification may be imprecise when labeling for down-stream predictive models.

MRI Cohort

From our methylation cohort, we identified 78 tumors with matched MM MRI data (Supplementary Figure 3). Overall,the MRI cohort consisted of 51 IDH-gliomas and 27 GBM-IDHwt tumors. Among the IDH-mutant tumors, 13 represented tumors in the O-IDH subclass and 38 represented AIDH or AIDH-HG molecular classes. Due to the relatively restricted representation, the AIDH and AIDH-HG tumors were combined into a single AIDH class for MRI models. Among the GBM-IDHwt tumors, 13 of tumors were assigned to the GBM-RTKII and 14 were assigned to the GBM-MES groups. The patient information and demographics for the MRI cohort is presented in Table 1.

Prediction of Methylation Class from Radiomic Features

As initial evaluation of the feasibility to predict methylation class from MRI data, we used 724 previously published radiomic features derived from multiple MRI sequences to build supervised models to predict the class labels in our MRI cohort.[27] We trained three different models including a deep neural net, random forest, and support vector machine to predict methylation families (IDH-glioma vs GBM-IDHwt tumors) or subclasses (AIDH vs OIDH tumors or GBM-MES vs GBM-RTK tumors).

To generalize the proposed tumor classification methods, we conducted experiments with 10 bootstraps and fivefold validation for radiomic feature-based classification models. At the family level, the deep neural net achieved the highest accuracy (94.6%), followed by the RF (82.1%) and SVM (78.4%), despite failing to show a statistically significant difference between models. The deep neural net significantly outperformed the conventional machine learning models at the subclass levels, achieving a median accuracy to predict OIDH vs AIDH gliomas of 95.9% and for distinguishing GBM-IDHwt subclasses of 93.8% (Figure 4). To confirm the learned representations for the highest performing models, we visualized the learned feature representations from the DNN models for training and test data and demonstrated that for each classification task, the individual modeled classes can be separated by their learned feature representations and test samples project into the same feature space (Supplementary Figure 5). The finding suggests that extracted radiomic features carry sufficient features to distinguish methylation families or subclasses using machine learning.

Direct Prediction of Methylation Family and Subclass from MRI Data

Next, to test whether MRI images contained sufficient information to directly predict methylation class from MRI data, we constructed deep convolutional neural networks (DCNN) using multiple previous published architectures. Because previous studies have demonstrated good classification performance using T2WI images alone,[34–36] we trained models to predict methylation family or subclass using multimodality or T2 images alone (Figure 5).

The classification accuracies of the DCNN models for family-level tumor classification are shown in Figure 5A. The VGG-Net and denseNet with MM showed the highest accuracy, achieving an accuracy of 80.59% ± 2.13 and 80.73% ± 2.08, respectively. For each model, the MM models showed higher accuracy but no statistically significant difference to their respective counterpart models utilizing T2 sequences alone.

In the IDH-glioma subtype classification tasks, the VGG-Net model showed the best performance compared to ResNet and denseNet, achieving mean accuracies of 87.21% ± 1.59 and 87.60% ± 2.64, respectively (Figure 5B). For this classification task, an appreciable difference between MM and T2 models was not detected for any of the model architectures. For subclassification of GBM, IDH-wildtype tumors, the VGG-Net model again showed superior performance achieving a mean accuracy of 89.00% ± 4.20 for MM models and 83.4% ± 6.30 for the T2 alone model (Figure 5C).

Overall, the VGG-Net demonstrated the best overall stability and performance compared to the ResNet and DenseNet models for multimodality and T2-modality based learning. Multimodality models were characterized by overall higher accuracy scores and lower variability compared to T2-alone models; however, multimodality models were not statistically better than their T2-alone counterparts.

