Radiomics of Pediatric Low-Grade Gliomas: Toward a Pretherapeutic Differentiation of \textit{BRAF}-Mutated and \textit{BRAF}-Fused Tumors


\textit{AJNR Am J Neuroradiol} published online 11 February 2021
http://www.ajnr.org/content/early/2021/02/11/ajnr.A6998
Radiomics of Pediatric Low-Grade Gliomas: Toward a Pretherapeutic Differentiation of BRAF-Mutated and BRAF-Fused Tumors


BACKGROUND AND PURPOSE: B-Raf proto-oncogene, serine/threonine kinase (BRAF) status has important implications for prognosis and therapy of pediatric low-grade gliomas. Currently, BRAF status classification relies on biopsy. Our aim was to train and validate a radiomics approach to predict BRAF fusion and BRAF V600E mutation.

MATERIALS AND METHODS: In this bi-institutional retrospective study, FLAIR MR imaging datasets of 115 pediatric patients with low-grade gliomas from 2 children's hospitals acquired between January 2009 and January 2016 were included and analyzed. Radiomics features were extracted from tumor segmentations, and the predictive model was tested using independent training and testing datasets, with all available tumor types. The model was selected on the basis of a grid search on the number of trees, opting for the best split for a random forest. We used the area under the receiver operating characteristic curve to evaluate model performance.

RESULTS: The training cohort consisted of 94 pediatric patients with low-grade gliomas (mean age, 9.4 years; 45 boys), and the external validation cohort comprised 21 pediatric patients with low-grade gliomas (mean age, 8.37 years; 12 boys). A 4-fold cross-validation scheme predicted BRAF status with an area under the curve of 0.75 (SD, 0.12) (95% confidence interval, 0.62–0.89) on the internal validation cohort. By means of the optimal hyperparameters determined by 4-fold cross-validation, the area under the curve for the external validation was 0.85. Age and tumor location were significant predictors of BRAF status (P values = .04 and <.001, respectively). Sex was not a significant predictor (P value = .96).

CONCLUSIONS: Radiomics-based prediction of BRAF status in pediatric low-grade gliomas appears feasible in this bi-institutional exploratory study.

ABBREVIATIONS: AUC = area under the curve; JPA = juvenile pilocytic astrocytoma; NPV = negative predictive value; pLGG = pediatric low-grade glioma; PPV = positive predictive value; ROC = receiver operating characteristic

Pediatric low-grade gliomas (pLGGs) are the most common brain tumors in children, accounting for approximately 40% of central nervous system tumors in childhood.1 pLGGs comprise a heterogeneous variety of tumors classified by the World Health Organization as grades I or II and include juvenile pilocytic astrocytoma (JPA), ganglioglioma, dysembryoplastic neuroepithelial tumor, pleomorphic xanthoastrocytoma, and diffuse low-grade glioma.2 A mainstay of pLGG therapy is surgical excision when possible, which may be curative in case of total resection.2 When total resection is not possible, pLGGs become a chronic disease with protracted reduction in the quality of life.2,3 While death from these tumors is rare with standard chemotherapy and radiation, 10-year progression-free survival is <50%.4,5 Thus, many patients will have multiple recurrences requiring multimodal therapy, leading to considerable morbidity.

In addition to patients with neurofibromatosis type I who develop pLGG, molecular characterization of sporadic pLGG has
also identified frequent alterations in the mitogen-activated protein kinas pathway, most commonly fusions or mutations in the B-Raf proto-oncogene, serine/threonine kinase (BRAF) gene. The 2 major BRAF gene alterations are BRAF fusion and BRAF V600E point mutation (p.V600E). The chromosomal alteration in BRAF fusion involves the duplication of the BRAF oncogene, followed by its insertion into one of several fusion targets, most often the KIAA1549 gene.6 The transcript of this duplication/fusion contains the kinase terminus of the BRAF protein but lacks the auto-regulatory domain, resulting in constant up-regulation of several downstream pathway elements. BRAF V600E point mutations constitutively activate BRAF, causing a deregulation in the mitogen-activated p.V600E protein kinase pathway.7

