Bridging Pixels to Genes

Mana Moassefi, MD • *Bradley J. Erickson, MD, PhD*

Mana Moassefi, MD, is a postdoctoral research fellow at the Radiology Informatics Lab, Mayo Clinic. Her research focuses on advancements in neuroradiology and abdominal radiology. She is an alumnus of the Academy Council of Early Career Investigators in Imaging, class of 2024, and received the Roger A. Bauman, MD Award from the Society for Imaging Informatics in Medicine (SIIM). Additionally, she serves on SIIM's education and machine learning education subcommittees.

Bradley J. Erickson, MD, PhD, is professor of radiology and informatics at the Mayo Clinic. He directs the Radiology Informatics Lab and has been a clinical neuroradiologist for 30 years.

Glioma is a life-threatening brain tumor influenced by its genetic and histologic features, notably isocitrate dehydrogenase *(IDH)* mutations, a key factor in tumor development and progression. The reference-standard method to identify *IDH* status involves obtaining tissue samples and conducting histologic assays. Given the risks associated with invasive biopsy, combining noninvasive techniques, including imaging, radiomics, and machine learning (ML) algorithms, to predict *IDH* mutation status can be highly beneficial.

Numerous studies have used ML algorithms to analyze extracted radiomics features to classify glioma tumors as either *IDH* mutated or wild type, with mixed success rates (1–3). A recently published meta-analysis of 26 studies involving 3280 patients revealed that radiomics features achieved a pooled sensitivity of 79% and specificity of 80% in detecting *IDH* mutations (4). This evidence underscores the potential of radiomics, combined with ML, to predict *IDH* mutation status from MRI scans, which could significantly improve patient care by reducing the need for invasive diagnostic procedures.

In this issue of *Radiology: Artificial Intelligence*, Truong and colleagues introduced an ML-based tool that classifies gliomas into *IDH* mutated and wild-type categories using MRI radiomics features (5). They used retrospective MRI scans from six datasets, including three public sources and three private datasets from distinct institutions. The data included patients newly diagnosed with glioma with known *IDH* mutation status and available preoperative MRI scans. The MRI data for this study, sourced from various providers, were subjected to different preprocessing tools due to the unavailability of raw data in all cases. The substantial size of the testing set, combined with the rigorous validation across diverse datasets, significantly enhances the reliability and ensures the generalizability of the predictions, contributing to the high quality and trustworthiness of the study's findings.

To address the challenge of dataset imbalance, where the prevalence of *IDH*-mutated tumors was approximately 27%, Truong et al implemented a two-stage training framework. This strategy achieved an accuracy of 79.9% for the mutated class and 93.5% for the wild-type class. In contrast, using the synthetic minority oversampling technique, which augments the mutated class to equal the number of wild-type cases, resulted in accuracies of 73.6% and 94.9%, respectively. They conclude that the two-stage approach effectively tackled the imbalance, potentially leading to more reliable diagnostic predictions.

Radiomics features from brain tumors were extracted using the PyRadiomics package, focusing on two regions of interest (ROI): the whole tumor and the nonenhancing, necrosis, and edema regions. This broad and inclusive approach diverges from previous studies that targeted specific tumor subcompartments like enhancing tumor regions; the approach allowed for training and testing across all cases in the datasets, which boosts model generalizability.

A notable aspect of this study involved enhancing the extraction of radiomics features by generating 12 derived images for each MRI sequence in addition to the original images. These derived images were created using a variety of filters, including wavelet filtering at four levels, Laplacian of Gaussian filtering with varying blurring levels, and mathematical transformations such as squaring, square root, logarithm, and exponential functions. By emphasizing different characteristics of the original images, these derived images significantly enriched the diversity and depth of the radiomics features extracted. The Boruta feature selection method was used to identify critical radiomics features for predicting *IDH* mutations by generating random shadow features to benchmark the significance of actual features. A total of 1197 features were extracted per ROI for each of the MRI sequences (T1-weighted, postcontrast T1-weighted, T2-weighted, and T2-weighted fluid-attenuated inversion

From Department of Radiology, the Radiology Informatics Laboratory, Mayo Clinic, 200 1st St SW, Rochester, MN 55905. Received April 29, 2024; revision requested May 3; revision received May 3; accepted May 9. **Address correspondence to** B.J.E. (email: *Bje@mayo.edu*).

Authors declared no funding for this work.

Conflicts of interest are listed at the end of this article.

See also article by Truong et al in this issue.

recovery), using both original and derived images enhanced by filters. The original, squared, and Laplacian of Gaussian images were the main contributors and informative among the image types. Similarly, from the PyRadiomics feature extractions, first-order statistics, 22 gray-level co-occurrence matrix, and 16 gray-level size zone matrix features were the most frequently selected feature classes for better prediction.

