Cell & Developmental Biology
ISSN: 2168-9296
OPEN ACCESS Freely available online
Research Article

The Use of Molecular Diagnostics and Artificial Intelligence in the Classification and Prognosis of Gliomas: A Systematic Review and Meta-analysis

Costas Fantis*

Department of Neurology, University of Buckingham, Buckingham MK18 1EG, United Kingdom

Supplementary Table 1: AI applied specifically to MRI

Paper	Biomarkers investigated	AI used	Demographics/sample size/dataset	Accuracy/effectiveness
Riahi Samani et al. 2023	Peritumoral region differences with overall survival and *IDH1* mutation status	CNN:Convolutional Neural Networks, FERNET: Freewater estimatoR using Interpolated Initialization	381 patients with adult-type diffuse gliomas (CNS WHO grade 4) and 50 patients with brain metastases, and they were randomly divided into three datasets.	Results showed significant differences in the proposed markers between patients with different overall survival and *IDH1* mutation status (t test, Wilcoxon rank sum test, linear regression; p<0.01).Clustering of patients using the proposed markers reveals distinct survival groups (log rank; p<10-5, Cox hazard ratio=1.82; p<0.005).
Do et al. 2022	MGMT methylation	Random Forest (RF), XG Boost, and Support Vector Machine (SVM),	The Cancer Imaging Archive (TCIA), The Cancer Genome Atlas (TCGA) Glioblastoma multiforme (GBM) dataset, which included 154 patients with annotated MGMT status	A sensitivity of 0.894, specificity of 0.966, and accuracy of 0.925 for predicting the MGMT methylation status in GBM.
Xu et al. 2022	BRAF V600E mutation in pediatric low-grade gliomas	Machine learning algorithms	113 cases, with 79 patients in the training group (47 men, 32 women; mean age 9.86 ± 5.20) and 34 patients in the testing group (20 men, 14 women; mean age 10.97 ± 5.14)	AUC of 0.91, accuracy of 0.93, sensitivity of 0.83, and specificity of 0.97.
H. Luo et al. 2021	*IDH1/2*, 1p19q status	Radiomic Signature based radiomics model	655 glioma patients	Average accuracies of histological diagnosis and molecular subtyping were 89.8 and 86.1% in the cross-validation cohort, respectively, while these numbers reached 83.9 and 80.4% in the independent testing cohort
Y.-Z. Sun et al. 2021	Differentiating pseudoprogression from true progression in GBM patients after standard treatment	'E1', 'caret', and 'unbalanced' R packages	77 patients with glioblastoma multiforme (GBM): 40 men and 37 women with a mean age of 49.1 ± 10.5 years (range 17–76 years).	The radiomic classifier demonstrated ACC, sensitivity, and specificity of 72.78% 78.36% and 61.33%. The accuracy, sensitivity and specificity of three radiologists' assessments were lower on all criteria, therefore TCE-based radiomics showed better classification performance

Correspondence to: Costas Fantis, Department of Neurology, University of Buckingham, Buckingham MK18 1EG, United Kingdom, Tel: 078132929[illegible]; E-mail: cosfantis@encodehealth.com

Received: [illegible] Manuscript No. CDB-24-31036; **Editor assigned:** 06-May-2024, Pre QC No. CDB-24-31036 (PQ); **Reviewed:** 20-May-2024, QC No. CDB-24-31036; **Revised:** 27-May-2024, Manuscript No. CDB-24-31036 (R); **Published:** 03-Jun-2024, DOI: 10.35841/2168-9296.24.13.339.

Citation: Fantis C (2024) The Use of Molecular Diagnostics and Artificial Intelligence in the Classification and Prognosis of Gliomas: A Systematic Review and Meta-analysis. Cell Dev Biol. 13:339.

Yu Han et al. 2020	*IDH1* mutation	Support Vector Machine with Recursive Feature Elimination (SVM-RFE)	59 grade II/III glioma patients with known *IDH1* mutation status	SVM model was established [illegible]g 19 features selected with SVM-RFE for radiomic analysis. The AUC and accuracy for mutation on [illegible]ain-ing set were 0.892 and 0.952, [illegible] testing set were [illegible] 0.84, respectively
Lo et al. 2020	WT and mutant *IDH*	Logistic regression	The data set used in the experiment was from The Cancer Imaging Archive (TCIA), established by the National Cancer Institute. The study collected an image database composed of 32 WT and 7 mutant *IDH* cases.	His system achieved an accu[illegible]y of 90%, a sensitivity of 57% and a specific[illegible] of 97%.
Z. Zhang et al. 2020	LGGs vs HGGs	SVM models (using an iterative information gain algorithm under the leave-one-out cross-validation strategy and used the AUC the optimization criteria. VGG (Visual Geometry Group) is a Convolutional Neural Network (CNN) architecture developed by the Visual Geometry Group at the University of Oxford	108 patients, of which 43 were LGGs and 65 were HGGs (25 WHO III and 40 WHO IV patients)	AUC=0.9[illegible]cy= 0.[illegible], sensitivity=0.98, [illegible] specificity=0.[illegible] in classifying LGGs from H[illegible] AUC=0.99, accuracy=0.98, sensitivity=0.[illegible] nd specificity=1.00 in classifying [illegible]de III from IV.
Zhao et al. 2020	WHO grade II from III oligodendrogliomas	Conditional inference random forest classifier	36 pati[illegible] with T1CE were included (19[illegible] 17 women, mean age=45 years; age[illegible] 5 years) and classified into two groups: ODG2 (n=19; mean age=46 years, age [illegible] years) and ODG3 (n=17, mean a[illegible]=44 years, age range=9-65 years).	The AUC, sensitivity, specificity and accuracy of radiomics were 0.798 (95%CI 0.699–0.896), 0.672, 0.789, 0.735 for T1 CE, 0.774 (95%CI 0.671–0.877), 0.700, 0.683, 0.689 for FLAIR, 0.861 (95%CI 0.783–0.940), 0.778, 0.783, 0.781 for their combination, respectively. The radiomic strategy performed superior to that of radiologists.
Shboul, Chen, and M Iftekharuddin 2020	MGMT methylation, *IDH* mutations, 1p/19q co-deletion, ATRX mutation, and TERT mutations	LGG-XGBoost Extreme Gradi[illegible]ting [illegible]GBoo[illegible]	10[illegible]pre-operative LGG patients	The prediction models achieve a test performance AUC of 0.70 to 0.84
X. Sun et al. 2020	Ki-67 and p5[illegible]	LASSO regressio[illegible] model	110 patients with glioma.	The best Area Under the Curve (AUC) for the Ki-67 model was 0.773 for T2 weighted imaging in solid glioma (sensitivity, 0.818; specificity, 0.833), followed by a less reliable AUC of 0.773 (sensitivity, 0.727; specificity 0.667) in 20 mm peritumoral areas. The highest AUC for the p53 model was 0.709 (sensitivity, 1; specificity, 0.4) for T2 weighted imaging in 10-mm peritumoral areas.
Mzoughi et al. 2020	L[illegible]r HGG	LASSO, SVM	Preoperative MR images from 272 patients with primary WHO grade II/III gliomas, including 179 with grade II and 93 with grade III. The proportion of patients with a p53 mutation was 42.2% and 48.9% in the training and validation sets, respectively.	An overall accuracy of 96.49% using the validation dataset

