

MRI Radiogenomics of Pediatric Medulloblastoma: A Multicenter Study

Michael Zhang, MD • Samuel W. Wong, MS • Jason N. Wright, MD • Matthias W. Wagner, MD • Sebastian Toescu, MBChB • Michelle Han, MD • Lydia T. Tam, BS • Quan Zhou, PhD • Saman S. Ahmadian, MD • Katie Shpanskaya, MD • Seth Lummus, DO • Hollie Lai, MD • Azam Eghbal, MD • Alireza Radmanesh, MD • Jordan Nemelka • Stephen Harward II, MD, PhD • Michael Malinzak, MD, PhD • Suzanne Laughlin, MD • Sébastien Perreault, MD • Kristina R. M. Braun, MD • Robert M. Lober, MD, PhD • Yoon Jae Cho, MD • Birgit Ertl-Wagner, MD, PhD • Chang Y. Ho, MD • Kshiti Mankad, FRCR • Hannes Vogel, MD • Samuel H. Cheshier, MD, PhD • Thomas S. Jacques, FRCPath, PhD • Kristian Aquilina, MD • Paul G. Fisher, MD • Michael Taylor, MD • Tina Poussaint, MD • Nicholas A. Vitanza, MD • Gerald A. Grant, MD • Stefan Pfister, MD • Eric Thompson, MD • Alok Jaju, MD • Vijay Ramaswamy, MD, PhD • Kristen W. Yeom, MD

From the Departments of Neurosurgery (M.Z., Q.Z.), Otolaryngology–Head and Neck Surgery (Q.Z.), and Pathology (S.S.A., H.V.), Stanford Hospital and Clinics, Stanford, Calif; Departments of Radiology (M.Z., K.W.Y.), Neurology (P.G.F.), and Neurosurgery (G.A.G.), Lucile Packard Children's Hospital, Stanford University, 725 Welch Rd, G516, Palo Alto, CA 94304; Department of Statistics, Stanford University, Stanford, Calif (S.W.W.); Department of Radiology (J.N.W.) and Division of Pediatric Hematology/Oncology, Department of Pediatrics (N.A.V.), Seattle Children's Hospital, Seattle, Wash; Department of Radiology, Harborview Medical Center, Seattle, Wash (J.N.W.); Departments of Diagnostic Imaging (M.W.W., S. Laughlin, B.E.W.) and Surgery (M.T.) and Division of Haematology/Oncology, Department of Pediatrics (V.R.), The Hospital for Sick Children, Toronto, Canada; Departments of Neurosurgery (S.T., K.A.), Radiology (K.M.), and Developmental Biology & Cancer (T.S.J.), Great Ormond Street Institute of Child Health, London, UK; Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, Pa (M.H.); Stanford School of Medicine, Stanford University, Stanford, Calif (L.T.T.); Departments of Radiology (K.S., M.M.) and Neurosurgery (S.H.), Duke Children's Hospital & Health Center, Durham, NC; Department of Physiology and Nutrition, University of Colorado–Colorado Springs, Colorado Springs, Colo (S. Lummus); Department of Radiology, Children's Hospital of Orange County, Orange, Calif (H.L., A.E.); Department of Radiology, New York University Grossman School of Medicine, New York, NY (A.R.); Division of Pediatric Neurosurgery, Department of Neurosurgery, and Huntsman Cancer Institute, University of Utah School of Medicine, Intermountain Healthcare Primary Children's Hospital, Salt Lake City, Utah (J.N., S.H.C., E.T.); Division of Child Neurology, Department of Pediatrics, Centre Hospitalier Universitaire Sainte-Justine, Université de Montréal, Montréal, Canada (S. Perreault); Department of Clinical Radiology & Imaging Sciences, Riley Children's Hospital, Indianapolis, Ind (K.R.M.B., C.Y.H.); Division of Neurosurgery, Dayton Children's Hospital, Dayton, Ohio (R.M.L.); Department of Pediatrics, Doernbecher Children's Hospital, Portland, Ore (Y.J.C.); Department of Radiology, Boston Children's Hospital, Boston, Mass (T.P.); Department of Pediatrics, Hopp Children's Cancer Center, Heidelberg, Germany (S. Pfister); and Department of Medical Imaging, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, Ill (A.J.). Received August 25, 2021; revision requested October 26; revision received November 9; accepted February 8, 2022. Address correspondence to K.W.Y. (email: kyeom@stanford.edu).

M.Z. is funded by the National Institutes of Health (5T32CA009695-27). S.T. was supported by payment to the University College London (UCL) by Great Ormond Street Hospital (GOSH) Children's Charity. S.H.C. was supported by the Kathryn S.R. Lowry Endowed Chair at the University of Utah, Department of Neurosurgery. T.S.J. is supported by the Institute of Child Health – Newcastle University – Institute of Cancer Research High-Risk Childhood Brain Tumour (INSTINCT) Network and the Everest Centre, which are funded by The Brain Tumour Charity, as well as supported by the GOSH Children's Charity, Children with Cancer UK, Cancer Research UK, and the Olivia Hodson Cancer Fund. All research at the GOSH National Health Service Foundation Trust and UCL Great Ormond Street Institute of Child Health is made possible by the National Institute for Health Research GOSH Biomedical Research Centre. V.R. is supported by the Canadian Institutes for Health Research, the Canadian Cancer Society Emerging Scholars Award, the Garron Family Cancer Centre, and the C.R. Younger Foundation. K.W.Y. is supported by the American Brain Tumor Association (DG1800019).

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

See also the editorial by Verschakelen in this issue.

Radiology 2022; 000:1–11 • <https://doi.org/10.1148/radiol.212137> • Content codes:  

Background: Radiogenomics of pediatric medulloblastoma (MB) offers an opportunity for MB risk stratification, which may aid therapeutic decision making, family counseling, and selection of patient groups suitable for targeted genetic analysis.

Purpose: To develop machine learning strategies that identify the four clinically significant MB molecular subgroups.

Materials and Methods: In this retrospective study, consecutive pediatric patients with newly diagnosed MB at MRI at 12 international pediatric sites between July 1997 and May 2020 were identified. There were 1800 features extracted from T2- and contrast-enhanced T1-weighted preoperative MRI scans. A two-stage sequential classifier was designed—one that first identifies non-wingless (*WNT*) and non-sonic hedgehog (*SHH*) MB and then differentiates therapeutically relevant *WNT* from *SHH*. Further, a classifier that distinguishes high-risk group 3 from group 4 MB was developed. An independent, binary subgroup analysis was conducted to uncover radiomics features unique to infantile versus childhood *SHH* subgroups. The best-performing models from six candidate classifiers were selected, and performance was measured on holdout test sets. CIs were obtained by bootstrapping the test sets for 2000 random samples. Model accuracy score was compared with the no-information rate using the Wald test.

Results: The study cohort comprised 263 patients (mean age \pm SD at diagnosis, 87 months \pm 60; 166 boys). A two-stage classifier outperformed a single-stage multiclass classifier. The combined, sequential classifier achieved a microaveraged F1 score of 88% and a binary F1 score of 95% specifically for *WNT*. A group 3 versus group 4 classifier achieved an area under the receiver operating characteristic curve of 98%. Of the Image Biomarker Standardization Initiative features, texture and first-order intensity features were most contributory across the molecular subgroups.

Conclusion: An MRI-based machine learning decision path allowed identification of the four clinically relevant molecular pediatric medulloblastoma subgroups.

© RSNA, 2022

Online supplemental material is available for this article.

Abbreviations

AUC = area under the receiver operating characteristic curve, IBSI = Image Biomarker Standardization Initiative, MB = medulloblastoma

Summary

MRI-based machine learning designed with sequential decision steps reliably differentiated four unique medulloblastoma molecular subgroups.

Key Results

- A sequential two-stage radiomics classifier trained on MRI scans of 263 patients from a multinational cohort of 12 pediatric centers achieved an F1 score (a measure of accuracy) of 88% to identify the four medulloblastoma (MB) molecular subgroups.
- An MRI radiogenomics approach reliably identified a therapeutically relevant, low-risk wingless (*WNT*) MB molecular subgroup (F1 score, 95% for *WNT* MB).
- Machine learning uncovered quantitative MRI features within the sonic hedgehog (*SHH*) MB subgroup, such as the high-risk, childhood *SHH*.

MRI predictors of medulloblastoma (MB) molecular subgroup offer a noninvasive, presurgical path toward MB risk stratification (1–5). Molecular subgroup status confers prognosis and potentially a key to more precise, tailored therapy (6). Studies have also shown incidence of perioperative complications (eg, cerebellar mutism), and the value of extent of surgical resection is highly subgroup dependent, highlighting a need for reliable, presurgical MB-subgroup prediction (7–9). However, molecular testing is unavailable outside major cancer centers, which limits efficient triage of trial candidates and pursuit of optimal treatment strategies (9–12). Moreover, the upcoming revised World Health Organization classification of central nervous system tumors that will mandate robust molecular subgrouping presents a considerable challenge in many clinical pathology laboratories, particularly those in under-resourced geographic centers.

