Supplemental Table 1. Cohort sizes and demographics of patients with embryonal tumors, high-grade glioma (HGG) and ependymoma (EP).

		Embryonal						HGG		EP
		ETMR	HGNET	NB	PB	NOS	ATRT	GBM	AA	
N		7	3	3	7	7	23	75	52	54
% Total		3.03%	1.30%	1.30%	3.03%	3.03%	9.96%	32.47%	22.51%	23.38%
Age	Mean	64.7	138.5	56.2	122.9	72.4	46.1	147.9	124.1	87.3
	SD	55.9	97.4	19.6	88.5	68.2	61.7	79.1	65.2	65.3
Sex	Female	2	3	1	5	1	9	33	30	26
	%	28.6%	100.0%	33.3%	71.4%	14.3%	39.1%	44.0%	57.7%	48.1%
	Male	5	0	2	2	6	14	42	22	28
	%	71.4%	0.0%	66.7%	28.6%	85.7%	60.9%	56.0%	42.3%	51.9%

Supplemental Table 2. Performance metrics of the best classifier for each binary evaluation on the holdout test set. Accuracy is statistically compared to No Information Rate.

	Classifier	**Positive Class**	**Sens**	**Spec**	**PPV**	**NPV**	**AUC**	**Accuracy**	**95% CI**	**NIR**	**p**
Embryonal - HGG	LR	Embryonal	0.8461	0.9062	0.7857	0.9354	0.9807	0.8888	0.8000 - 0.9778	0.7176	< 0.0001***
Embryonal - EP	XGB	Embryonal	0.9285	0.6923	0.7647	0.9	0.8241	0.8148	0.6296 - 0.963	0.5238	0.0012**
EP - HGG	NN	EP	0.8181	0.9428	0.8181	0.9428	0.9558	0.9130	0.8261 - 0.9783	0.7017	< 0.0001***

**p < 0.01, *** p < 0.001

Supplemental Table 3. Listing of contributing institutions by pathology.

		Embryonal							HGG			EP
		ATRT	ETMR	HGNET	NB	PB	NOS		GBM	AA		
Lurie Children's Hospital of Chicago		16	0	3	1	0	5		23	28		25
Children's Hospital of Orange County		1	0	0	1	1	1		0	0		0
Dayton Children's Hospital		1	1	0	0	0	0		5	3		2

Indiana University Riley Hospital for Children		0	0	0	0	0	1		8	1		0
Seattle Children's Hospital		0	0	0	0	0	0		0	0		5
Stanford Lucile Packard Children's Hospital		4	6	0	1	6	0		29	17		15
Intermountain Primary Children's Hospital		1	0	0	0	0	0		10	3		7

Supplemental Table 4a. A list of the variables identified by feature reduction and submitted for model training in each binary pairing.

T1-MRI Features	T2-MRI Features	Demographics
Embryonal vs HGG		
t1_log-sigma-3-mm-3D_glcm_InverseVariance	t2_log-sigma-1-mm-3D_glcm_ClusterShade	Age
t1_original_shape_Flatness	t2_log-sigma-1-mm-3D_glcm_InverseVariance	
t1_wavelet-HHH_glszm_SizeZoneNonUniformityNormalized	t2_log-sigma-3-mm-3D_glszm_LargeAreaHighGrayLevelEmphasis	
t1_wavelet-HHH_glszm_SmallAreaEmphasis	t2_original_glcm_Imc2	
t1_wavelet-HHH_glszm_SmallAreaLowGrayLevelEmphasis	t2_original_shape_Flatness	