Tian et al. 2020	Telomerase Reverse Transcriptase (TERT) promoter mutations	LASSO regression and univariate and multivariate logistic regression analyse	126 patients who suffered from HGG were divided into the training cohort (88, 70%) and the validation cohort (38, 30%). There was no significant difference in terms of age (p=0.986), gender (p=0.975), grade (p=0.327), or location (p=0.421) between the training and validation cohorts	To verify the discovery, 4 models were established, and there was only one [illegible]e factor (CNV) in Model A than Model B. The AUC of Model A in both training cohort and validation cohort were high[illegible] in the ROC analysis (0.[illegible]5 *vs.* 0.917, 0.889 vs 0.868, respectively).
Gates et al. 2020	Predict the World Health Organization grade of gliomas	Random forest method	23 patients with gliomas who were enrolled in a prospective clinical imaging trial between 2013 and 2016. The patients had a range of glioma grades (7 grade II, 9 grade III, and 7 grade IV).	Tumor grade was predict[illegible]t 96% accuracy [k=0.9[illegible]] using 4 inputs [T2, DC, CBV, and transfer constant from dynamic contrast-enhanced imaging].
Wu et al. 2019	*IDH* genotype	RF, r-SVM, NN, l-SVM, kNN, NB, FDA, Adaboost	A total of 126 patients were enrolled for analysis. Overall[illegible]4 radiomic features extracted from the pre-operative MRI[illegible]ges were analyzed. The patient[illegible]e randomly assigned to either the training set or the validation set at a ratio of 2:1.	Random Forest (RF) showed high predictive performance (accuracy 0.885 ± 0.041, AUC 0.931 ± 0.0[illegible]ural Network (NN) (accuracy 0.829 ± 0.06[illegible] AUC 0.878 ± 0.052) Flexible discriminant analysis (FDA) (accuracy 0.851 ± 0.049, AUC 0.875 ± 0.057) displayed low predictive performance. With regard to stability, RF also showed high robustness against data perturbation (relative standard deviations, RSD 3.87%).
H. Zhou et al. 2019	*IDH* and 1p19q genotype in grade II-IV gliomas	Random forest classification	Preoperative MRIs of 538 glioma patients from three institutions were used as training cohort. The model was validated using MRIs from glioma patients in cancer imaging archive.	The model predictive of *IDH* achieved an area under the receiver operating characteristic curve (AUC) of 0.921 in the training cohort and 0.919 in the validation cohort. Based on the top 15 features, the AUC was 0.917 and 0.916 for the training and validation cohort, respectively. The overall accuracy for 3 group prediction (*IDH*-wild type, *IDH*-mutant and 1p19q co-deletion, *IDH*-mutant and 1p19q non-codeletion) was 78.2% (155 correctly predicted out of 198).
Batchala et al. 2019	1p/19q-codeletion status among *IDH*-mutant lower grade gliomas	Multivari[illegible]stic Regression (MLR) analyse	102 patients with *IDH*-mutant LGGs in the training data- set, 51% were women (n=52) and 49% were men (n=50). The median age was 41 years (range, 20.0–75.0 years; interquartile range, 33.053.0 years). Of the 102 LGGs, 62.7% (n=64) were *IDH*mut-Noncodel and 37.3% (n=38) were *IDH*mut-Codel, 57.8% (n=59) were WHO grade II, and 42.2% (n=43) were WHO grade III.	Training dataset analysis produced a 2-step classification algorithm with 86.3% codeletion prediction accuracy, based on the following: 1) the presence of the T2-FLAIR mismatch sign, which was 100% predictive of noncodeleted lower grade gliomas, (=21); and 2) a logistic regression model based on texture, patient age, T2* susceptibility, primary lobe, and hydrocephalus. Independent validation of the classification algorithm rendered codeletion prediction accuracies of 81.1% and 79.2% in 2 independent readers. The metrics used in the algorithm were associated with moderate substantial interreader agreement (k=0.56-0.79).
Fan et al. 2019	Differentiation [illegible]n Glioblastoma (GBM) and Anaplastic Oligodendroglioma (AO).	Distance correlation, LASSO, and GBDT and two classifiers LDA and SVM (LDA-based and SVM-based models)	126 patients were enrolled in the current study, 76 patients were diagnosed with GBM, and 50 patients with AO.	The diagnostic performance of machine learning models was superior to human readers. Both classifiers showed promising ability in discrimination with AUC more than 0.900 when combined with suitable feature selection method. For LDA-based models, the AUC of models were 0.986, 0.994, and 0.970 in the testing group, respectively. For the SVM-based models, the AUC of models were 0.923, 0.817, and 0.500 in the testing group, respectively. The over-fitting model was GBDT+SVM, suggesting that this model was too volatile that unsuitable for classification.