Lassaletta et al8 recently demonstrated that patient prognosis differed in pLGGs on the basis of the underlying molecular alteration. Patients with BRAF fusion and neurofibromatosis type 1 have a favorable outcome, while those with the BRAF V600E mutation, particularly in association with cyclin dependent kinase inhibitor 2A (CDKN2A) deletion, are at increased risk of progression and transformation.8,9 This finding has led to clinical trials using mitogen-activated protein kinase pathway–targeted agents such as mitogen-activated protein kinase enzyme inhibitors and BRAF V600E inhibitors for patients with molecular evidence of BRAF alterations. These new therapies seem promising, and many pLGGs that were refractory to traditional chemotherapy have had meaningful responses to these targeted agents.10,11

In the past decade, radiomics has emerged as an imaging-based method to link quantitative features extracted from medical images to outcomes, such as cancer genotype or survival.12,13 Radiomic signatures have been extensively investigated for different cancer sites including liver cancer,14 bone tumors,15 and adult brain tumors including glioblastoma,16 medulloblastoma,17 and midline high-grade glioma.18,19 To our knowledge, no prior study has investigated radiomic approaches to subtype pLGGs.

Using a bi-institutional cohort, we aimed to develop and validate a radiomic signature that is predictive of the BRAF status of pLGGs.

**MATERIALS AND METHODS**

**Patients**

This retrospective study was approved by the institutional review board or research ethics board of the 2 participating academic institutions: The Hospital for Sick Children (Toronto, Ontario, Canada) and the Lucile Packard Children’s Hospital (Stanford University, Palo Alto, California). Because of the retrospective nature of the study, informed consent was waived by the local research ethics boards. An interinstitutional data-transfer agreement was obtained for data-sharing. All patients were identified from the electronic health record data base at Toronto and Stanford from January 2009 to January 2016. Patient inclusion criteria were the following: 1) 0–18 years of age, 2) availability of molecular information on BRAF status in histopathologically confirmed pLGG, and 3) availability of preoperative brain MR imaging with a non-motion-degraded FLAIR sequence. Patients with histone H3 K27M mutation were excluded. Spinal cord tumors were also not considered for this study.

**Molecular Analysis**

BRAF fusion status was determined using an nCounter Metabolic Pathways Panel (NanoString Technologies) or fluorescence in situ hybridization, while the BRAF p.V600E mutation was determined using immunohistochemistry or droplet digital polymerase chain reaction as previously described.20 For most patients, molecular analysis was performed with formalin-fixed paraffin-embedded tissue that was obtained at the time of the operation. Nineteen patients had molecular subtyping based on frozen tissue.

**MR Imaging Acquisition, Data Retrieval, and Image Segmentation**

All patients from The Hospital for Sick Children, Toronto, underwent brain MR imaging at 1.5T or 3T across various vendors (Signa, GE Healthcare; Achieva, Philips Healthcare; Magnetom Skyra, Siemens). Sequences acquired included 2D axial and coronal T2 FLAIR (TR/TE, 7000–10,000/140–170 ms; 3- to 6-mm section thickness; 3- to 7.5-mm gap), 2D axial T2-weighted fast spin-echo, 3D axial or sagittal precontrast, and 3D axial gadolinium-based contrast agent–enhanced T1-weighted turbo or fast-field echo. Patients from the Lucile Packard Children’s Hospital, Stanford, underwent brain MR imaging at 1.5T or 3T from a single vendor (Signa or Discovery 750; GE Healthcare). MRIs were performed using the brain tumor protocol of the institution, which included 2D axial T2-weighted fast spin-echo, 2D axial or sagittal precontrast T1-weighted spin-echo, 2D axial T2 FLAIR (TR/TE, 7000–10,000/140–170 ms; 4- to 5-mm section thickness; 1- to 1.5-mm gap), and 2D axial gadolinium-based contrast agent–enhanced T1-weighted spin-echo sequences. All MR imaging data were extracted from the respective PACS and were de-identified for further analyses.