t1_wavelet-HLH_firstorder_Mean	t2_wavelet-HHH_glcm_Idn	
t1_wavelet-LHL_firstorder_Median	t2_wavelet-HHH_glszm_ZoneVariance	
t1_wavelet-LHL_glcm_Imc2	t2_wavelet-HHL_firstorder_Median	
t1_wavelet-LLH_glcm_Idn	t2_wavelet-HHL_glcm_MCC	
	t2_wavelet-HHL_glrlm_LongRunHighGrayLevelEmphasis	
	t2_wavelet-HLH_firstorder_Skewness	
	t2_wavelet-LLH_firstorder_Skewness	
	t2_wavelet-LLL_firstorder_Skewness	
Embryonal vs EP		
t1_log-sigma-1-mm-3D_glcm_Imc2	t2_log-sigma-5-mm-3D_firstorder_Median	Age
t1_wavelet-HLL_firstorder_Skewness	t2_wavelet-LLL_firstorder_Kurtosis	
EP vs HGG		
t1_log-sigma-1-mm-3D_glcm_InverseVariance	t2_log-sigma-1-mm-3D_glcm_ClusterShade	Age
t1_log-sigma-3-mm-3D_firstorder_Skewness	t2_log-sigma-5-mm-3D_glszm_GrayLevelNonUniformityNormalized	
t1_log-sigma-3-mm-3D_glcm_InverseVariance	t2_original_glcm_Imc2	
t1_log-sigma-5-mm-3D_glszm_LargeAreaEmphasis	t2_original_glrlm_LongRunHighGrayLevelEmphasis	
t1_original_glrlm_LongRunHighGrayLevelEmphasis	t2_wavelet-HHH_glcm_Idmn	
t1_original_shape_Flatness	t2_wavelet-HHH_glrlm_LongRunLowGrayLevelEmphasis	

t1_wavelet-HHH_glszm_GrayLevelNonUniformityNormalized	t2_wavelet-HHH_glszm_GrayLevelNonUniformityNormalized	
t1_wavelet-HHH_glszm_SmallAreaEmphasis	t2_wavelet-HHH_glszm_SmallAreaEmphasis	
t1_wavelet-HHL_firstorder_Mean	t2_wavelet-HHL_firstorder_Median	
t1_wavelet-HHL_glcm_ClusterShade	t2_wavelet-HHL_firstorder_Skewness	
t1_wavelet-HHL_glszm_ZoneEntropy	t2_wavelet-HHL_glszm_GrayLevelNonUniformityNormalized	
t1_wavelet-HLH_firstorder_Mean	t2_wavelet-HLH_firstorder_Skewness	
t1_wavelet-HLH_glcm_Imc1	t2_wavelet-HLH_glcm_MCC	
t1_wavelet-HLL_glrlm_RunLengthNonUniformity	t2_wavelet-HLH_glszm_SizeZoneNonUniformityNormalized	
t1_wavelet-LHL_glcm_MCC	t2_wavelet-HLL_glcm_Idmn	
t1_wavelet-LLH_firstorder_Kurtosis	t2_wavelet-HLL_glcm_InverseVariance	
	t2_wavelet-HLL_glszm_GrayLevelNonUniformityNormalized	
	t2_wavelet-LLH_firstorder_Skewness	

Supplemental Table 4b. Number of features retained by the final classifier model for each binary pairing.

Embryonal vs HGG	
Age	1

1st order	6
shape	2
glcm	8
glszm	5
glrlm	1
Embryonal vs EP	
1st order	3
shape	0
glcm	1
glszm	0
glrlm	0
EP vs HGG	
1st order	8
shape	1
glcm	11
glszm	10
glrlm	4

Supplemental Table 5. Listing of the top three features for each binary pairing.

Feature	Interpretation	Higher Group
Embryonal-HGG		
age	Age	HGG
t2_log-sigma-1-mm-3D_glcm_ClusterShade	A measure of the skewness and uniformity Higher cluster shade implies greater asymmetry about the mean.	HGG
t1_wavelet-HLH_firstorder_Mean	The Mean gray level intensity within the ROI	Embryonal
Embryonal-EP		
t2_wavelet-LLL_firstorder_Kurtosis	The 'peakedness' of the distribution of values Higher signifies greater mass distribution in tail; lower signifies concentration toward peak/mean	Embryonal
t1_log-sigma-1-mm-3D_glcm_Imc2	The correlation between probability distributions I and j **quantifying the complexity** of the texture Range (0 – 1): value of 0 representing two independent distributions (no mutual information) and **value of 1** representing two **fully dependent/uniform distribution** (maximal mutual information)	EP
t1_wavelet- HLL_firstorder_Skewness	The asymmetry of the distribution about the Mean Positive is longer right tail. → higher Skewness in EP → EP darker	Embryonal