Sasaki et al. 2019	MGMT methylation status	Least Absolute Shrinkage and Selection Operator (LASSO) and Supervised Principal Component Analysis (SPCA)	Preoperative MRI scans from 201 newly diagnosed GBM patients were included in this study. A total of 489 texture features including the first-order feature, second-order features from 162 datasets, and location data from 182 datasets were collected.	Predictive accuracy of the pMGMT methylation status was 67% when modeled by two significant radiomic features. A significant survival difference was observed among the combined high-risk group and low-risk group (this group consists of radiomic low risk and pMGMT-unmet or radiomic high risk and pMGMT-met), and combined low-risk group (. Log-rank test)
C. Jiang et al. 2019	MGMT promoter methylation status in LGG patients	3D-CE-T1-weighte d single radiomics model, T2-weighted single radiomics model, fusion radiomics model, linear combination radiomics model, clinical integrated model	122 pathology confirmed LGG patients were retrospectively reviewed, with 87 local patients as the training dataset, and 35 from The Cancer Imaging Archive as independent validation. A total of 1702 radiomics features were extracted from three-dimensional contrast-enhanced T1 (3D-CE-T1) weighted and T2-weighted MRI images.	The fusion radiomics model which constructed from the concatenation of both series, displayed the best performance, with an accuracy of 0.849 and an area under the curve (AUC) of 0.970 (0.9 .000) in the training dataset, and an accuracy and an AUC of 0.898 (0.786-1.000) in the validation dataset.
Lee et al. 2019	*IDH1* mutation status in glioblastoma	K-nearest neighbors, support vector classification, decision tree, random forest, adaboost, naive bayes, linear discriminant analysis, gradient boosting.	Retrospectively identified patients with newly diagnosed GB After semiautomatic segmentation of the lesions, 21 features were extracted from preoperative multiparametric magnetic resonance images. training cohort (n=88) was used train machine learning-based classifiers, with internal validation machine-learning technique was validated in an external dataset of patients with GBM.	Accuracies of 87.3% in the training cohort were achieved and 81.7% in the validation cohort using the K-nearest neighbors model. Accuracies of 86.2% in the training cohort and 82.9% in the validation cohort using the AdaBoost model.
Y. Li et al. 2019	Phosphatase and Tensin homolog (PTEN) mutation status	Support Vector Machine (SVM) classifier	A total of 109 patients were included in the study, 69 patients in the training set and 40 patients in the validation set. The training set was collected from The Cancer Genome Atlas (TCGA) database, and the validation set was retrospectively collected from the Chinese Glioma Genome Atlas. Six features were selected using the mRMR algorithm, including two features derived from contrast-enhanced images and four features derived from T2-weighted images.	The predictive performance of the machine learning model for the training and validation sets were 0.925 and 0.787, respectively
L. Han and Kamdar. 2018	MGMT methylation status	Bidirectional Convolutional Recurrent Neural Network architecture (CRNN)	The study used a compendium of brain MRI scans of GBM patients collected from The Cancer Imaging Archive (TCIA) combined with methylation data from The Cancer Genome Atlas (TCGA). The training dataset consisted of 344 positive MRI scans and 351 negative scans, which corresponded to 117 patients. The validation dataset consisted of 21 patients, with 73 positive scans and 62 negative scans. The test set also had 21 patients, with 62 positive and 62 negative scans. After data augmentation, this resulted in 62,550 examples in the training set, and 12,150 in the validation set, and 11,160 in the test set.	An accuracy of 67% on the validation data and 62% on the test data, with precision and recall both at 67%.

Liang et al. 2018	*IDH* genotype	121, 161, 169, and 201 layers M3D-DenseN and Grouped-SNet and Grouped-MNet	To evaluate the performance, BRATS-2017 and The Cancer Genome Atlas breast invasive carcinoma (TCGA-BRCA) dataset to get image data as input and gene mutation information as the target. Transfer learning techniques were applied to predict World Health Organization (WHO) grade status.	The best performance was achieved with accuracy 84.6%, sensitivity 78.5%, specificity 88.0%, and AUC 85.7% on the training dataset when M3D-DenseNet depth was set to layers, and with an accuracy of 84.6%, sensitivity 78.5%, specificity 88.0% and AUC 85.7% on the validation dataset when M3D-DenseNet depth was set to layers. The experimental result shown that multi-modality model was better than single-modality model, as the best accuracy of single-modality model on the validation dataset was 74.4% (AUC=81.6%), nevertheless accuracy of multi-modality model on the validation dataset was 84.6% (AUC=85.7%). Prediction of World Health Organization (WHO) grade status achieved a high accuracy of 91.4% (AUC=94.8%) on validation dataset.
Yang et al. 2018	*IDH* status (mutated *vs.* wildtype)	AlexNet and GoogLeNet (Both CNN architectures)	113 histologically confirmed glioma patients were retrospectively enrolled. Each participant underwent preoperative conventional and advanced MRI scans on a 3.0T MRI scanner (Discovery 750, GE Healthcare, Milwaukee, WI, United States of America) with an 8-channel head coil. The study group comprised 52 patients (grade II: 25, grade III: 27) with LGG and 61 patients with HGG. Number of raw images with tumor in LGG and HGG group was 568 and 499, respectively. The ages of the LGG and HGG cohort range from 10 to 66 years old and 13 to 87 years old, respectively. Tumor location of the 113 glioma patients was summarized according to Vasari MRI Visual Feature Guide.	The mean value of validation accuracy, test accuracy and test AUC of GoogLeNet was 0.867, 0.907, and 0.939, respectively. As for AlexNet, the mean value of validation accuracy, test accuracy and test AUC were 0.866, 0.855, and 0.895, respectively. Although GoogLeNet is deeper than AlexNet, it has less parameters owing to the inception structures, which can reduce the risk of overfitting.Although only T1+C images were used, the test accuracy and AUC reached 0.945 and 0.968 in GoogLeNet
Y. Li, Liu, et al. 2018	ATRX mutation status in lower-grade gliomas	Support Vector Machine (SVM) and LASSO regression model	Clinical and radiological features of 95 TCGA patients and 91 CGGA patients were obtained. ATRX mutation was detected in 38.8 % (37/95) of patients in TCGA cohort and 30.8 % (28/91) in the CGGA cohort. Patients from TCGA were divided into training (n=63) and validation (n=32) sets, and no significant difference was observed in age (p=0.506), sex (p=0.337), grade (p=0.071), ATRX status (p=0.837), *IDH1* status (p=0.588), 1p19q status (p=0.526), histology (p=0.138) and tumour location (p=0.516) between the two datasets.	The predictive efficiencies measured by the area under the curve were 94.0%, 92.5% and 72.5% in the training, validation and external validation sets, respectively.