Interpretability of MRI Models

One potential limitation of using deep learning models for biomedical applications is the lack of human interpretability and quality control of the output. To help aid

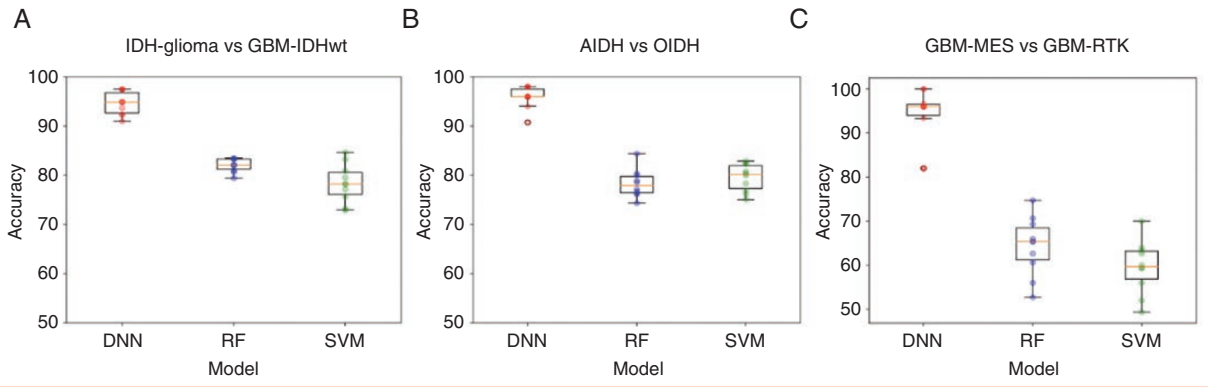

Figure 4. Accuracy of methylation class label prediction from MRI radiomic features. Seven hundred and twenty-four radiomic features from TCGA diffuse glioma MRI data were used to predict the IDH-mutant or GBM-IDHwt methylation families (A), to distinguish the OIDH from the AIDH/AIDH-HG methylation subclasses (B), or the GBM-RTK from GBM-MES methylation subclasses (C). For both family and subclass prediction a DNN, a random forest (RF), or a support vector machine (SVM) model was evaluated. The accuracy of each model is presented as ten bootstraps of the average accuracy under fivefold cross validation. A statistically significant difference was detected between subclass level models, with the DNN model showing significantly higher average accuracy compared to the RF and SVM models.

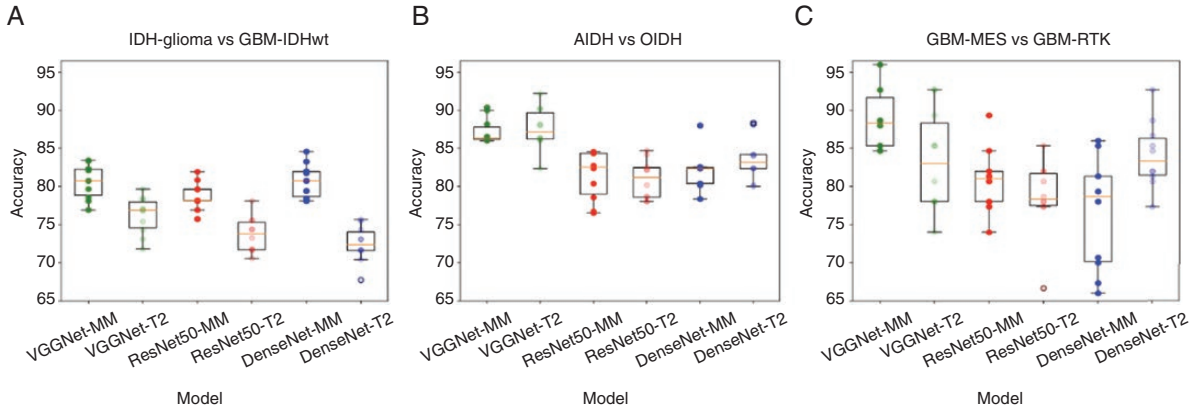

Figure 5. Accuracy of direct prediction of methylation class label prediction from multimodality MRI data. MM or T2WI MRI sequences from TCGA diffuse glioma MRI data were used to predict the IDH-mutant or GBM-IDHwt methylation families (A), to distinguish the OIDH from the AIDH/AIDH-HG methylation subclasses (B), or the GBM-RTK from GBM-MES methylation subclasses (C). For both family and subclass prediction a Deep Convolutional Neural Networks models including VGGNet, ResNet50, and DenseNet were evaluated. The accuracy of each model is presented as ten bootstraps of the average accuracy under fivefold cross validation. While a significant difference was not detected for family level classification tasks, with the VGGNet model showing overall higher average accuracy compared to the ResNet50 and DenseNet models for subclass level classification tasks. A significant difference between MM or T2 based models was not appreciated within individual models.