Although rarest among MB, wingless (*WNT*) mutations encounter the best outcomes and, thus, are most likely to benefit from therapy modification, including lower risk surgery or exclusion of cranial irradiation associated with long-term cognitive dysfunction. Unfortunately, immunohistochemistry alone may be insufficient for *WNT* characterization, and more advanced methods (eg, *CTNNB1* exon 3 sequencing, DNA methylation, gene profiling) can be costly or unavailable (10,11). Among the four sonic hedgehog (*SHH*) molecular subtypes (*SHH*- α , *SHH*- β , *SHH*- γ , *SHH*- δ), adult *SHH* mutations (*SHH*- δ) generally show favorable outcomes (6). However, among infantile *SHH* (*SHH*- β /*SHH*- γ), *SHH*- γ might not require intensified therapy, compared with an older childhood *SHH*- α that often harbors a high-risk *TP53* mutation (6). Finally, group 3 and group 4 tumors are often provisionally bundled as non-*WNT* and non-*SHH* MB due to lack of prevalent identifying driver mutations and require advanced processing with clustering, gene expression, and DNA methylation profiling (6,13,14). Unfortunately, group 3 tumors confer the worst prognosis and are prime targets for clinical trials and alternative therapies (15).

Machine learning offers an opportunity to mine high-dimensional image features and uncover quantitative features that aid precision analytics. Prior studies that piloted MB radiogenomics approaches found limited success and generalizability

largely due to small sample sizes across a single or a few centers (4,5). Recently described Image Biomarker Standardization Initiative (IBSI)-based radiomics features enable quantitative characterization of tumors, for example: shape (tumor size, volume, surface area, sphericity), first order statistics (distribution of voxel intensities), or texture (gray-level distribution within an image) in a standardized and replicable manner (16,17).

Herein, we assembled a large pediatric MB cohort across 12 major centers, including in the United States, Canada, and the United Kingdom, to develop IBSI-based radiogenomics strategies that identify the four clinically significant MB molecular subgroups.

Materials and Methods

Study Population

We conducted a multicenter (Table E1 [online]) retrospective study after institutional review board approval (number 51059). Our study was compliant with the Health Insurance Portability and Accountability Act of 1996. Consent was waived due to minimal risk to participants. The inclusion criteria were as follows: consecutive patients with pathologically confirmed MB from July 1997 to May 2020; age 19 years or less at diagnosis; preoperative MRI scans with both axial contrast-enhanced T1-weighted and T2-weighted sequences; and molecular subgroup analysis. Patients with nondiagnostic MRI scans were excluded. A subset of the participants was included in prior qualitative and non-IBSI-based MRI computational analysis, distinct from the present study (4,18,19).

Imaging Techniques

A brain MRI was performed with at 1.5- or 3-T magnet (GE Healthcare, Siemens Healthineers, Philips Healthcare, and Toshiba Canon Medical Systems USA). The T2-weighted MRI parameters were as follows: turbo spin-echo constant level appearance and sensitivity encoding, fast spin echo, periodically rotated overlapping parallel lines with enhanced reconstruction, syngo BLADE (Siemens Healthineers), driven equilibrium radio frequency reset pulse, and a section thickness of 0.8–5 mm. T1-weighted MRI comprised either three-dimensional isovolumetric or two-dimensional spin-echo imaging, with a section thickness of 0.8–5 mm.

Molecular Analysis

The four MB molecular subgroups (*WNT*, *SHH*, group 3, group 4) were determined with fluorescence in situ hybridization, specialized testing (including RNA methylation array and DNA methylation array), and next-generation sequencing panels (13,20–22).

Immunohistochemistry of Group 3 and Group 4

Given the sparse vascular information on group 3 and group 4 (14,15,23), and to gain insight into potential correlative image phenotypes, we conducted immunohistochemistry on eight random formalin-fixed specimens for endothelial cells (ETS-related gene antibodies, 1:1000; Abcam; product #ab92513) and tight junction proteins (Claudin-5 antibodies, 1:500; Thermo Fisher

Scientific; product #34–1600). We calculated staining index as the product of intensity and fraction of positive tumor cells by using Aperio ImageScope (Leica) (24).

Feature Extraction and Reduction

Tumor volume was delineated on MRI scans with consensus review by board-certified neuroradiologists (K.W.Y. and A.J., with >10 years of experience). We extracted 1800 IBSI-based (16,17) features using PyRadiomics (2.2.0.post7+gac7458e) within the quantitative image feature pipeline (Appendices E1, E2 [online]) (25). We applied z -score intensity normalization before feature extraction, as it improves robustness of MRI-based radiomics features and classification (26). Extracted features underwent sparse regression analysis by using a least absolute shrinkage and selection operator on RStudio software (version 1.2.5033) (Appendix E2 [online]).

Primary, Two-Stage Multiclass Classifier

We developed a two-stage model, with each step comprising a binary classifier for three classes (*WNT*, *SHH*, non-*WNT* and non-*SHH*). Given the molecular overlap, group 3 and group 4 were combined as non-*WNT* and non-*SHH* (hereafter, group 3/4). The first stage discriminated *WNT/SHH* from group 3/4. We then built a second binary classifier to distinguish *WNT* from *SHH*. For each stage, we conducted feature reduction for

each paired subgroup. Using the corresponding reduced-feature set, we identified the best-performing algorithm among six candidate classifiers (support vector machine, logistic regression, k -nearest neighbor, random forest, extreme gradient boosting, and neural network) at each stage (Appendix E2 [online]). Training and test sets were randomly allocated from the total cohort in a 75:25 ratio. The training cohort underwent resampling to correct for sample imbalance. Optimal classifier parameters were estimated with a grid search (Table E2 [online]). Relative influences of imaging features were calculated with logistic regression based on coefficients used in the weighted sum.

In the combined, final model, a holdout test set underwent classification in the first stage whereby its outputs were subsequently passed onto a second-stage classifier. Figure 1 illustrates the staged classifier configuration. Overall performance was assessed for the combined stages. The final radiomics multiclass classifier was guided by maximizing the F1 score, measured as the weighted average between the precision (positive predictive value) and recall (sensitivity). The F1 score is also known as the Dice similarity coefficient.

Single-Stage Multiclass Classifier Model

To compare the performance of our primary model to that of a simpler single-stage model, we used the same six candidate classifiers to perform a multiclass classification of *WNT*, *SHH*, and