Jakola et al. 2018	*IDH* prediction	Binary classifier system	25 patients fulfilled the inclusion criteria, nine were women and average age was 44. 20 patients had *IDH* mutation and 5 were *IDH* wild type	The textural parameter hom eneity could discriminate between LGG patien with *IDH* mutation (0.12, IQR 0.10-0.15) and *ID* wild type (0.07, IQR 0.06-0.09, p=0.005). Non of the other four an (energy, entropy, c rel on and inertia) were associated with molecular s tus. Using ROC curves, the area under curve f predicting *IDH* mutati 905 for hom eneity, 0.840 f tumor vo nd 0.940 r the combi ed parameters o or volume and hom eneity. Molecular st s could not be predic using the four othe chosen texture param (energy, entropy correlation and inertia). L ith *IDH* mu ation with or without 1p19q co uld not be separated.
Shofty et al. 2018	Low-grade Gliomas (LGG) classified into three distinct groups based on *IDH1* mutation and 1p/19g codeletion status	Linear SVM, quadratic SVM, cubic SVM, fine Gaussian SVM, medium Gaussian SVM, coarse Gaussian SVM. 17 machine learning classifiers in total	47 patients diagnosed with LGG with *IDH1*-mutated tumors nd a genetic analysis for 1p/19 deletion status were included in is dy. A total of 152 features cluding location and texture, w re extracted from fluid-atten d inversion recovery; weig ted; and post-contrast images.	Best ification was obtained using the Ensemble B d Trees classifier, with sensitivity 92%, specific 83% and accuracy 87%, and with a ea under the curve 0.87.
X. Zhang et al. 2018	*IDH* and *TP53* mutations	SVM model	The data t used in the study w s the LGG p ct of The Cancer m aging Archi CIA), which on tained 199 sub er, the qualified subjects in this study were 103. Thef 103 LrGG patients were s development (n=73) and v datio =30) cohorts.	The accuracies and area under the curves for *IDH* and *TP53* mutations were reported to be 84.9%, 0.830, and 92.0%, 0.949, respectively, on the development cohort, and 80.0%, 0.792, and 85.0%, 0.869, respectively, on the validation cohort.
Yuqi Han et al. 2018	1p/19q co-deletion in lower-grade glioma	Random forest / radiomics signature	patients histopathologically diag osed with lower-grade glioma. Clini l p rameters were recorded for each patient. Radiomics analysis performed by extracting 647 MRI- d features and applied random forest algorithm to generate a radiomics signature for predicting 1p/19q co-deletion in the training cohort (n=184).	The study reported the accuracy of the radiomics signature in predicting 1p/19q co-deletion, with areas under curve (AUCs) of 0.887 and 0.760 on the training and validation cohorts, respectively. The study also compared the radiomics signature to a clinical model and found that the radiomics signature outperformed the clinical model in predicting 1p/19q co-deletion.
Citak-Er et al. 2018	Grade III d IV gli as	A three-step SVM- based evaluation ro ch	The patients were newly diagnosed as having glioma, with a mean age of 49.5 ± 12.8 years (range, 23–79 years) 2. There were 18 females and 25 males, and the glioma grades were determined through the histopathologic analysis of surgical specimens	A machine-learning model based on support vector machine algorithm with linear kernel achieved an accuracy of 93.0%, a specificity of 86.7%, and a sensitivity of 96.4% for the grading of gliomas using ten-fold cross validation based on the proposed subset of the mp-MRI features.
De Looze et al. 2018	Glio rade and *IDH*	Random forest	The study included patients with a histopathologically proven diffuse glioma (grade II-IV) from a single neuropathological database in the center, selected from 2015 to 2017 The study employed 161 features derived from MR images to analyze using both Gabor texture and image intensity characteristics.	96% level of accuracy in differentiating between a low-grade and high-grade glioma. grade II/III; area under the receiver operating characteristic curve (AUC)=98%, sensitivity=0.82, specificity=0.94; grade II/IV; AUC=100%, sensitivity=1.0, specificity=1.0; grade III/IV: AUC=97%, sensitivity=0.83, specificity=0.97. Discrimination of *IDH* status: AUC of 88%, sensitivity=0.81, specificity=0.77.