In this comprehensive study, Truong et al have developed a robust MRI radiomics-based ML model to predict *IDH* mutation status in gliomas, leveraging a remarkably extensive dataset from various institutions. Their best-performing models trained on The Cancer Imaging Archive (TCIA) dataset achieved the area under the receiver operating characteristic curve (AUC) values of 0.89 for UT Southwestern Medical Center (UTSW), 0.86 for New York University (NYU), 0.93 for University of Wisconsin-Madison (UWM), 0.94 for University of California San Francisco Preoperative Diffuse Glioma MRI dataset (UCSF), and 0.88 for Erasmus Glioma Database (EGD) test sets. The best-performing models trained on the UTSW dataset achieved slightly higher AUCs: 0.92 for TCIA, 0.88 for NYU, 0.96 for UWM, 0.93 for UCSF, and 0.90 for EGD.

The variations in MRI data properties, particularly the use of older imaging techniques in the TCIA datasets, including images post-1983, might result in the performance discrepancies observed in the models trained on TCIA and UTSW data (6). Specifically, the models trained using the UTSW dataset, which utilized more recent imaging technologies, demonstrated superior performance. This observation suggests that advancements in MRI technology could enhance the accuracy of radiomics-based prediction models. It also underscores the need for a detailed analysis of imaging protocols and their impact on model performance.

In this study, lack of access to the original Digital Imaging and Communications in Medicine data forced the authors to use three different skull stripping and coregistration tools along with a mix of automated and manual techniques for tumor regional segmentation. Despite the application of only one type of preprocessing in the training set, the study demonstrated strong performance on a large test set prepared with different techniques. Variations in MRI scanner types, acquisition techniques, preprocessing, and segmentation also introduced variability in the test sets. It is reassuring that the model performed consistently well across these diverse conditions, suggesting that it has high generalizability.

It is widely accepted that the accuracy of tumor segmentation is vital to assure the reliability of extracted radiomic features. This study demonstrated the feasibility of relying on automated segmentation for feature extraction, validated through testing on a substantial test set. This success sets the stage for broader adoption and trust in automated segmentation for radiomics features in other applications.

It seems that a more thorough evaluation and report of the findings, especially the selection and number of features used, would have been valuable. For instance, a reduced feature set might maintain performance and simplify the models without reducing efficacy. It is unclear if the authors did this investigation or simply used the set of features provided by the Boruta method. Additionally, the use of derived images, such as those produced through Laplacian of Gaussian filtering, should be critically evaluated to assess their necessity, considering the consistent performance metrics observed across different models.

This study applied only traditional radiomics metrics to extract information. Although the performance of this approach was good, it is valuable to explore deep learning techniques to more effectively handle the complexities inherent in the dataset and extracted features, potentially enhancing model performance beyond traditional ML methods.

The application of ML methods to routine MRI scans has opened the door for radiology to participate in the -omics revolution (5). Studies such as this one demonstrate a great deal of information present in images that humans are not able to appreciate. There is also the potential for more information to be found: Do the cases that ML gets "wrong" have some unique clinical features reflecting how tumor genetics are expressed? While molecular markers are clearly important for managing patients, is there additional information that radiology can provide to improve further the precision of care for patients? This study and others reflect that radiology has a great future in molecular medicine.

Disclosures of conflicts of interest: M.M. No relevant relationships. **B.J.E.** Research committee chair for SIIM; consultant to the editor for *Radiology: Artificial Intelligence*.

References

1. Li ZC, Bai H, Sun Q, et al. Multiregional radiomics profiling from multiparametric MRI: Identifying an imaging predictor of IDH1 mutation status in glioblastoma. Cancer Med 2018;7(12):5999–6009.
2. Lu J, Xu W, Chen X, Wang T, Li H. Noninvasive prediction of IDH mutation status in gliomas using preoperative multiparametric MRI radiomics nomogram: A mutlicenter study. Magn Reson Imaging 2023;104:72–79.
3. Zhou W, Wen J, Huang Q, et al. Development and validation of clinical-radiomics analysis for preoperative prediction of IDH mutation status and WHO grade in diffuse gliomas: a consecutive L-[methyl-11C] methionine cohort study with two PET scanners. Eur J Nucl Med Mol Imaging 2024;51(5):1423–1435.
4. Di Salle G, Tumminello L, Laino ME, et al. Accuracy of radiomics in predicting IDH mutation status in diffuse gliomas: a bivariate meta-analysis. Radiol Artif Intell 2024;6(1):e220257.
5. Truong NCD, Yogananda CGB, Wagner BC, et al. A two-stage training framework using multi-contrast MRI radiomics for IDH mutation status prediction in glioma. Radiol Artif Intell 2024;6(4):e230218.
6. Styliara EI, Astrakas LG, Alexiou G, et al. Survival outcome prediction in glioblastoma: insights from MRI radiomics. Curr Oncol 2024;31(4):2233–2243.