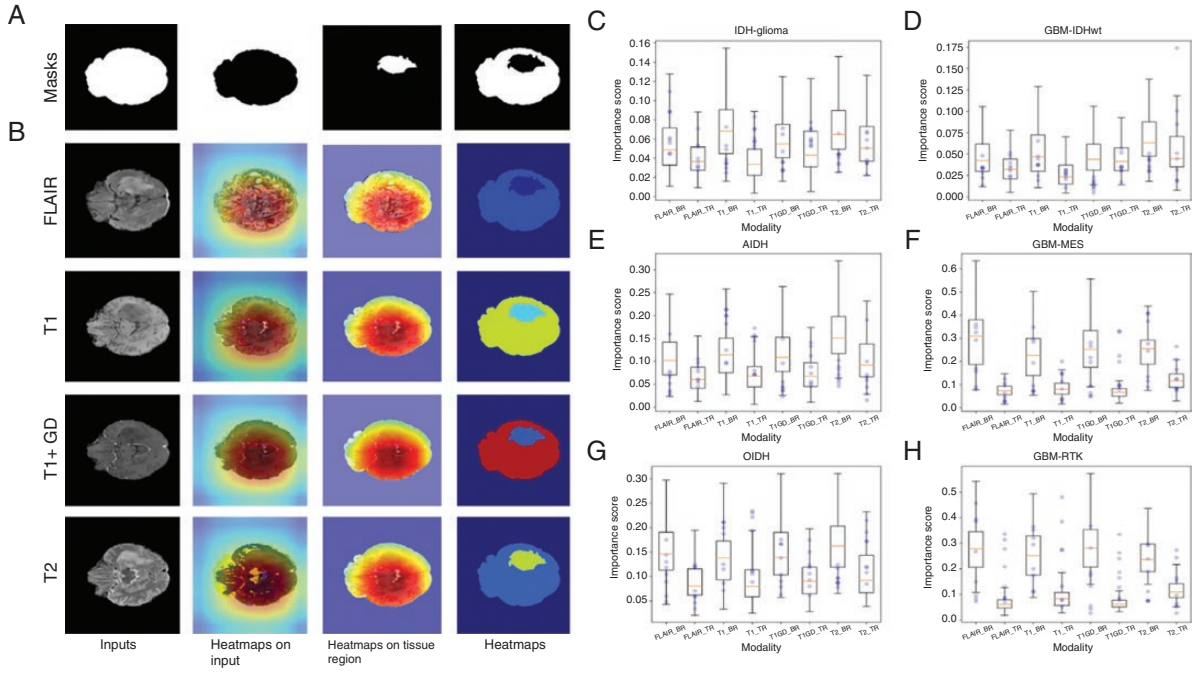

Figure 6. Explainable outputs derived from DCNN feature representation from multimodality MRI data. The brain mask, nontissue mask, predicted tumor region mask, and nontumor brain region are shown from left to right in panel (A), respectively. The MRI sequence input, the heatmap on input MRI sequences, heatmap on tissue region, and the final outputs for four different modalities are demonstrated in panel (B) from left to right, respectively. Cumulative importance scores were generated for each molecular tumor type and the box plots for family and subclass-level importance are shown for MM or T2 weighted models in the panels from (C) to (H). TR, Tumor region; BR, Brain region.

human-level interpretation to our classification models, we constructed an explainable system for the VGG-net based models to align importance scores to image region and modality. We found importance was focused on the tissue regions even in the absence of segmentation. The models used a combination of tumor and nontumor regions for each classification task, with focus split between tumor and the deep grey structures. The latter suggests the use of location specific cues in addition to tumor intrinsic signals in class predictions (Figure 6). We observed variability

with respect to the MRI sequences most important for classification between the individual tasks, but the FLAIR and T2 weighted images were consistently of high importance. Contrast-enhanced sequences within the tumor region showed the greatest importance in differentiating IDH-glioma which largely represent lower grade gliomas from GBM-IDHwt tumors (Figure 6).

Because the T2-FLAIR mismatch sign has previously been reported to distinguish AIDH from OIDH tumors,[37] we evaluated whether our models may also be relying on this feature. We found the T2-FLAIR mismatch sign in 43% of sampled slice images of AIDH tumors compared to 0% of OIDH tumors. There was no statistically significant difference between be correctly called AIDH and having the T2-FLAIR mismatch sign, and selected examples were miscalled OIDH even the presence of the sign (Supplementary Figure 6). While the findings do not exclude the T2-FLAIR mismatch as a contributing factor, it cannot be the primary determinant of class call in the IDH-glioma subtype model.