Lu et al. 2018	*IDH* mutation, 1p/19q codeletion, and Telomerase Reverse Transcriptase (TERT) mutation, ATRX	The study tested nine different machine-learning models and selected the best model for each binary classification task based on the performance of the model on the training dataset SVM models: Linear, cubic, Bootstrap-aggrega ted (bagged) tree algorithm, AdaBoost algorithm with decision tree, RUSBoost algorithm with decision tree.	The data from 456 subjects with gliomas were obtained from The Cancer Imaging Archive. Overall, 214 subjects, including 106 cases of glioblastomas and 108 cases of lower grade gliomas with preoperative MRI, survival data, histology, *IDH*, and 1p/19q status were included.	The *IDH* and 1p/19q status of gliomas can be classified by radiomics and machine-learning approaches, with areas under ROC curves between 0.922 and 0.975 and accuracies between 87.7% and 96.1% estimated on the training dataset. The test on the model showed a comparable model performance with that on the training dataset, suggesting the efficacy of the trained classifiers. The classification of 5 molecular subtypes solely based on the MR phenotypes showed 81.8% accuracy and a higher accuracy of 8[illegible] was achieved when the histology diagnosis was available.
Y. Li, Qian,et al. 2018	Prediction of p53, *CDKN2A/B* homozygous deletion, *IDH* mutations, 1p/19q co-deletions, and methylation profiles.	LASSO, SVM	The study analyzed preoperative MR images from 272 patients with primary WHO grade II/III glioma, including 179 with grade II and 93 with grade III. The proportion of patients with a p53 mutation was 42.2% and 48.9% in the training and validation sets, respectively. The study did not find any significant differences in age, sex, grade, p53 status, or tumor location between the two sets of patients. Data was collected from the Chinese Glioma Genome Atlas database.	The prediction accuracies based on the area under the curve were 89.6% in the training set and 76.3% in the validation set, which were better than individual features.
Chang et al. 2018	*IDH* Status in low- and high-grade gliomas	Residual neural networks	Preoperative imaging was acquired for 201 patients from the Hospital of University of Pennsylvania (HUP), 157 patients from Brigham and Women's Hospital (BWH), and 138 patients from The cancer Imaging Archive (TIA) and divided into training, validation and testing sets.	Achieved *IDH* prediction accuracies of 82.8% (AUC=0.90), 83.0% (AUC=0.93), and 85.7% (AUC=0.94) within training, validation, and testing sets, respectively. When age at diagnosis was incorporated into the model, the training, validation, and testing accuracies increased to 87.3% (AUC=0.93), 87.6% (AUC =0.95), and 89.1% (AUC=0.95), respectively.
H. Zhou et al. 2017	wild-type *IDH* versus *IDH1* mutation; *IDH1* mutation with 1p/19q codeletion versus *IDH1* mutation without 1p/19q codeletion; grade II versus grade III LGGs; and progression versus non progression of LGGs	Multivariable texture model constructed using logistic regression	The study identified 165 patients with diffuse low- and intermediate-grade gliomas (histological grades II and III) from TCGA who have overlapping pre-surgical MRI data from The Cancer Imaging Archive (TCIA).	Interrater analysis showed significant agreement in all imaging features (k=0.703-1.000). On multivariate Cox regression analysis, no enhancement and a smooth non-enhancing margin were associated with longer Progression-Free Survival (PFS), while a smooth non-enhancing margin was associated with longer Overall Survival (OS) after taking into account age, grade, tumor location, histology, extent of resection, and *IDH1* 1p/19q subtype. Using logistic regression and bootstrap testing evaluations, texture models were found to possess higher prediction potential for *IDH1* mutation, 1p/19q codeletion status, histological grade, and progression of LGGs than VASARI features, with areas under the receiver-operating characteristic curves of 0.86 ± 0.01, 0.96 ± 0.01, 0.86=0.01, and 0.80 ± 0.01, respectively.
Bakas et al. 2017	LGG and GBM	GLISTR and GLISTRboost	Pre-operative multimodal Magnetic Resonance Imaging (MRI) (n=243) of the multi-institutional glioma collections of The Cancer Genome Atlas (TCGA), publicly available in The cancer Imaging Archive (TIA). Pre-operative scans were identified in both glioblastoma (TCGA-GBM, n=135) and low-grade-glioma (TCGA-LGG, n=108) collections *via* radiological assessment.	The median DICE values with their corresponding Inter-Quartile Ranges (IQR) for the three evaluated regions, i.e., WT, TC, ET, were equal to 0.92 (IQR: 0.88–0.94), 0.88 (IQR: 0.81–0.93) and 0.88 (IQR: 0.81–0.91), respectively. Equivalently, the 95th percentile of the Hausdorff distance for WT, TC and ET were equal to 3.61 (IQR: 2.39–8.15), 4.06 (IQR: 2.39–7.29), and 2 (IQR: 1.41–2.83), respectively.

Supplementary Table 2: AI applied in assorted application types

Paper	Application type	Biomarkers Investigated	AI used	Demographics/Sample size/ Dataset	Accuracy/Effectiveness
Jin et al. 2021	Low-cost transformed slide scanner to capture histopathological images of glioma tissue sections, FISH, pyrosequencing (direct sequencing)	*IDH1* at codon 132, *IDH2* at codon 172, chromosome 1p and 19q codeletion, *CDKN2A/B* homozygous deletion and methylation profiles	"AI Neuropathologist", deep neural networks and Convolutional Neural Networks (CNNs)	The CNNs were trained and verified on over 79 990 histological patch images from 267 patients. CNN-based independent diagnostic testing was performed on data from 56 patients with 172[illegible]2 histological patch image	Patch [illegible]ccuracy of 86.5% and patient level accuracy of 87.5% for distinguishing different histopathological subtypes of glioma
Pellegrino et al. 2021	Next-generation sequencing	Pathogenic Single Nucleotide Variants (SNVs), Single Nucleotide Polymorphisms (SNPs), Multiple Nucleotide Variants (MNVs), insertions, and deletions.	Random forest	102,011 SNVs, [illegible] and indel variants from a [illegible] panel (18[illegible] genes sequenced for glioma). The dataset included a total of 7,301 samples. [illegible]ouse glioma panel was [illegible] define glioma grades and [illegible]rget the following genes by [illegible]uencing full sequences and [illegible]pots of *AKT1*, *ATRX*, *BRAF*, *CDKN2A*, CIC, EGFR, FGFR1, *H3F3A*, *HIST1H3B*, *IDH1*, *IDH2*, *NOTCH1*, *PIK3CA*, *PIK3R1*, *PTEN*, *PTPN11*, *TERT* and *TP53*.	Best error rate for RF was 0.22% (ntree=500 and mtry=4), with an AUC of 0.99. The final trained model with the neural network achieved an accuracy of 98% and an ROC-AUC of 0.99 with validation data. RF model was tested to interpret more than 2000 variants from our NGS database: 20 variants were misclassified (error rate <1%). The errors were nomenclature problems and false positives. After adding false positives to the training database and implementing the RF model routinely, the error rate was always <0.5%. The RF model shows excellent results for oncosomatic NGS interpretation.
N. Zhang et al. 2021	Weighted Gene Co-expression Network Analysis (WGCNA)	Gene expression profiles from three CGGA cohorts, 300 genes were identified from 806 genes. By performing elastic net regression analysis and PCA algorithm, 33 monocyte-related genes were derived from the [illegible] and their [illegible]efficient[illegible]ere [illegible]btained.	Neural network classifiers and [illegible]pport Vector Mac[illegible]	[illegible]ffuse glioma samples were collected from three databases: The Cancer [illegible]enome Atlas (TCGA), Chinese Glioma Genome Atlas (CGGA), and Gene Expression Omnibus (GEO). For the TCGA cohort (672 glioma samples), the RNA-seq data and corresponding clinical information were retrieved from TCGA database (http://cancergenome.nih.gov/). Three CGGA validation cohorts were employed in this study, including two RNA-seq cohorts (CGGA325 and CGGA693) and a microarray cohort (CGGAarray). The RNA-seq and microarray data, clinical and survival information were downloaded from the CGGA database (http://www.cgga.org.cn). Expression matrices of GSE108474 (414 glioma samples) were obtained from the GEO database (https://www.ncbi.nlm. nih.gov/geo/)	After establishing monocyte density as a suitable marker for survival prediction of gliomas, it was further investigated its prediction efficiency by developing a prognostic nomogram. Combing prognostic factors, including risk scores, patient ages, tumor grades, *IDH* mutation, and chromosome 1p/19q codeletion, a prognostic nomogram was developed. In TCGA dataset, the Kaplan-Meier survival curve demonstrated a good discrimination of survival probabilities of the two clusters ($p<0.0001$). The ROC curve confirmed the discriminative ability of this nomogram (AUC=0.802). Predicted probabilities corresponded well with the actual one- to five-year overall survival rates of glioma patients. The efficiency of the prognostic model was validated in CGGA 693 cohort. The Kaplan-Meier survival curve demonstrated a good discrimination of survival probabilities of the two clusters ($p<0.0001$). The ROC curve confirmed the discriminative ability of this nomogram (AUC=0.737). Predicted probabilities corresponded well with the actual four-year overall survival rates of glioma patients.