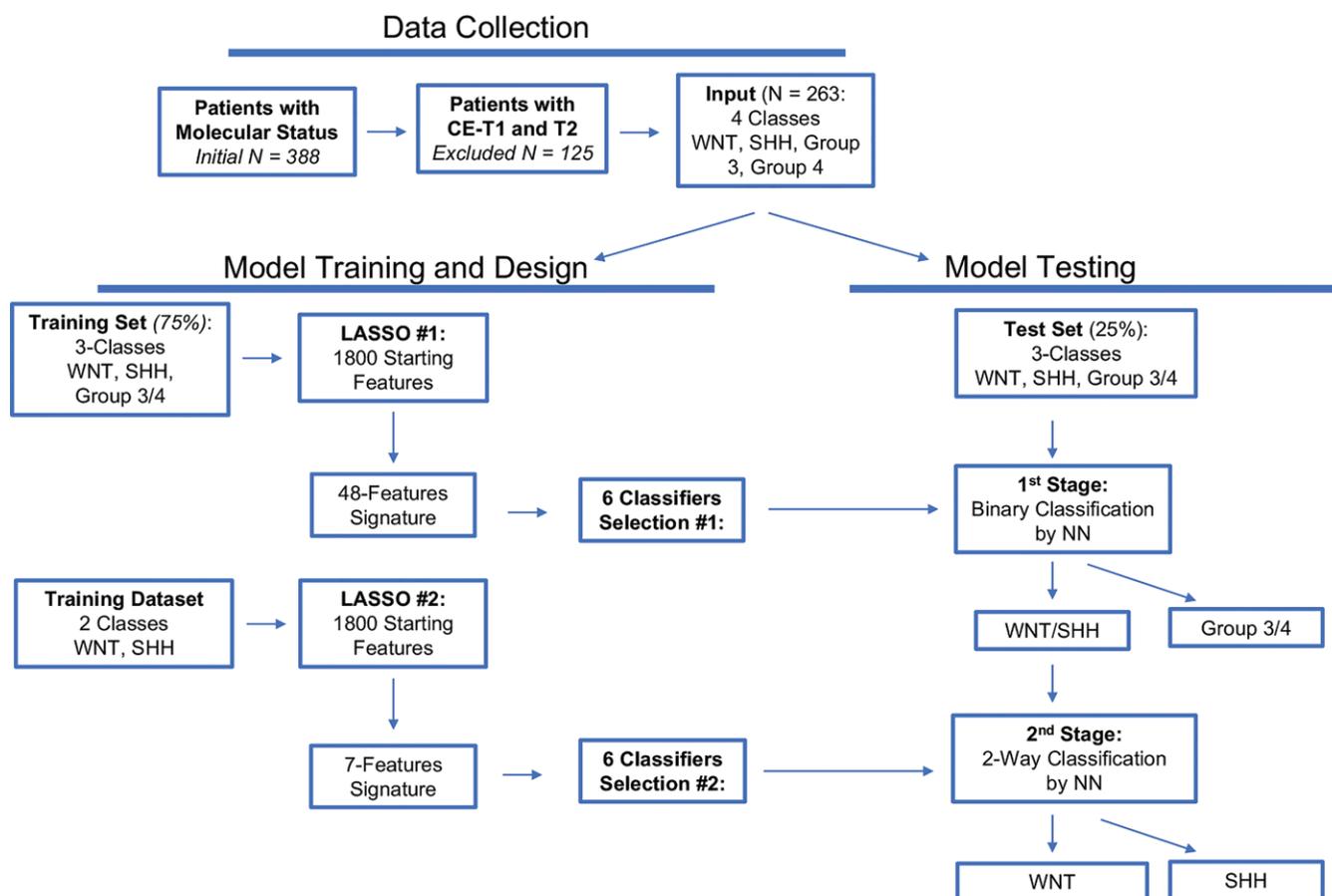


Figure 1: Flowchart shows workflow for training and testing of a two-stage classifier. Each stage consists of a binary classifier optimized for its own respective reduced-feature set obtained by sparse regression analyses. The first stage passes the subgroup composed of wingless (*WNT*) and sonic hedgehog (*SHH*) to the second stage for further separation. CE = contrast enhanced, LASSO = least absolute shrinkage and selection operator, NN = neural network.

group 3/4 MB. We performed a single feature-reduction step for the entire cohort and for model training and testing.

Subgroup Classifier Models

Due to ambiguities in molecular diagnosis, not all non-*WNT* and non-*SHH* tumors were eligible for group 3 versus group 4 classifier training. Thus, in the sequential model, we had classified these observations under the group 3/4 label. To explore the possibility of discriminating group 3 from group 4, we constructed an independent binary classifier using a methylation-confirmed reduced cohort (group 3 and group 4, $n = 48$ and $n = 64$, respectively). Furthermore, we conducted an independent binary analysis to distinguish infantile (age, ≤ 60 months) versus childhood (age, >60 months) *SHH* ($n = 50$ and $n = 33$, respectively).

Statistical Analysis

$P < .05$ was considered significant. A statistics expert (S.W.W., with 2 years of experience) analyzed performance metrics. CIs were obtained by bootstrapping the test sets for 2000 random samples. Model accuracy score was compared with no-information rate (prevalence rate) using the Wald test. Classifier development was performed using Python 3.8.5; feature reduction and statistics were calculated with RStudio, version

1.2.5033. The F1 score was calculated as the weighted average between the precision score, positive predictive value, and recall score (sensitivity).

Code Availability

Code for data processing and analysis is provided at https://github.com/sandymule/tumor-classification/tree/master/MB_Molecl_final_for_submission.

Results

Patient Cohort

A total of 263 patients (mean age at diagnosis, 87 months \pm 60 [SD]; 166 boys) met the inclusion criteria. Twenty-six patients (9.9%) had *WNT* MB, 83 (31.6%) had *SHH* MB, and 154 (58.6%) had group 3/4 MB (Table 1); mean ages at diagnosis were 121, 75, and 88 months, respectively. Table 1 summarizes patient demographics and molecular subgroup distribution.

Model Development

Tables 2 and 3 summarize model performances. Figure 2 illustrates sample probability outputs of the holdout test sets on the four MB subgroups. Table E2 (online) identifies the top retained features that contributed to predictive modeling within each classifier, including definitions and qualitative interpretations.

First-Stage Classifier Model: *WNT* and/or *SHH* versus Group 3/4 MB

In the first stage, least absolute shrinkage and selection operator regression identified 48 relevant IBSI-based radiomic features (Table E3 [online]), with two clinical variables, 15 from T1-weighted MRI and 31 from T2-weighted MRI, including one shape, seven first-order, and 29 textural (10 gray-level co-occurrence matrix, 14 gray-level zone size matrix, and five gray-level run length matrix) features. Among the six classifier models, neural network exhibited the best performance (F1 score, 0.90) (Table E4 [online]). Sensitivity, specificity, accuracy, and area under the receiver operating characteristic curve (AUC) of the neural network were 88% (37 of 42 patients), 88% (21 of 24 patients), 88% (58 of 66 patients; CI: 79, 95), and 96% (range, 89%–99%), respectively.

Second-Stage Classifier Model: *WNT* versus *SHH* MB

The second least absolute shrinkage and selection operator regression for *WNT* versus *SHH* classification identified seven features. These features entailed one clinical feature, three from T1-weighted MRI and three from T2-weighted MRI, including two first-order and four textural (two gray-level co-occurrence matrices and two gray-level run length matrices) features. Among the six classifier models, the neural network showed the highest performance (F1 score, 0.96 [CI: 86, 100]) (Table E4 [online]), with sensitivity, specificity, accuracy, and AUC of 96% (22 of 23 patients), 80%

Table 1: Comparison of Clinical Features in Patients with *WNT*, *SHH*, and Group 3/4 Medulloblastoma

Characteristic	<i>WNT</i>	<i>SHH</i>	Group 3/4	<i>P</i> Value
No. of patients	26 (9.9)	83 (31.6)	154 (58.6)	
Age*				
Mean	121	75	88	.002
Median	110	47	82	
Range	0–401	0–252	3–246	
SD	78	65	51	
Male sex	10 (38.5)	42 (50.6)	114 (74.0)	<.001
Female sex	16 (61.5)	41 (49.4)	40 (26.0)	

Note.—Data are numbers of patients with percentages in parentheses, unless otherwise stated. *SHH* = sonic hedgehog, *WNT* = wingless.

* Age is in months at diagnosis.

Table 2: Binarized and Summated Performance Metrics of the Final Sequential Classifier

Positive Class Cohort	Sensitivity	Specificity	F1 Score	Accuracy	NIR
Binary					
Group 3/4	88	88	90 (82, 96)	88 (79, 96)	64
<i>SHH</i>	86	90	77 (56, 91)	89 (80, 97)	21
<i>WNT</i>	90	100	95 (80, 100)	98 (95, 100)	15
Multiclass					
Micro-average	88	94	88 (76, 96)	92 (80, 97)	...

Note.—Data are percentages, and data in parentheses are 95% CIs. The final sequential classifier uses a staged neural network binary classifier for wingless (*WNT*) and sonic hedgehog (*SHH*) versus group 3/4, followed by a neural network binary classifier for *WNT* versus *SHH*. NIR = no-information rate.