Tumor segmentation was performed by a fellowship-trained pediatric neuroradiologist with 6 years of neuroradiology research experience (M.W.W.) using 3D Slicer (Version 4.10.2; 21 http://www.slicer.org). The scripted loadable module SlicerRadiomics extension was used to obtain access to the radiomics feature-calculation classes implemented in the pyradiomics library (http://pyradiomics.readthedocs.io/). This extension selects all available feature classes and ensures isotropic resampling under “resampled voxel size” when extracting 3D features. The bin width was set to 25 (ie, default), and symmetric gray level co-occurrence matrix was enforced. Semiautomated tumor segmentation on FLAIR images was performed with the level tracing-effect tool. This semiautomatic approach had been found superior to multiuser manual delineation with regard to the reproducibility and robustness of results.17 Final and proper placement of ROIs was confirmed by a board-certified neuroradiologist (B.B.E.-W., with 15 years of postfellowship experience).

**Radiomic Feature-Extraction Methodology**

A total of 851 MR imaging–based radiomic features were extracted from the ROIs on FLAIR images. Radiomic features included histogram, shape, and texture features with and without wavelet-based filters. Features of Laplacian of Gaussian filters were not extracted. All features are summarized in the Online Supplemental Data. Bias field correction before z score normalization was used to standardize the range of all image features.
features.\textsuperscript{22,23} Once the features were extracted, we applied z score normalization again followed by L2 normalization to the features of cohort 1 and used the distribution of the features in cohort 1 (training data) to normalize cohort 2 (test data). Details of pre-processing and radiomic feature extraction in 3D Slicer and other software have been described elsewhere.\textsuperscript{12,16,24}

**Statistical Analysis**

**Feature Selection, Radiomics, and Machine Learning Approach.**

We used random forest as the classification model\textsuperscript{25} and performed both internal cross-validations using cohort 1 data (n = 94) as well as external validation using cohort 2 (n = 21) with the molecular subtype as the end point.

**Internal Cross-Validation.** First, we used cohort 1 in k-fold cross-validation to find the best hyperparameter for the random forest model, namely the number of trees in the random forest. Once the optimal number of trees was found, it was used to perform 4-fold internal cross-validations using cohort 1.

**External Validation.** Next, using the optimal number of trees found in the previous step, the entire dataset in cohort 1 was used to train a random forest model, which was then tested on cohort 2. Cohort 2 was never used in any stage of the training of the random forest model and was only used for external validation.

Next, the area under the receiver operating characteristic (ROC) curve (AUC) was calculated for both internal and external validations. In addition, the top 10 features that contributed the most to the random forest model were extracted.

**Clinical Factors.** For clinical factors (age, sex, anatomic location of tumor), logistic regression was performed to determine the predictive power of each factor in determining the molecular subtypes.

**RESULTS**

**Patients**

A total of 115 patients were included (The Hospital for Sick Children, n = 94, Lucile Packard Children’s Hospital, n = 21) comprising 57 boys; mean age, 9.21 (SD, 10.81) years (Table 1). Patient demographic and pathologic information consisted of age at diagnosis, sex, histologic diagnosis, molecular diagnosis regarding the \textit{BRAF} status, and anatomic location of the tumor (supratentorial versus infratentorial). We used the patient data from The Hospital for Sick Children (cohort 1, n = 94) for internal validation using cross-validation. We then used cohort 1 to train an optimized model and tested it (external validation) on the patient data from the Lucile Packard Children’s Hospital (cohort 2, n = 21).