Nuechterlein et al. 2021	Microarray based CGH	Mutational status of *IDH2* and the presence or absence of co-deletion of whole chromosome arms 1p and 19q	Random forest classifier	SCNA data from 786 adult diffuse gliomas in The Cancer Genome Atlas (TCGA) to develop a two-stage classification system	Cross-validated results on TCGA SCNA data showed near perfect classification results. The astrocytic *IDH* mutation model validated well on four additional datasets (AUC=0.99, AUC=0.96) as did the 1p/19q-codeleted oligodendroglioma screen on the two datasets that contained oligodendrogliomas (AUC=0.97, MCC=0.94).
Matsui et al. 2020	Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), and Computed Tomography (CT)	Diffuse astrocytoma, *IDH*-wildtype diffuse astrocytoma, *IDH*-mutant oligodendroglioma	A deep learning model	The dataset used in the study consisted of 217 patients diagnosed with WHO grade II/III glioma at the Department of Neurosurgery at Tokyo Women's Medical University between January 2000 and March 2019	The model could predict the molecular subtype with an accuracy of 96.6% for the training dataset and 68.7% for the test dataset. The model achieved test accuracies of 62.5%, 60.4%, and 59.4% when it contained only MRI, MRI and PET, and MRI and CT data, respectively. This is the first model to double the expected value for a 3-group classification problem, when predicting the LGG molecular subtype.
C.J.Park et al. 2020	MRI/DTI histogram	DH status	Random forest	118 patients with pathologically confirmed lower-grade gliomas who underwent preoperative MRI at newly diagnosed lower-grade gliomas.	Adding DTI radiomics to conventional radiomics significantly improved the accuracy of *IDH* status subtyping (AUC, 0.900 [95% CI, 0.855–0.945], p=0.006), whereas adding DTI histogram parameters yielded nonsignificant trend toward improvement (0.869 [95% CI, 0.816–0.922], p=0.150) compared with the model with conventional radiomics alone (0.835 [95% CI, 0.773–0.896]).
Acs, Rantalainen, and Hartman, 2020	Digital Image analysis of hematoxylin and eosin-stained slides	Glioma grade	Deep neural network models	1061 Whole Slide Images (WSIs) with patient follow-up data from TCGA. Histology images and genomic data into a single unified prediction framework called Genomic Survival Convolutional Neural Network (GSCNN model).	The GSCNN model achieved higher prognostic accuracy than the current WHO paradigm based on genomic classification and histologic grading.
Kocher et al. 2020	PET, MRI, MR spectroscopy	*CDKN2A/B* homozygous deletion, *IDH* mutations, 1p/19q deletion, methylation profiles	Deep learning based machine learning methods	The accuracy was measured using a machine learning algorithm trained on a dataset of 120 patients with WHO grade III and IV gliomas	An accuracy of approximately 90% in assessing WHO grade in newly diagnosed gliomas
Cakmakci et al. 2020	1H HRMAS NMR spectroscopy	Pilocytic Astrocytoma (PAST-I), Astrocytoma Grade II (AST-II), Astrocytoma Grade III (AST-III), Glioblastoma(GBM), Oligodendroglioma Grade II(ODG-II),Oligodendroglioma Grade II-III (ODG-II_III), Oligodendroglioma Grade III(ODG-III or ODIII), Oligoastrocytoma grade II(OAST-II),Oligoastrocytoma grade III (OAST-III), Oligoastrocytoma grade II-III(OASTII-III), Ganglioglioma Grade II (GG-II), Ganglioglioma Grade III (GG-III), Dysembryoplastic Neuroepithelial Tumors (DNET), Gliosarcoma (GS)	Random forest-based machine learning The code is released at http://github.com/ciceklab/HRMAS_NC	247 primary brain tumor samples from 218 patients, 74 non-tumor brain tissue samples from epilepsy surgery of 54 patients, and spectra of 244 samples from excision cavity of patients 4. The study used a new and large dataset of glioma and control samples (n=565) that were labeled with a quantitative pathology analysis. The study was conducted in France, and the dataset is publicly available at https://zenodo.org/record/3951448.	A median AUC of 87.1% and AUPR of 96.1% for the best performing method, which was a random forest-based approach.