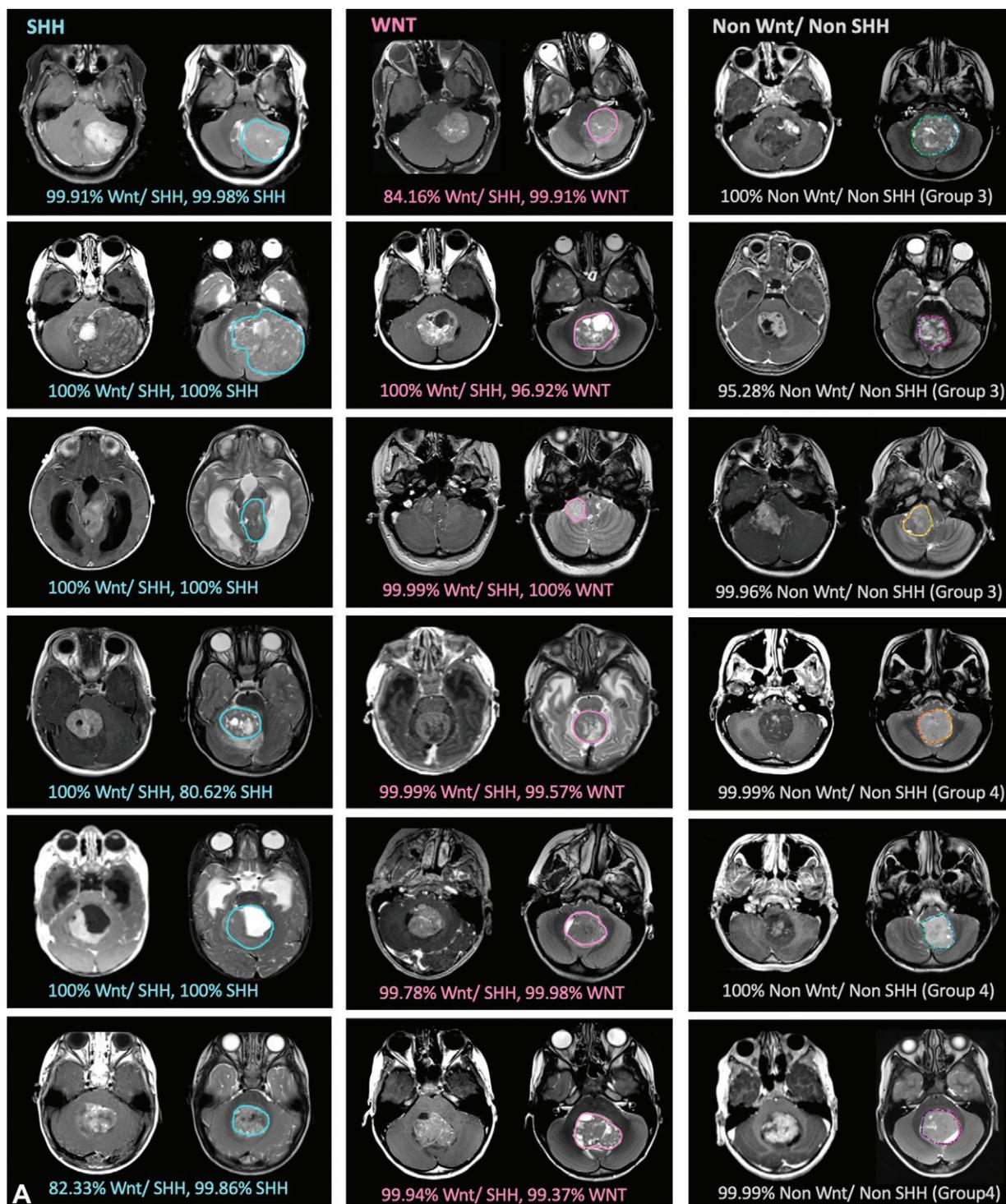


Figure 2: Examples of probability output on contrast-enhanced T1-weighted (left) and T2-weighted (right) MRI scans from the medulloblastoma (MB) test subset that did not participate in the model development. **(A)** Results of a staged primary classifier model are shown with probability outputs of non-wingless (*WNT*) and non-sonic hedgehog (*SHH*) and subsequent outputs of *WNT* and *SHH* generated from *WNT* and *SHH* (Fig 2 continues).

(four of five patients), 93% (26 of 28 patients; CI: 79, 100), 95% (CI: 80, 100), respectively. The top three relevant features as defined by IBSI included T1-Correlation (gray-level co-occurrence matrix), T2-InterquartileRange (first-order intensity), and T2-Kurtosis (first-order intensity) (Fig 3, Table E2 [online]) (16,17). Figure 4 illustrates MRI examples of qualitative features of *WNT* and *SHH* tumors.

Combined Sequential Model

Finally, the two classifiers were performed sequentially such that the output of the initial neural network classifier containing the combined *WNT* and *SHH* group was fed to the second neural network classifier (Table 3). The metrics for the combined model were as follows: sensitivity, 88% (58 of 66 patients); specificity, 94% (124 of 132 patients); accuracy, 92% (182 of 198 patients);

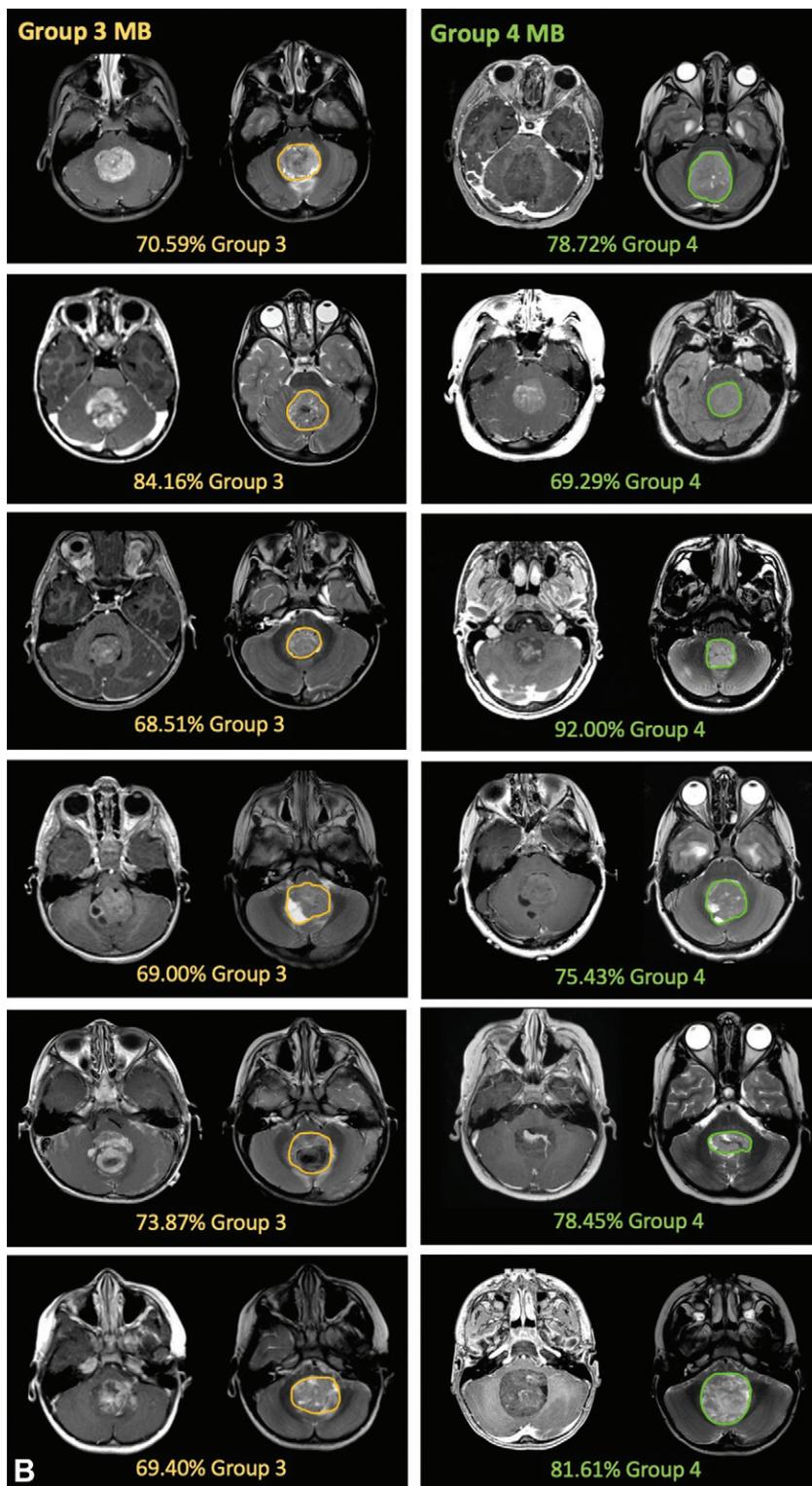


Figure 2: (continued) (B) Examples of tumors from an independent binary classifier model that differentiates between group 3 and group 4 are shown.

80%–97%); and F1 score, 88% (76%–96%). Accuracy for the final model was compared with the no-information rate and was better than random guessing for all subgroups ($P < .001$, Wald test). The sequential model outperformed a single-stage model trained to classify among *WNT*, *SHH*, and group 3/4 simultaneously (Table E5 [online]).

Subgroup Classifier Models

For infantile versus childhood *SHH* analysis, feature reduction identified 15 features (nine from T1-weighted MRI and six from T2-weighted MRI). The random forest algorithm achieved the highest performance, yielding a sensitivity of 83% (five of six patients), specificity of 93% (14 of 15 patients), accuracy of 90% (19 of 21 patients; CI: 76, 100), and AUC of 89% (CI: 47, 99) (Table 3). Predictive features for distinguishing between infantile *SHH* versus childhood *SHH* included T1-Gray Level Nonuniformity, Normalized; T1-Zone Entropy; and T2-Inverse Difference Moment, Normalized (Figure E1 [online], Table E2 [online]).

For group 3 versus group 4 analysis, feature reduction identified 16 features (one clinical variable, nine from T1-weighted MRI and six from T2-weighted MRI), including one shape, seven first order, and seven textural (three gray-level co-occurrence matrix, three gray-level size zone matrix, and one gray-level run length matrix) features. An extreme gradient boosting classifier produced the highest metrics, with a sensitivity of 93% (14 of 15 patients), specificity of 92% (12 of 13 patients), accuracy of 93% (26 of 28 patients; CI: 82, 100), and AUC of 98% (CI: 85, 100) (Table 3). Important predictors for distinguishing group 3 and group 4 included T1- and T2-Mean Voxel Intensity, as well as T1-Run Length Nonuniformity (Fig 5, Table E2 [online]).