**Radiomics Model Evaluation**

Internal Validation. The number of trees, best-performing features, AUC, and other classification metrics for the 4-fold cross-validation are shown in Tables 2 and 3. For the internal validation, only data from cohort 1 were used. The ROC curve with a 4-fold cross-validation scheme to predict \textit{BRAF} status is shown in Fig 1. The internal validation yielded an AUC of 0.75 (SD, 0.1) (95% CI, 0.62–0.89) for the 4-fold cross-validation. The mean sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were 0.72, 95% CI, 0.60–0.84; 0.86, 95% CI, 0.76–0.95; 0.73, 95% CI, 0.60–0.87; and 0.85, 95% CI, 0.80–0.91, respectively.

**External Validation.** By means of the optimal hyperparameters obtained from 4-fold internal validation, the AUC for external validation was 0.85 (Fig 2). The Youden J statistic\textsuperscript{26} was used to determine the optimal threshold on the external ROC curve to calculate sensitivity, specificity, PPV, and NPV, which are listed in the Online Supplemental Data.

**Identification of Discriminative Clinical Factors**

**Clinical Factors.** The distribution of infratentorial and supratentorial tumors is shown in Table 1. Predictive clinical factors for \textit{BRAF} status were analyzed on cohort 1 (Table 4). Older age was a predictor of \textit{BRAF} V600E mutation (\textit{P} value = .04; OR, 1.14; 95% CI, 1.008–1.30) and as expected, supratentorial tumor location was a very strong predictor of \textit{BRAF} V600E (\textit{P} value < .001; OR, 18.80; 95% CI, 4.96–94.6). Sex was not a predictor (\textit{P} value = .96).

**Combined Clinical and Radiomics Model Evaluation**

Internal Validation. We appended the 2 predictive clinical factors for \textit{BRAF} status (age and tumor location) to the radiomics model outlined above. For the internal validation, only data from cohort 1 were used. The internal validation yielded an AUC of 0.77 (SD, 0.10) (95% CI, 0.65–0.88) for 4-fold cross-validation. The mean sensitivity, specificity, PPV, and NPV were 0.72, 95% CI, 0.60–0.84; 0.86, 95% CI, 0.78–0.93; 0.73, 95% CI, 0.63–0.83; and 0.86, 95% CI, 0.80–0.91, respectively. The improvement of our internal cross-validation compared with the radiomics-only model was not statistically significant (\textit{P} value > .05).

### Table 1: Patient demographics

<table>
<thead>
<tr>
<th></th>
<th>Toronto</th>
<th>Stanford</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>94</td>
<td>21</td>
</tr>
<tr>
<td>Age (mean) (yr)</td>
<td>9.4</td>
<td>8.37</td>
</tr>
<tr>
<td>Male sex (No.) (%)</td>
<td>45 (48)</td>
<td>12 (57)</td>
</tr>
<tr>
<td>Histologic diagnosis (No.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JPA</td>
<td>54</td>
<td>12</td>
</tr>
<tr>
<td>GG</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>LGA</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>PMA</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>PKA</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>DNET</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>GC</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>ODG</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Molecular subgroup (No.) (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{BRAF} fusion</td>
<td>62 (66)</td>
<td>14 (66)</td>
</tr>
<tr>
<td>\textit{BRAF} mutation</td>
<td>32 (34)</td>
<td>7 (34)</td>
</tr>
<tr>
<td>FLAIR availability (No.)</td>
<td>94</td>
<td>21</td>
</tr>
<tr>
<td>Supratentorial (No.) (%)</td>
<td>43 (46)</td>
<td>6 (28)</td>
</tr>
<tr>
<td>Transtentorial (No.) (%)</td>
<td>0</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Infratentorial (No.) (%)</td>
<td>51 (54)</td>
<td>14 (67)</td>
</tr>
</tbody>
</table>

**Note:** –GG indicates ganglioglioma; LGA, low-grade astrocytoma; PMA, pilomyxoid astrocytoma; PKA, pleomorphic xanthoastrocytoma; DNET, dysembryoplastic neuroepithelial tumor; DA, diffuse astrocytoma; GC, gangliocytoma; ODG, oligodendroglioma.
After we appended the 2 predictive clinical factors to the radiomics model, the AUC for external validation decreased to 0.67. The Youden J statistic was used to determine the optimal threshold on the external ROC curve to calculate sensitivity, specificity, PPV, and NPV, which are listed in the Online Supplemental Data.