Herrera-Rios et al. 2020	Next-Generation Sequencing (NGS)	Notch1	Agilent Whole Human Genome Array platform, and Gene Set Variation Analysis (GSVA) analysis using the gsva package in R	The use of NGS of a glioma-tailored gene panel in 50 glioma patients from Switzerland and the expression of ALDH1A3 in 301 glioblastoma patients from China	Shows Brontictuzumab impairs cellular invasion of Glioma Stem-like Cells (GSCs) and reduces the expression of stem cell markers, suggesting that it may have potential in targeting the stem cell-like cell population that promotes the pathology of glioblastoma.Further investigation of Brontictuzumab as a targeted therapy for GSCs and shows that high Notch1 activation and the amount of suppression of Hey1/Hes1 mRNA and cleaved Notch1 protein upon treatment serve as predictive for resistance of GSCs against Brontictuzumab. These studies suggest that Brontictuzumab may have potential as a targeted therapy for glioblastoma patients, particularly those with elevated Notch1 transcription or stem cell-like cell population.
Uroshlev et al. 2020	NGS	Methylation level of cytosine-guanine (CpG) dinucleotides	Support Vector Machine (SVM) classifier with a linear kernel	100 randomly selected whole genome sequencing datasets from the 1000 Genomes project that contained reads mapped on a reference genome (*GRCh37*)	A simple support vector machine classifier based on this algorithm shows an accuracy of 84%. The method allows the detection of epigenetic markers purely based on mechanochemical DNA fragmentation, which can be detected by a simple analysis of the NGS sequencing data.
Orzan et al. 2020	RNA-sequencing analysis and immunohistoche mical profiling	Proneural (PN); Classical (CL); Mesenchymal (ME) of GBM molecular subgroups	The study used a machine learning approach to integrate immunohistoc hemical profiles based on expression of a restricted panel of gene classifiers to generate a GliTS based on protein quantification that allowed an efficient GliTS assignment when applied to an extended cohort.	The study analyzed a total of 248 glioblastoma samples, including 51 representative samples for RNA-sequencing analysis and 197 extended cohort samples for immunohistochemical profiling. he study dataset consisted of newly diagnosed pathologically confirmed GBMs (n=197) retrieved from the Institutional database of the Department of Pathology, Spedali Civili of Brescia, Italy.	The study reported high concordance between profiles obtained either by molecular or IHC-based approach, being 81.3% for CL and up to 90% for MES 12. This suggests that the integrated molecular and immunohistochemistry-based algorithm developed in the study has high accuracy and reliability in predicting glioblastoma transcriptional subtypes.
Suchorska et al. 2019	High-resolution contrast-enhanced T1-weighted MRI, pyrosequencing	*IDH1* and *IDH2* mutations, co-del 1p/19q	Random survival forest.	301 patients with either WHO grade II (n 181) or WHO grade III (n 120) were included.	In *IDH* mutation tumours only, both conventional Cox regression modelling and RSF analyses showed that CE on initial MRI is a prognostic factor for survival with dependence on volume ($p<0.05$).
Levitin et al. 2019	Single-cell RNA sequencing (scRNA-seq), low-pass Whole-Genome Sequencing (WGS) and MRI	Differentially expressed genes and factors associated with astrocyte-like glioma cells. In-house glioma panel was made to define glioma grades and to target the following genes by sequencing full sequences and hotspots of *AKT1*, *ATRX*, *BRAF*, *CDKN2A*, CIC, *EGFR*,*FGFR1*, *H3F3A*,*HIST1H3B*,*IDH1*, *IDH2*,*NOTCH1*, *PIK3CA*, *PIK3R1*, *PTEN*,*PTPN11*, *TERT* and *TP53*.	Single-cell Hierarchical Poisson Factorization (scHPF) Code is available at: https://github.com/simslab/scHPF.	102,011 SNVs, SNPs, and indel variants from a 27-gene panel sequenced for colorectal cancer, melanoma, lung cancer, and 18 genes sequenced for glioma cancer. The dataset included a total of 7,301 samples. scRNA-seq data: Gene Expression Omnibus GSE116621 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE116621)	Data available at: Gene Expression Omnibus GSE116621 (https://www.ncbi.nlm.nih.gov /geo/query/acc.cgi?acc=GSE 116621)

Wei et al. 2019	MRI, pyro-sequencing, spectrophotomet ry, PCR amplification	MGMT promoter methylation in astrocytomas	The maximum Relevance and Minimum Redundancy (mRMR) algorithm	3,051 features extracted from the Region of Interest (ROI) on both the tumour and the peritumoral habitats	Exhibited supreme power for predicting MGMT promoter methylation, with area under the curve values of 0.925 in the training cohort and 0.902 in the validation cohort. Performance of the radiomics signature surpassed that of clinical factors and ADC parameters. Moreover, the radiomics approach successfully divided patients into high- risk and low-risk groups for overall survival after TMZ chemotherapy (p=0.03).
Fukuma et al. 2019	MRI, immunohistoche mistry and DNA sequencing	*IDH* wild type, *IDH* and pTERT co- mutations, *IDH* mutant and pTERT wild type	Linear Support Vector Machine (SVM), Deep Neural Networks (DNN) and Convolutional Neural Networks (CNN) (AlexNet)	The dataset included 199 cases from 11 institutions, with patients aged 20 years or older, frozen or fresh tissue available for genomic analysis, pre-operative MRI available for radiomic analysis, and a local diagnosis of lower grade (WHO grade II/III) glioma based on the fourth edition of the WHO classification	The classification accuracy between the lesion and normal tissue was highest when the texture features from conv5 were used to train the SVM classifier (98.4 ± 1.9% (0.999 ± 0.003) Using all features, an accuracy of 63.1% was achieved, which was significantly higher than the accuracy obtained from using either the radiomic features or patient age alone. In particular, prediction of the pTERT mutation was significantly improved by the CNN texture features.
van der Voort et al. 2019	MRI imaging Fluorescence In Situ Hybridization (FISH), targeted next-generation sequencing panel using an Ion Torrent Personal Genome Machine	1p/19q codeletion status of presumed low-grade glioma	Ensemble of five SVMs	Preoperative brain MR images from 284 (424 patients identified in the EMC/HMC dataset after initial screening of presumed low grade glioma) patients who had undergone biopsy or resection of presumed LGG, which were used to train the support vector machine algorithm. The performance of the algorithm was compared with tissue diagnosis on an external validation dataset of MR images from 129 patients with LGG from The Cancer Imaging Archive (TCIA) There were 159 patients screened from the TCIA dataset.	The algorithm had a higher predictive performance than the average of the neurosurgeons (AUC 0.52) but lower than that of the neuroradiologists (AUC of 0.81). There was a wide variability between clinical experts
Z.-C.Li et al. 2018	MRI, Pyrosequencing analysis (Polymerase Chain Reaction (PCR)	*IDH1* (R132 region)	CNN, Boruta algorithm, Random Forest	651 patients from The Cancer Imaging Archive (TCIA) publicly-available dataset (www. cancerimagingarchive .net) and three local institutions between January 2013 and July 2017 were analyzed. Finally: Retrospective multi-center study, 225 patients were included. A total of 1614 multiregional features were extracted from enhancement area, non-enhancement area, necrosis, edema, tumor core, and whole tumor in multiparametric MRI. The training cohort comprised 58 patients from TCIA and 60 patients from Sun Yat-sen university cancer center. The validation cohort comprised 45 patients from The 3rd Affiliated Hospital of Sun Yat-sen University and 62 patients from Guangzhou general hospital of guangzhou military command.	The edema model achieved the best accuracy of 96% and the best F1-score of 0.75 while the non-enhancement model achieved the best Area Under the receiver operating characteristic Curve (AUC) of 0.88 in the validation cohort. The overall performance of the tumor-core model (accuracy 0.96, AUC 0.86 and F1-scorE 0.75) and the whole-tumor model (accuracy 0.96, AUC 0.88 and F1-score 0.75) was slightly better than the single-regional models. The 8-feature all-region radiomics model achieved an improved overall performance of an accuracy 96%, an AUC 0.90, and an F1-score 0.78. Among all models, the model combining all-region imaging features with age achieved the best performance of an accuracy 97%, an AUC 0.96, and an F1-score 0.84.