Immunohistochemistry Analysis of Group 3 and Group 4 MB

The degree of ETS-related gene endothelial staining was greater for group 4 ($n = 4$) than group 3 ($n = 4$; $P = .03$, Wilcoxon rank-sum) (Fig E2 [online]). There was no significant difference in the Claudin-5 expression between the group 3 and group 4 tumors ($P = .69$). Figure 6 illustrates MRI examples of group 3 and 4 tumors and their correlative vascular histologic markers.

Discussion

Molecular subgroup information is a key to tailored therapy for pediatric medulloblastoma (MB) (12,27). Although molecular testing currently serves as the only validated means of attributing risk, it is not uniformly accessible. A validated image-based option could offer a cost-effective alternative or an orthogonal path to subgroup prediction (12,27). Herein, we applied machine learning on presurgical MRI scans from 12 pediatric centers to construct decision algorithms predictive of the four MB molecular subgroups.

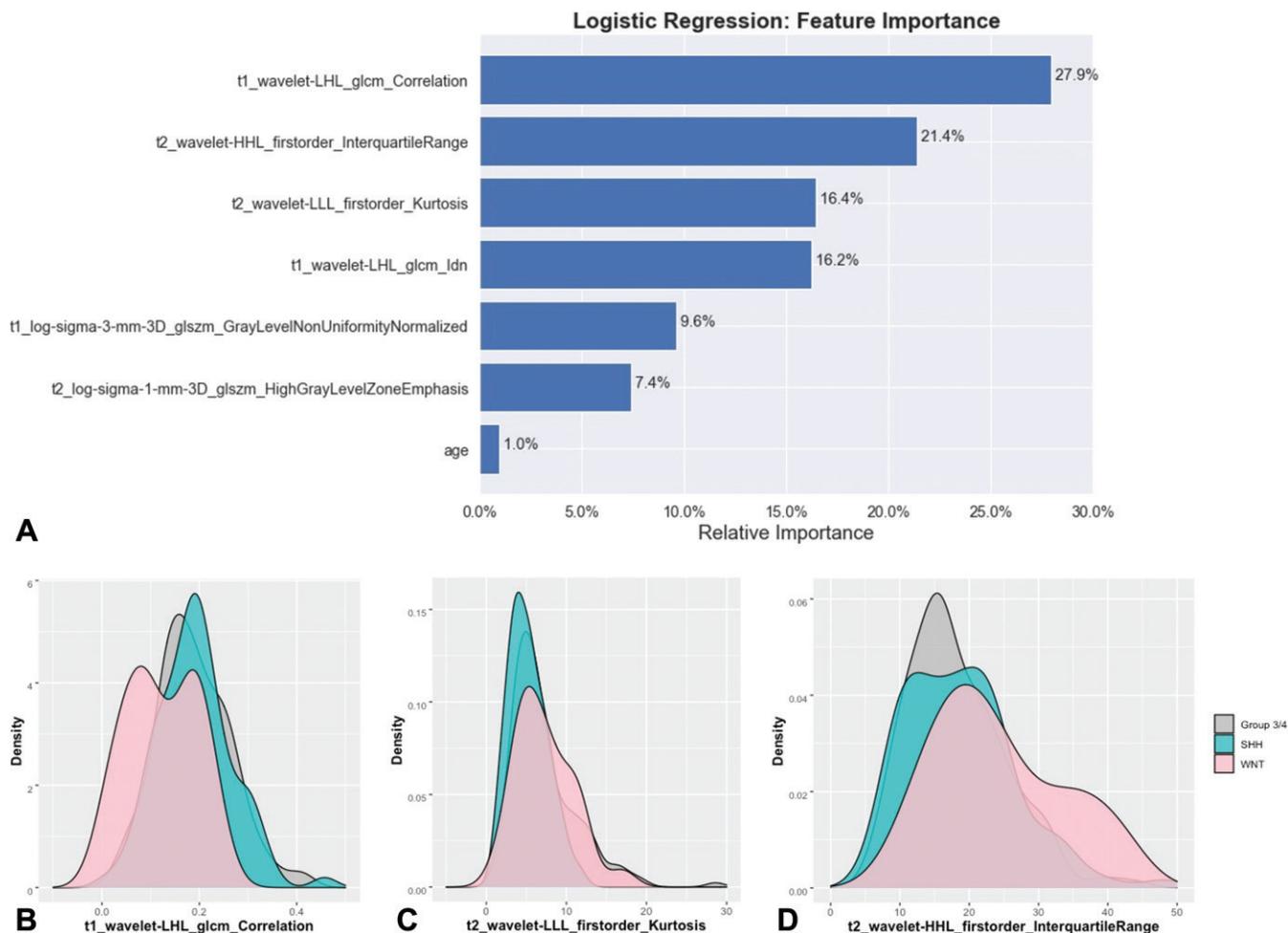


Figure 3: (A) Bar plot shows the relative influence as calculated with logistic regression of the seven reduced features for the second stage, a binary classifier trained to distinguish wingless (*WNT*) from sonic hedgehog (*SHH*) medulloblastoma. (B–D) Density plots of the top three Image Biomarker Standardization Initiative features, including (B) T1-Correlation, (C) T2-Kurtosis, and (D) T2-Interquartile Range.

On the basis of human visual inspection, studies have suggested tumor location and enhancement patterns might infer the MB subgroup, albeit with mixed performances (18,19,28). Since then, a small cohort study by Iv et al (4) described potential for machine learning, despite limited predictability (AUCs, 70%–83%) on a subset of MB subgroups (2,5). Because their pilot work preceded radiomics standardization, the computed features are nonreplicable and incongruous with the current IBSI standards (16,17). More recently, Yan et al (5) proposed a classification scheme better predictive of *WNT* (AUC, 83%) but less robust on other subgroups (AUCs, <70%), while deep-learning approaches suggested by Chen et al (2) lacked feature interpretability and had a lower predictive performance than the present study (4).

Herein, we describe a staged approach that outperforms all earlier approaches, with high predictive performance across all subgroups (2,4,5). Using the largest MB cohort to date, we combined multiple radiomic models in a clinically rational design that optimizes the classification performance at a level consistent with modern advanced genomic testing (13). We achieved this by prioritizing the less aggressive *WNT* identification, as this group could benefit most from less aggressive

surgery or reduced neurotoxic therapies. We also distinguished high-risk group 3 from group 4 and identified potential features that underlie a high-risk, childhood *SHH* subgroup, that is, the α -*SHH* group.

We found that intensity-based IBSI features contributed to the differentiation between *WNT* and *SHH*. For example, T1-Correlation, a global measure of homogeneity, is greater for *SHH* but lower for *WNT*. Similarly, Patay et al (29) described more heterogeneous contrast enhancement of *WNT* compared with *SHH* (19). *WNT* tumors often manifest vascular fragility with blood-brain barrier breakdown, which might also render tumor intensities more heterogeneous (14,15,23). Prior studies have suggested that group 4 tumors enhance less in volume than group 3, regardless of degree of T1-shortening (18,19,28). Counterintuitively, the average gray values (T1- and T2-Mean) were higher for group 4 than group 3. Higher vascular density, as suggested by ETS-related gene staining, of group 4 might create a tumor environment with higher T1-Mean Voxel Intensity, possibly from paramagnetic iron within densely packed blood volume. In our proxy classifier of infantile versus childhood *SHH*, infantile *SHH* showed global heterogeneity (T1-Zone Entropy), particularly with clusters of related voxels (T1-Size