Discussion

In this study, we show that methylation class can be predicted either indirectly from extracted radiomic features or directly from MM MRI images across histomorphologic tumor types using tumor methylation class labels from TCGA diffuse glioma cohorts assigned using a custom DNN. In addition, using explainable AI methods, we demonstrate that our direct prediction models utilize tumor intrinsic and nontumor intrinsic tissue regions and show variability in specific MRI sequence importance, providing important quality control and adding additional insight to the classification output.

Our results validate methylation class as a viable target for MRI-based prediction models, extending the list of genomic targets that have been successfully predicted from MRI data. Initial attempts at radiomic modeling were geared toward histomorphologic diagnosis or tumor grade.[27] Subsequent efforts successfully implemented models to predict genomic features in diffuse gliomas such as IDH-mutation status, 1p19q co-deletion,[8,38,39] or cytogenetic abnormalities.[9,40]Using methylation class for radiogenomic prediction offers advantages over histomorphology-defined tumor types and individual genomic features such as SNV or copy number abnormalities. Histopathologic tumor types are known to contain molecular heterogeneity which may add noise to predictive models. Similarly, tumor grade may be unreliable due to sampling bias. This was highlighted by our observation that high-confidence methylation class assignments for adult-type high-grade glioma classes in the GBM-IDHwt family were identified among cases in the TCGA LGG cohort. A tumor's genome-wide DNA methylation signature is objective, and the core signature is maintained between diagnostic, recurrent, and metastatic samples.[41–43] Because methylation class can be used as a predictive target even in tumor types that lack highly recurrent SNVs or copy number, the approach substantially expands the number and types of tumors that can be modeled using these methods.

Using a custom dataset, Kickingereder *et al.* previously used MRI radiomic features to predict three methylation classes of glioblastoma.[20] Subsequently, the same group used methylation class combined with radiomic features to demonstrate that radiomic features were an independent predictor of risk in their cohort.[21] Our findings, relying on the widely used and comprehensive TCGA diffuse glioma cohort, validates the previous use of radiomic features to predict methylation classes and expands the types of diffuse adult-type gliomas to include lower grade tumor classes. While we observed a slight improvement in performance metrics compared to Kickingereder *et al.*, this can probably be explained by our use of contemporary activation functions and model architectures and the use of a DNN compared to classical machine learning methods. In addition to previous studies, our findings also represent the first direct prediction of methylation class from multisequence MRI data without the pre-extraction of radiomic features.

One criticism of using machine learning for biomedical applications is that the models often lack interpretability. Our explainable models provide important quality control and add insight to our implementation. Remarkably, even though we did not perform specific tumor segmentation when training our direct classification modes, we demonstrate that our predictive models rely on tumor intrinsic elements for the specific classification task. In each classification task, the T2 weighted sequence showed overall high importance across all tumor regions. In contrast, the importance of contrast enhanced sequences showed the overall lowest contribution and seemed to be only slightly more important in the non-tumor regions. The high importance of T2WI in classification tasks may explain why models using T2 alone are nearly equivalent to multimodality models. Of note, our explainable models focused attention on both tumor and nontumor brain regions suggesting that location within the brain may be an important feature related to methylation class. This would be supported by previous studies that have demonstrated distinct locations within the brain associated with molecular tumor subtypes.[44] Because the focus is in deep grey structures, suggesting general location cues, it isn't clear to what extent location specific signal could confuse the models.

There are significant limitations to our current models and approach which would need to be addressed prior to clinical implementation. While the accuracy was relatively high for individual tasks, the cohort was restricted to adult-type diffuse gliomas. The clinical utility would be enhanced if the models were expanded to be more comprehensive, and representative of the full spectrum of tumor types encountered. Several limitations were related to the restricted training examples available with matched MRI and methylation data. Our implementation relied on multiple binary classification tasks, whereas single comprehensive models could be constructed with sufficient training data. Although the current model performance is slightly better if extracted radiomic features are use as the input, we have demonstrated feasibility for direct image classification. The latter would prevent significant preprocessing steps when constructing input data for training.