L. Zhou et al. 2018	Microarray data analysis	9,341 DEGs were identified, including 9,175 upregulated genes and 166 downregulated genes	The limma package in R software, Morpheus online software, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses	The original data for the study was downloaded from the Gene Expression Omnibus (GEO) database, specifically the GSE15824 and GSE51062 datasets. These datasets are publicly available and were used for gene expression analysis and identification of differentially expressed genes in glioblastoma. The datasets were based on the Affymetrix GPL570 platform (Affymetrix Human Genome U133 Plus 2.0 Array) and included 5 normal samples and 74 GBM samples.	The study identified several genes that are good prognostic biomarkers for glioblastoma including RANBP17, ZNF734, NLRP2, GPR1, CCDC81, SH3RF1, and TM7SF4.
Liu et al. 2018	MRI and RNA sequencing	Nine textural features from Contrast-Enhanced(CE) MRI images that were associated with angiogenesis and could predict patient survival	Random Survival Forest (RSF)	The algorithm was trained on the RNA sequencing data from the training set of 91 anaplastic glioma from the The Cancer Genome Atlas (TCGA) patients and validated on an independent set of 64 patients from the CGGA database.	A 48-gene signature for CE was identified in TCGA and validated in CGGA dataset (area under the curve=0.9787). Seven genes derived from the CE-specific signature could stratify AG patients into two subgroups based on overall survival time according to corresponding risk score. Nine prognostic quantitative radiomic features of CE were found and the underlying biological processes of them was investigated.
Blanc-Durand et al. 2018	Positron Emission Tomography (PET)	To demonstrate the feasibility of an automated 18F-Fluoro-Ethyl-Tyrosine (18F-FET) PET lesion detection and segmentation relying on a full 3D U-Net Convolutional Neural Network (CNN).	Full 3D U-Net convolutional neural network	Thirty-seven patients were included (26 [70%] in the training set and 11 [30%] in the validation set).	All 11 lesions were accurately detected with no false positive, resulting in a sensitivity and a specificity for the detection at the tumor level of 100%. After 150 epochs, DSC reached 0.7924 in the training set and 0.7911 in the validation set. After morphological dilatation and fixed thresholding of the predicted U-Net mask a substantial improvement of the DSC to 0.8231 (+4.1%) was noted. At the voxel level, this segmentation led to a 0.88 sensitivity [95% CI, 87.1 to,88.2%] a 0.99 specificity[99.9 to 99.9%], a 0.78 positive predictive value:[76.9 to 78.3%], and a 0.99 negative predictive value[99.9 to 99.9%].
Bolis et al. 2017	RNA-sequencing	21 genes that were co-expressed in a tumor-type independent manner	ATRA-21 model	The data set used in the study was RNA-sequencing data for 10,080 patients and 33 different tumor types derived from The Cancer Genome Atlas (TCGA) and Leucegene datasets. The RNA-sequencing data was completely re-processed and used for machine learning methods and network analysis	On average, lower grade gliomas are predicted to be even more sensitive to ATRA than Acute Promyelocytic Leukemia (APL)

Way et al. 2017	Multi-array analysis	NF1 ans associated causal genes, including QPRT, RSL1D1 (CSIG), PPEF, PEG10, and ATF5	Logistic regression classifiers	Using RNAseg data from The Cancer Genome Atlas (TCGA), an ensemble of 500 logistic regression classifiers was trained that integrated mutation status with whole transcriptomes to predict NF1 inactivation in glioblastoma (GBM).	On TCGA data the classifier detected NF1 mutated tumors (test set area under the receiver operating characteristic curve (AUROC) quantile=0.53-0.95) over 50 random initializations. On RNA-Seq data transformed into the space of expression microarrays, this method produced a classifier with similar performance (test set AUROC mean=0.77, 95% quantile=0.53-0.96).
Brunelli Et al. 2017	Multi-array analysis	NF1 ans associated causal genes, including QPRT, RSL1D1 (CSIG), PPEF, PEG10, and ATF5	Machine learning classifier	Using RNAseq data from The Cancer Genome Atlas (TCGA), an ensemble of 500 logistic regression classifiers were trained that integrated mutation status with whole transcriptomes to predict NF1 inactivation in Glioblastoma (GBM).	On TCGA data, the classifier detected NF1 mutated tumors (test set Area Under the Receiver Operating characteristic Curve (AUROC) mean=0.77, 95% quantile=0.53-0.95) over 50 random initializations. On RNA-Seq data transformed into the space of gene expression microarrays, this method produced a classifier with similar performance (test set AUROC mean=0.77, 95% quantile=0.53-0.96).The ensemble classifier trained on the transformed TCGA data was applied to a microarray validation set of 12 samples with matched RNA and NF1 protein-level measurements. The classifier's NF1 score was associated with NF1 protein concentration in these samples.
Wang et al. 2010	PCR analysis	*ATM*, *ATR*, Chk1 and Chk2	Silico analysis	Tissues from ten normal brains and thirty human gliomas were utilized for the first real-time PCR analysis, and another twelve normal brain tissues and forty gliomas were used for confirmation	Expression of *ATM*, *ATR*, Chk1 and Chk2 genes in gliomas decreased to 83.9%, 26.1%, 33.4% and 10.9%, respectively, relative to normal brain tissues. Significant differences were observed in *ATR*, Chk1 and Chk2 but not in *ATM* between glioma tissues and non-tumor tissues, indicating the loss of repair and checkpoint control in the glioma tissues. Chk2 had the most significant reduction in expression and was ~ 10-fold less in glioma compared to normal controls ($P<0.0001$). These data suggest that the down- regulation of *ATR*, Chk1 and Chk2 genes may be a characteristic of gliomas.