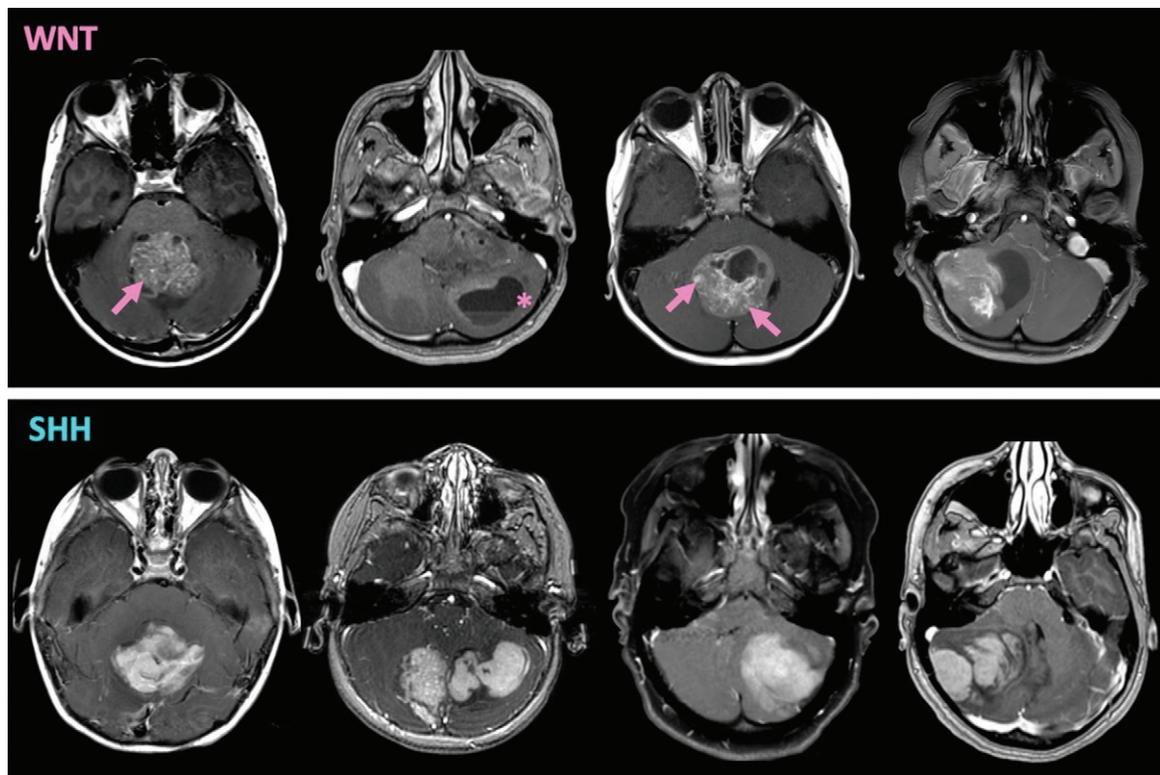


Figure 4: Examples of axial contrast-enhanced T1 MRI scans of wingless (*WNT*) and sonic hedgehog (*SHH*) medulloblastoma. T1-Correlation, a global measure of homogeneity, was greater for *SHH* on contrast-enhanced T1-weighted MRI scans. At a macroscopic level, *SHH* tumors appear to have more homogeneous distribution of high signal intensity across pixels on T1-weighted MRI scans compared with more heterogeneous enhancement of *WNT* tumors, which might relate to higher vascular fragility and associated hemorrhagic components. Note that hemorrhagic fluid-level (*) and stippled and curvilinear foci of enhancement (arrows) are seen in the patient with *WNT* tumor.

Table 3: Binarized and Summated Performance Metrics of the Six Classifiers for Group 3 versus Group 4 and Infantile versus Childhood *SHH* Classifiers

Independent Binary Task and Classifier	Sensitivity	Specificity	Accuracy	F1 Score	AUC
Group 3 versus group 4					
Support vector machine	73	92	82	82	95
Logistic regression	93	92	93	93	93
K-nearest neighbor	73	77	75	76	87
Random forest	73	85	79	79	91
eXtreme gradient boost	93	92	93	93	98
Neural network	80	85	82	83	92
<i>SHH</i> infantile versus childhood					
Support vector machine	33	100	81	50	77
Logistic regression	50	93	81	60	90
K-nearest neighbor	50	93	81	60	89
Random forest	83	93	91	83	89
eXtreme gradient boost	17	87	67	22	68
Neural network	50	100	86	67	89

Note.—Data are percentages. Performance metrics for these binary classifiers are based on the holdout test. AUC = area under the receiver operating curve, *SHH* = sonic hedgehog.

Zone Nonuniformity, Normalized) that might reflect nodular architecture known to occur with *SHH*- γ (6).

There are many clinical implications. Lesions that lateralize to cerebellar peduncle, and, thus, are at higher risk of mutism, may

benefit from biopsy as *WNT* tumors may respond to chemoradiation alone (9,18). The availability of preoperative image phenotypes predictive of *WNT* could potentially offer a surgeon the option of less invasive or lower risk surgery. Preoperative

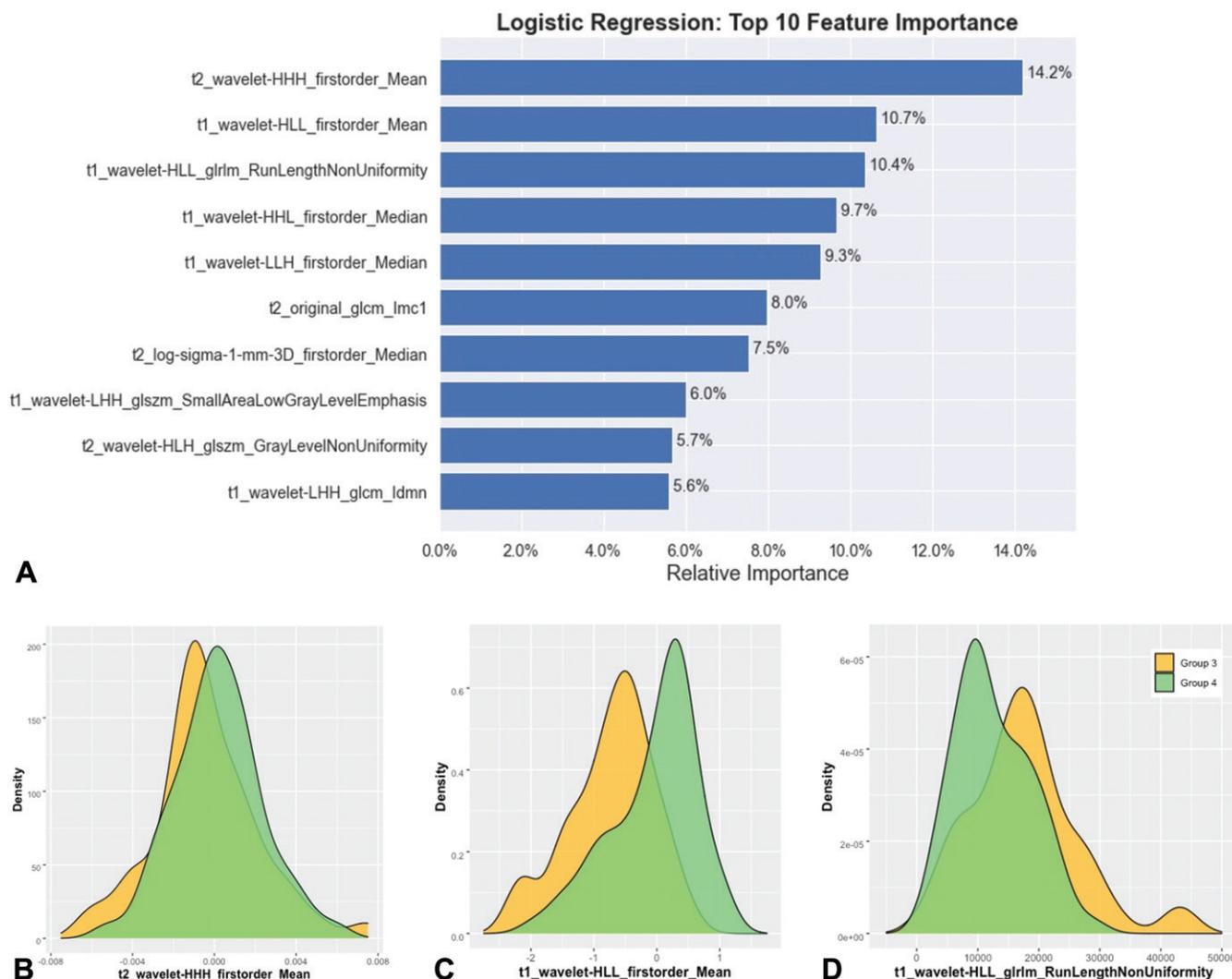


Figure 5: (A) Bar plot shows the relative influence as calculated by logistic regression of the top 10 reduced features for the follow-up binary classifier trained to distinguish group 3 from group 4. (B–D) Density plots of the top three Image Biomarker Standardization Initiative features, including (B) T2-Mean, (C) T1-Mean, and (D) T1-Run Length Nonuniformity.

differentiation of group 3 versus group 4 might also assist therapy planning. For example, group 4 tumors, typically more prone to cerebellar mutism, might qualify for neoadjuvant chemotherapy that first aims to decrease tumor size before safer, delayed surgical resection (7,8,30). Infantile *SHH* subgroup may also qualify for bypassing radiation therapy compared with higher risk *SHH- α* (31–33). Thus, at ages where there is considerable overlap between these groups, a radiomic classifier might serve as an adjunct to methylation arrays.