Overall, our results demonstrate the feasibility of indirect or direct prediction of methylation class from MRI data which

theoretically expands the tumor types which could be modeled using radiogenomics to nearly all known brain tumor types. Future studies should expand the number and types of tumors used to match methylation and MRI data which should enhance model robustness and improve accuracy.

Supplementary material

Supplementary material is available online at *Neuro-Oncology Advances* online.

Keywords

glioma | radiogenomics | MRI | DNA methylation profiling | brain tumor classification

Acknowledgments

This work was partially supported by the National Cancer Institute support Grant (P30 CA021765) and the American Lebanese Syrian Associated Charities (ALSAC).

Conflict of interest statement

The authors have no conflict of interest related to this work.

Funding

This work was partially supported by the National Cancer Institute support Grant (P30 CA021765) and the American Lebanese Syrian Associated Charities (ALSAC).

Authorship Statement

Conceptualization: ZA, BAO. Methodology: ZA, QTT, BAO. Clinical correlation and interpretation: BAO, AKB, JTL. Manuscript writing and editing: ZA, QTT, AKB, JTL, BAO. Statistical analysis: QTT. Supervision of the overall study: BAO.

References

1. Weller M, Wick W, Aldape K, et al. Glioma. *Nat Rev Dis Primers*.2015;1(1):15017.
2. Ceccarelli M, Barthel FP, Malta TM, et al; TCGA Research Network. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell*.2016;164(3):550–563.
3. 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*. 2016;131(6):803–820.
4. Hartmann C, Hentschel B, Wick W, et al. Patients with IDH1 wild type anaplastic astrocytomas exhibit worse prognosis than IDH1-mutated glioblastomas, and IDH1 mutation status accounts for the unfavorable prognostic effect of higher age: implications for classification of gliomas. *Acta Neuropathol*.2010;120(6):707–718.
5. Brat DJ, Verhaak RG, Aldape KD, et al. Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *N Engl J Med*.2015;372(26):2481–2498.
6. Eckel-Passow JE, Lachance DH, Molinaro AM, et al. Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors. *N Engl J Med*.2015;372(26):2499–2508.
7. Itakura H, Achrol AS, Mitchell LA, et al. Magnetic resonance image features identify glioblastoma phenotypic subtypes with distinct molecular pathway activities. *Sci Transl Med*.2015;7(303):303ra–30138.
8. Hsieh KL, Chen CY, Lo CM. Radiomic model for predicting mutations in the isocitrate dehydrogenase gene in glioblastomas. *Oncotarget*.2017;8(28):45888–45897.
9. Zhou H, Vallières M, Bai HX, et al. MRI features predict survival and molecular markers in diffuse lower-grade gliomas. *Neuro Oncol*.2017;19(6):862–870.
10. Jain R, Johnson DR, Patel SH, et al. "Real world" use of a highly reliable imaging sign: "T2-FLAIR mismatch" for identification of IDH mutant astrocytomas. *Neuro Oncol*.2020;22(7):936–943.
11. Hoadley KA, Yau C, Hinoue T, et al. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell*.2018;173(2):291–304 e6.
12. Hovestadt V, Remke M, Kool M, et al. Robust molecular subgrouping and copy-number profiling of medulloblastoma from small amounts of archival tumour material using high-density DNA methylation arrays. *Acta Neuropathol*.2013;125(6):913–916.
13. Sturm D, Orr BA, Toprak UH, et al. New Brain tumor entities emerge from molecular classification of CNS-PNETs. *Cell*.2016;164(5):1060–1072.
14. Sturm D, Witt H, Hovestadt V, et al. Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell*.2012;22(4):425–437.
15. Capper D, Jones DTW, Sill M, et al. DNA methylation-based classification of central nervous system tumours. *Nature*.2018;555(7697):469–474.
16. Capper D, Stichel D, Sahm F, et al. Practical implementation of DNA methylation and copy-number-based CNS tumor diagnostics: the Heidelberg experience. *Acta Neuropathol*.2018;136(2):181–210.
17. Koelsche C, Schrimpf D, Stichel D, et al. Sarcoma classification by DNA methylation profiling. *Nat Commun*.2021;12(1):498.
18. Santana-Santos L, Kam KL, Dittmann D, et al. Validation of whole genome methylation profiling classifier for central nervous system tumors. *J Mol Diagn*.2022;24(8):924–934.
19. Kumar R, Liu APY, Orr BA, Northcott PA, Robinson GW. Advances in the classification of pediatric brain tumors through DNA methylation profiling: from research tool to frontline diagnostic. *Cancer*.2018;124(21):4168–4180.
20. Kickingereder P, Bonekamp D, Nowosielski M, et al. Radiogenomics of glioblastoma: machine learning-based classification of molecular characteristics by using multiparametric and multiregional mr imaging features. *Radiology*.2016;281(3):907–918.
21. Kickingereder P, Neuberger U, Bonekamp D, et al. Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma. *Neuro Oncol*.2018;20(6):848–857.
22. Liu J, Lichtenberg T, Hoadley KA, et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell*.2018;173(2):400–416 e11.

23. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*.2014;30(10):1363–1369.
24. Triche TJ, Jr, Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res*.2013;41(7):e90.
25. Pedregosa F, Varoquaux G, Gramfort G, et al. Scikit-learn: machine learning in python. *J Mach Learn Res*.2011;12(25):2825–2830.
26. Abadi M, Barham P, Chen J, et al. *Tensorflow: a system for large-scale machine learning. in Osdi*. 2016. Savannah, GA, USA.
27. Bakas S, Akbari H, Sotiras A, et al. Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci Data*.2017;4(1):170117.
28. 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.
29. Reinhold JC, Dewey BE, Carass A, et al. Evaluating the impact of intensity normalization on mr image synthesis. *Proc SPIE Int Soc Opt Eng*.2019;10949:890–898.
30. He K, Ren S, Sun J, et al. Deep residual learning for image recognition, in Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. 2016, IEEE. p. 770–778.
31. Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
32. Huang Y, Wang JY, Wei XM, et al. Bioinfo-Kit: A sharing software tool for Bioinformatics. In *Applied Mechanics and Materials*. 2014. Trans Tech Publ.
33. Brat DJ, Aldape K, Colman H, et al. cIMPACT-NOW update 3: recommended diagnostic criteria for "Diffuse astrocytic glioma, IDH-wildtype, with molecular features of glioblastoma, WHO grade IV.". *Acta Neuropathol*.2018;136(5):805–810.
34. Radbruch A, Lutz K, Wiestler B, et al. Relevance of T2 signal changes in the assessment of progression of glioblastoma according to the Response Assessment in Neurooncology criteria. *Neuro Oncol*2012;14(2):222–229.
35. Kern M, Auer TA, Picht T, Misch M, Wiener E. T2 mapping of molecular subtypes of WHO grade II/III gliomas. *BMC Neurol*.2020;20(1):8.
36. Gu W, Fang S, Hou X, Ma D, Li S. Exploring diagnostic performance of T2 mapping in diffuse glioma grading. *Quant Imaging Med Surg*2021;11(7):2943–2954.
37. Patel SH, Poisson LM, Brat DJ, 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(20):6078–6085.
38. Chang K, Bai HX, Zhou H, et al. Residual convolutional neural network for the determination of IDH status in low- and high-grade gliomas from MR imaging. *Clin Cancer Res*.2018;24(5):1073–1081.
39. Pasquini L, Napolitano A, Tagliente E, et al. Deep learning can differentiate IDH-mutant from IDH-Wild GBM. *J Pers Med*2021;11(4):290–301.
40. 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.
41. Bormann F, Rodríguez-Paredes M, Lasitschka F, et al. Cell-of-origin DNA methylation signatures are maintained during colorectal carcinogenesis. *Cell Rep*2018;23(11):3407–3418.
42. Gull N, Jones MR, Peng P-C, et al. DNA methylation and transcriptomic features are preserved throughout disease recurrence and chemoresistance in high grade serous ovarian cancers. *J Exp Clin Cancer Res*.2022;41(1):232.
43. Kumar R, Smith KS, Deng M, et al. Clinical outcomes and patient-matched molecular composition of relapsed medulloblastoma. *J Clin Oncol*.2021;39(7):807–821.
44. Patay Z, DeSain LA, Hwang SN, et al. MR imaging characteristics of wingless-type-subgroup pediatric medulloblastoma. *AJNR Am J Neuroradiol*.2015;36(12):2386–2393.