**DISCUSSION**

In this bi-institutional study, we generated and validated a radiomic signature predictive of the \( \text{BRAF} \) status of pLGGs. The optimal random forest model achieved an AUC of 0.85 on the external validation dataset.

Currently, the molecular signature of pLGG is assessed through analysis of the tumor tissue. To that end, patients with nonresectable tumors are submitted to surgical procedures. Prognostication and targeted therapy depend on the mutational status. In this context, imaging could play a pivotal role if it allows identification of pLGG molecular subgroups. However, to date, we lack accurate imaging biomarkers that may facilitate this task.

Although genetic alterations of pLGGs are well-analyzed, little is known about the correlation between molecular markers and imaging characteristics. While many studies investigated the use of qualitative and quantitative features derived from conventional and advanced sequences to differentiate high- and low-grade pediatric brain tumors, only a few studies tried to link imaging characteristics to molecular markers. Ho et al described different MR imaging patterns based on 15 cases of \( \text{BRAF} \) V600-mutated diencephalic PLGGs and 25 cases of \( \text{BRAF} \) V600 wild-type JPA/pilomyxoid astrocytomas. Among their findings, which were based on analysis of T2WI and contrast-enhanced T1 sequences, they reported that \( \text{BRAF} \) V600 wild-type JPA/pilomyxoid astrocytoma presented predominantly as a solitary solid mass with homogeneous or heterogeneous contrast enhancement, whereas the mutated pLGG appeared multiloculated or multinodular following contrast administration. Quantitative imaging features differentiating pLGG molecular subgroups were studied in only 1 small case.

---

**Table 2: Performance of radiomic features**

<table>
<thead>
<tr>
<th>No. of Folds</th>
<th>No. of Trees</th>
<th>AUC (SD) (95% CI)</th>
<th>Mean Sensitivity (95% CI)</th>
<th>Mean Specificity (95% CI)</th>
<th>Mean PPV (95% CI)</th>
<th>Mean NPV (95% CI)</th>
<th>Top 10 Predictive Features on the External Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>25</td>
<td>0.75 (SD, 0.12)</td>
<td>0.72 (0.60–0.84)</td>
<td>0.86 (0.76–0.95)</td>
<td>0.73 (0.60–0.87)</td>
<td>0.85 (0.80–0.91)</td>
<td>(585, 374, 761, 22, 17, 560, 344, 258, 148, 108)</td>
</tr>
</tbody>
</table>