We report several limitations. First, a larger sample size would be desirable. Given the lack of public MB image data sets, we compiled the largest MB cohort from 12 pediatric institutions to conduct machine learning on real-world images, thereby extending model generalizability. Also, MRI scans were obtained from heterogeneous protocols and scanners without phantom use. To mitigate variability in signal intensities across acquisition parameters, we performed *z*-score normalization before feature extraction. While we report results of presurgical imaging, future iterations that incorporate diffusion, MRI

fingerprinting, or perfusion might bolster model performance (18,19). We extracted radiomics features from isolated tumor volume and, thus, did not incorporate tumor-brain spatial relationships. Future studies should consider combining radiomics and deep learning to assimilate both tumor and global brain spatial features.

In conclusion, we present an MRI-based machine learning decision path predictive of the four clinically relevant molecular pediatric medulloblastoma subgroups. We describe contributory Image Biomarker Standardization Initiative–based radiomics signatures across the subgroups for artificial intelligence transparency and to gain insight into tumor phenotypes. While tumor diagnosis will continue to rely on tissue specimens, a validated machine learning option, including future imaging genomics investigations that combine prospective model developments and deployment, could offer a global, cost-effective alternative or adjunct to molecular-based risk assessment and widen future opportunities for risk-tailored therapies and trial design.

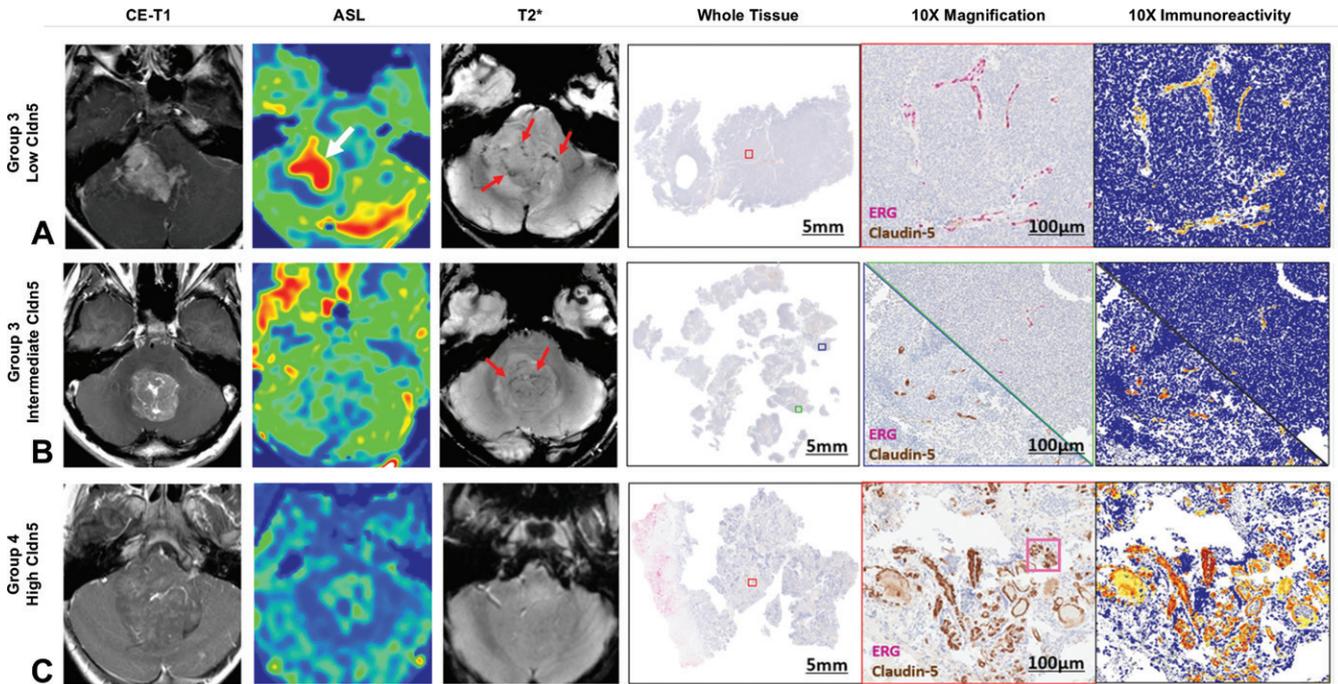


Figure 6: Contrast-enhanced (CE) T1-weighted, arterial spin labeling (ASL), and T2*-weighted MRI features of medulloblastoma and corresponding vascular immunohistochemistry. The ETS-related gene (ERG) antibody (pink) stains are for nuclei of endothelial cells, which line the interior surface of blood vessels and are a marker of vascular density. The Claudin-5 (Cldn5) antibody (brown) stains are for endothelial tight junction protein and mark blood-brain barrier integrity. **(A)** Images in a patient with group 3 right cerebellopontine mass. On T2*-weighted image, the mass shows heterogeneous enhancement with irregular T2* foci of blood products or deoxyhemoglobin of tumor vascularity (red arrows). Corresponding ASL image shows high perfusion (white arrow). Diffuse loss of Claudin-5 expression (lack of brown staining) along endothelial cells (pink ERG staining) suggests blood-brain barrier breakdown. **(B)** Images in a different patient with group 3 tumor show faint foci (red arrows) on T2*-weighted image, intermediate perfusion, and heterogeneous distribution of Claudin-5 expression, depending on the tumor specimen location (blue and green boxes). **(C)** Images in patient with group 4 tumor show mild, patchy enhancement and intermediate perfusion. No discrete blood products are present on T2*-weighted MRI scan. Diffuse Claudin-5 expression is seen associated with abundant endothelial cells (pink staining behind the brown staining) within the tumor, suggesting high vascular density associated with preserved blood-brain barrier integrity.

Author contributions: Guarantors of integrity of entire study, M.Z., M.H., H.L., N.A.V., K.W.Y.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, M.Z., S.W.W., M.W.W., K.S., S. Lummus, A.E., S. Laughlin, M.T., N.A.V., S. Pfister, A.J., K.W.Y.; clinical studies, S.S.A., K.S., S. Lummus, H.L., A.E., A.R., J.N., S.H., S. Laughlin, B.E.W., C.Y.H., K.M., T.S.J., K.A., N.A.V., G.A.G., A.J., K.W.Y.; experimental studies, M.H., L.T.T., Q.Z., S. Lummus, A.R., H.V., S.H.C., K.W.Y.; statistical analysis, M.Z., S.W.W., M.H., K.W.Y.; and manuscript editing, M.Z., S.W.W., J.N.W., M.W.W., L.T.T., Q.Z., S.S.A., K.S., S. Lummus, H.L., A.E., A.R., S.H., S. Laughlin, S. Perreault, R.M.L., B.E.W., C.Y.H., K.M., H.V., S.H.C., T.S.J., K.A., P.G.F., N.A.V., G.A.G., S. Pfister, E.T., A.J., V.R., K.W.Y.

Disclosures of conflicts of interest: M.Z. No relevant relationships. S.W.W. No relevant relationships. J.N.W. No relevant relationships. M.W.W. No relevant relationships. S.T. No relevant relationships. M.H. No relevant relationships. L.T. No relevant relationships. Q.Z. No relevant relationships. S.S.A. No relevant relationships. K.S. No relevant relationships. S. Lummus No relevant relationships. H.L. No relevant relationships. A.E. No relevant relationships. A.R. No relevant relationships. J.N. No relevant relationships. S.H. No relevant relationships. M.M. No relevant relationships. S. Laughlin No relevant relationships. S. Perreault No relevant relationships. K.R.M.B. No relevant relationships. R.M.L. No relevant relationships. Y.J.C. Hyundai Hope on Wheels research grant from the Ericksen Family Endowed Professorship. B.E.W. Deputy editor of *Radiology*; spouse owns stock in Siemens; spouse is employee of Siemens Healthineers. C.Y.H. No relevant relationships. K.M. European Course in Paediatric Neuroradiology lecture honorarium; Guerbet lecturer honorarium; Siemens lecturer honorarium; payment for expert medicolegal reports in the United Kingdom. H.V. Payment for deposition testimony from Miller Weisbrod; participant on a DataSafety Monitoring Board or Advisory Board for Stanford Neuropathology. S.H.C. No relevant relationships. T.S.J. Payment to institution from Cancer Research UK, The Brain Tumour Charity, Children with