**Note:** SD indicates Standard Deviation; PPV, Positive Predictive Value; NPV; Negative Predictive Value.

**Table 3: Predictive radiomic features**

<table>
<thead>
<tr>
<th>No.</th>
<th>Source</th>
<th>Feature Category</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>585</td>
<td>3D wavelet transform</td>
<td>Gray-level difference matrix</td>
<td>Small dependence low gray-level emphasis</td>
</tr>
<tr>
<td>374</td>
<td>3D wavelet transform</td>
<td>Gray-level size zone matrix</td>
<td>Zone percentage</td>
</tr>
<tr>
<td>761</td>
<td>3D wavelet transform</td>
<td>Gray-level difference matrix</td>
<td>Dependence entropy</td>
</tr>
<tr>
<td>22</td>
<td>Original</td>
<td>Gray-level difference matrix</td>
<td>Dependence nonuniformity normalized</td>
</tr>
<tr>
<td>17</td>
<td>Original</td>
<td>Gray-level difference matrix</td>
<td>Dependence entropy</td>
</tr>
<tr>
<td>560</td>
<td>3D wavelet transform</td>
<td>Gray-level size zone matrix</td>
<td>Zone percentage</td>
</tr>
<tr>
<td>344</td>
<td>3D wavelet transform</td>
<td>Histogram</td>
<td>Entropy</td>
</tr>
<tr>
<td>258</td>
<td>3D wavelet transform</td>
<td>Gray-level run-length</td>
<td>Gray-level variance</td>
</tr>
<tr>
<td>148</td>
<td>3D wavelet transform</td>
<td>Histogram</td>
<td>Uniformity</td>
</tr>
<tr>
<td>108</td>
<td>3D wavelet transform</td>
<td>Gray-level difference matrix</td>
<td>Gray-level variance</td>
</tr>
</tbody>
</table>

*Radiomic features are ranked from top to bottom according to their importance.*
series of 7 patients. Ishi et al found a lower T2WI signal and a larger T2WI/contrast-enhanced FLAIR mismatch to be indicative of BRAF V600E mutation in optic pathway gliomas. In their study, T2WI/contrast-enhanced FLAIR mismatch was defined as a mismatch of a tumor region with high signal intensity on T2WI or FLAIR sequences with enhancement on contrast-enhanced T1WI sequences.

In our study, we trained and validated radiomic features of FLAIR MR images to predict BRAF fusion or mutation status in pLGG. As expected, the location of the tumor and age at presentation significantly predicted the mutational status. Histologically and radiomorphologically, pLGGs are largely heterogeneous. On the FLAIR sequence alone, tumors display a variety of qualitative differences, including the volume of their cystic and solid components, sharp and indistinct borders, presence or absence of hemorrhage, location, and volume at initial presentation. Our training cohort reflected the large spectrum of pLGGs with regard to the prevalence of tumor types and imaging characteristics on the FLAIR sequence (Fig 3). However, the independent external validation cohort comprised JPA, ganglioloma, and pilomyxoid astrocytoma only. This feature may explain the difference between the internal and external prediction of our model (best model; internal AUC = 0.75, external AUC = 0.85) and warrants further investigation. A less comprehensive approach with prediction of BRAF status either restricted to 1 or a few pLGG subtypes or anatomic location such as the optic pathway or cervicomedullary junction may further improve prediction accuracy. Future studies could adopt a more restrictive approach and analyze molecular markers within a given pLGG type or anatomic location. Due to the need for a large sample size and the low prevalence of these tumors, radiomic studies may be limited to large multinational and multi-institutional collaborations.

Another factor that may further improve our model prediction is the incorporation of patient demographic information such as age at presentation and qualitative radiographic features such as tumor location. This may be particularly helpful for the BRAF V600E mutation, which is known to be strongly associated with supratentorial location as seen in our study.

Our study has limitations. Due to the retrospective and bi-institutional nature of the study, there was heterogeneity in the FLAIR sequence acquisition, including the use of different scanner vendors, field strengths, and imaging parameters. However, because the heterogeneity in image acquisition reflects clinical practice, a robust and predictive model needs to incorporate these technical variations. In addition, our exploratory study used only FLAIR images for feature discrimination and model development. Incorporating additional MR imaging sequences such as T2WI, DWI, and contrast-enhanced T1WI sequences could further increase random forest model performance.

CONCLUSIONS

We present the exploratory results for the application of radiomics and machine learning for the prediction of BRAF status in pLGGs using independent bi-institutional training and validation sets based on FLAIR images. The optimal random forest model achieved an AUC of 0.85 in the validation cohort. Future investigations with a larger sample size for all histologic tumor types are warranted to further improve BRAF classifier training and validation. The use of other imaging sequences, including DWI, T2WI, and contrast-enhanced T1WI, and patient age and tumor location, may also help improve prediction accuracy.
REFERENCES