Cancer UK, NIHR, Great Ormond Street Hospital Children's Charity, Olivia Hodson Cancer Fund; royalties for book publishing from Elsevier; honoraria for editing from Wiley; payment for lecture to an educational event organized by Bayer; payment via Neuropath Ltd, a company owned and operated by the author and their wife, for expert witness work in the courts in the UK and Ireland for expert neuropathology work, mostly in the case of unexpected child deaths for the HM courts; Editor in Chief for *Neuropathology and Applied Neurobiology* (paid a share of the profits of the journal); Lead for the Childhood Solid Tumour Domain of the Genomics England Clinical Interpretation Partnership; shareholder in Repath and Neuropath. K.A. No relevant relationships. P.G.F. NIH/NCI for co-investigator in Pediatric Brain Tumor Consortium; NIN/NHGRI for co-principal investigator in Undiagnosed Diseases Network; payment for role as Associate Editor for *The Journal of Pediatrics* by Elsevier Publishing; personal stock holdings in Johnson & Johnson. M.T. No relevant relationships. T.P. NIH funded Pediatric Brain Tumor Consortium Neuroimaging Center grant; payments from Springer Publishing; President of the American Society of Neuroradiology. N.A.V. No relevant relationships. G.A.G. Participant on a DataSafety Monitoring Board or Advisory Board for Stanford University. S. Pfister Grants from EU (IMI-2, ERC), BMBF, Brain Tumor Charity, DFG, Deutsche Kinderkrebsstiftung, Deutsche Krebshilfe; three patents in DNA methylation-based tumor classification; leadership or fiduciary role in SAB Princess Maxima Center, SAB INCA, and SAB BioSkryb. E.T. Grants or contracts from FDA, Cure Starts Now, American Brain Tumor Foundation, NIH; consulting fees for being a scientific advisor for Oncoheroes Biosciences; support for attending meetings and/or travel for Cure Starts Now; provisional patent on a drill for craniostylosis surgery; participant on a DataSafety Monitoring Board or Advisory Board for University of Alabama for a Phase I clinical trial of malignant pediatric brain tumors. A.J. No relevant relationships. V.R. No relevant relationships. K.W.Y. Participant on a DataSafety Monitoring Board or Advisory Board for Stanford University.

References

1. Attallah O. MB-AI-His: Histopathological Diagnosis of Pediatric Medulloblastoma and its Subtypes via AI. *Diagnostics* (Basel) 2021;11(2):359.

2. Chen X, Fan Z, Li KKW, et al. Molecular subgrouping of medulloblastoma based on few-shot learning of multitasking using conventional MR images: a retrospective multicenter study. *Neurooncol Adv* 2020;2(1):vdaa079.
3. Das D, Mahanta LB, Ahmed S, Baishya BK. Classification of childhood medulloblastoma into WHO-defined multiple subtypes based on textural analysis. *J Microsc* 2020;279(1):26–38.
4. Iv M, Zhou M, Shpanskaya K, et al. MR Imaging-Based Radiomic Signatures of Distinct Molecular Subgroups of Medulloblastoma. *AJNR Am J Neuroradiol* 2019;40(1):154–161.
5. Yan J, Liu L, Wang W, et al. Radiomic Features From Multi-Parameter MRI Combined With Clinical Parameters Predict Molecular Subgroups in Patients With Medulloblastoma. *Front Oncol* 2020;10:558162.
6. Cavalli FMG, Remke M, Rampasek L, et al. Intertumoral Heterogeneity within Medulloblastoma Subgroups. *Cancer Cell* 2017;31(6):737–754.e6.
7. Jabarkheel R, Amayiri N, Yecies D, et al. Molecular correlates of cerebellar mutism syndrome in medulloblastoma. *Neuro Oncol* 2020;22(2):290–297.
8. Ramaswamy V, Remke M, Shih D, et al. Duration of the pre-diagnostic interval in medulloblastoma is subgroup dependent. *Pediatr Blood Cancer* 2014;61(7):1190–1194.
9. Thompson EM, Hielscher T, Bouffet E, et al. Prognostic value of medulloblastoma extent of resection after accounting for molecular subgroup: a retrospective integrated clinical and molecular analysis. *Lancet Oncol* 2016;17(4):484–495.
10. Goschzik T, Zur Mühlen A, Kristiansen G, et al. Molecular stratification of medulloblastoma: comparison of histological and genetic methods to detect Wnt activated tumours. *Neuropathol Appl Neurobiol* 2015;41(2):135–144.
11. Pietsch T, Schmidt R, Remke M, et al. Prognostic significance of clinical, histopathological, and molecular characteristics of medulloblastomas in the prospective HIT2000 multicenter clinical trial cohort. *Acta Neuropathol (Berl)* 2014;128(1):137–149.
12. Shuangshoti S, Tadadontip P, Techavichit P, Thorner PS, Shuangshoti S, Teerapakpinyo C. Simplified Molecular Subtyping of Medulloblastoma for Reduced Cost and Improved Turnaround Time. *Appl Immunohistochem Mol Morphol* 2020;28(7):538–543.
13. Northcott PA, Shih DJH, Remke M, et al. Rapid, reliable, and reproducible molecular sub-grouping of clinical medulloblastoma samples. *Acta Neuropathol (Berl)* 2012;123(4):615–626.
14. Ramaswamy V, Remke M, Bouffet E, et al. Risk stratification of childhood medulloblastoma in the molecular era: the current consensus. *Acta Neuropathol (Berl)* 2016;131(6):821–831.
15. Menyhart O, Györfy B. Molecular stratifications, biomarker candidates and new therapeutic options in current medulloblastoma treatment approaches. *Cancer Metastasis Rev* 2020;39(1):211–233.
16. Zwanenburg A, Vallières M, Abdalah MA, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology* 2020;295(2):328–338.
17. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* 2017;77(21):e104–e107.
18. Perreault S, Ramaswamy V, Achrol AS, et al. MRI surrogates for molecular subgroups of medulloblastoma. *AJNR Am J Neuroradiol* 2014;35(7):1263–1269.
19. Mata-Mbemba D, Zapotocky M, Laughlin S, Taylor MD, Ramaswamy V, Raybaud C. MRI Characteristics of Primary Tumors and Metastatic Lesions in Molecular Subgroups of Pediatric Medulloblastoma: A Single-Center Study. *AJNR Am J Neuroradiol* 2018;39(5):949–955.
20. Capper D, Jones DTW, Sill M, et al. DNA methylation-based classification of central nervous system tumours. *Nature* 2018;555(7697):469–474.
21. George SL, Izquierdo E, Campbell J, et al. A tailored molecular profiling programme for children with cancer to identify clinically actionable genetic alterations. *Eur J Cancer* 2019;121:224–235.
22. Schwalbe EC, Hicks D, Rafiee G, et al. Minimal methylation classifier (MIMIC): A novel method for derivation and rapid diagnostic detection of disease-associated DNA methylation signatures. *Sci Rep* 2017;7(1):13421.
23. Phoenix TN, Patmore DM, Boop S, et al. Medulloblastoma Genotype Dictates Blood Brain Barrier Phenotype. *Cancer Cell* 2016;29(4):508–522.
24. Torp SH, Gulati S, Johannessen E, Dalen A. Coexpression of c-erbB 1-4 receptor proteins in human glioblastomas. An immunohistochemical study. *J Exp Clin Cancer Res* 2007;26(3):353–359.
25. Mattonen SA, Gude D, Echegaray S, Bakr S, Rubin DL, Napel S. Quantitative imaging feature pipeline: a web-based tool for utilizing, sharing, and building image-processing pipelines. *J Med Imaging (Bellingham)* 2020;7(4):042803.
26. Carré A, Klausner G, Edjlali M, et al. Standardization of brain MR images across machines and protocols: bridging the gap for MRI-based radiomics. *Sci Rep* 2020;10(1):12340.
27. Suresh SG, Srinivasan A, Scott JX, Rao SM, Chidambaram B, Chandrasekar S. Profile and Outcome of Pediatric Brain Tumors - Experience from a Tertiary Care Pediatric Oncology Unit in South India. *J Pediatr Neurosci* 2017;12(3):237–244.
28. Łastowska M, Jurkiewicz E, Trubicka J, et al. Contrast enhancement pattern predicts poor survival for patients with non-WNT/SHH medulloblastoma tumours. *J Neurooncol* 2015;123(1):65–73.
29. Patay Z, DeSain LA, Hwang SN, Coan A, Li Y, Ellison DW. MR Imaging Characteristics of Wingless-Type-Subgroup Pediatric Medulloblastoma. *AJNR Am J Neuroradiol* 2015;36(12):2386–2393.
30. Guerrini-Rousseau L, Abbas R, Huybrechts S, et al. Role of neoadjuvant chemotherapy in metastatic medulloblastoma: a comparative study in 92 children. *Neuro Oncol* 2020;22(11):1686–1695.
31. Baroni LV, Sampor C, Gonzalez A, et al. Bridging the treatment gap in infant medulloblastoma: molecularly informed outcomes of a globally feasible regimen. *Neuro Oncol* 2020;22(12):1873–1881.
32. Dhall G, O’Neil SH, Ji L, et al. Excellent outcome of young children with nodular desmoplastic medulloblastoma treated on “Head Start” III: a multi-institutional, prospective clinical trial. *Neuro Oncol* 2020;22(12):1862–1872.
33. Mynarek M, von Hoff K, Pietsch T, et al. Nonmetastatic Medulloblastoma of Early Childhood: Results From the Prospective Clinical Trial HIT-2000 and An Extended Validation Cohort. *J Clin Oncol* 2020;38(18):2028–2040.