EP-HGG		
t1_wavelet-HLH_firstorder_Mean	The Mean gray level intensity within the ROI	EP
t1_wavelet-HHL_glcm_ClusterShade	A measure of the skewness and uniformity Higher cluster shade implies greater asymmetry about the mean.	EP
t2_wavelet-HLH_glcm_MCC	MCC: complexity of the texture, with range 0 ≤ MCC ≤ 1. Greater MCC is with right shifted probability curve a "brighter (less homogenously gray-toned)"	HGG

Supplemental Table 6a. Area under-the-curve (AUC) of the 6 classifiers trialed in each of the binary classifiers. (Also, below is a brief description the unique aspects of the six models used.)

	micro-averaged AUC		
	Embryonal - HGG	Embryonal - EP	EP - HGG
SVM	0.9	0.78	0.71
LR	0.98	0.81	0.94
KNN	0.75	0.66	0.81
RF	0.92	0.8	0.85
XGB	0.95	0.82	0.84
NN	0.92	0.74	0.96

SVM: Support vector machine models identify an optimal separating line (or hyperplane) between predicted classes.

LR: Logistic regression interprets a generalized linear function such that the outcome variable is interpreted as the probability of given outcomes.

KNN: K-nearest neighbors evaluates the K-training points closest to a given datapoint to predict its classification.

RF: Random Forest aggregates the scoring from multiple decision trees to produce a classification for data points based on features.

XGB: XGBoost is another model of multiple decision trees (learners) and retroactively aims to learn from incorrectly identified datapoints at the potential cost of overfitting.

NN: Neural Networks are constructed with layers of nodes, where each node consists of a linear combination and a non-linear activation function, that collectively yield a final prediction. Excess layers can also lead to overfitting with small datasets.

Supplemental Table 6b. Area under-the-curve (AUC) of the 6 classifiers trialed in a 3-way classifier.

	micro-averaged AUC
	Embryonal - HGG - EP
SVM	0.75
LR	0.71
KNN	0.67
RF	0.7
XGB	0.75
NN	0.77

Supplemental Appendix 1. Configuration files for radiomic feature extraction.


```
setting:
  normalize: true
  normalizeScale: 100
  binWidth: 10
  label: 1
  interpolator: 'sitkBSpline' # This is an enumerated value, here None is not allowed
  resampledPixelSpacing: [1,1,1] # This disables resampling, as it is interpreted as None, to enable it, specify spacing in x, y, z as [x, y, z]
  weightingNorm: # If no value is specified, it is interpreted as None
  geometryTolerance: 0.0001
  correctMask: True

imageType:
  Original: {} # for dictionaries / mappings, None values are not allowed, '{}' is interpreted as an empty dictionary
  LoG: {'sigma': [5,3,1]}
  Wavelet: {}

featureClass:
  shape: ['VoxelVolume',
         'MeshVolume',
         'SurfaceArea',
         'SurfaceVolumeRatio',
```

'Sphericity',

'SphericalDisproportion',

'Maximum3DDiameter',

'Maximum2DDiameterSlice',

'Maximum2DDiameterColumn',

'Maximum2DDiameterRow',

'Elongation',

'Flatness'] # Only enable these shape descriptors (disables redundant Compactness 1 and Compactness 2)

firstorder: [] # specifying an empty list has the same effect as specifying nothing.

glcm: # for lists none values are allowed, in this case, all features are enabled

glrlm:

glszm:

Supplemental Appendix 2. Parameters for image pre-processing, feature extraction and feature reduction.

Image Pre-Processing

Prior to feature extraction, we normalized (normalize scale = 100) and resampled images to isotropic 1-mm voxels. Below is the link regarding the exact method: https://pyradiomics.readthedocs.io/en/latest/radiomics.html#radiomics.imageoperations.normalizeImage.

A bin width of 10 was used for grey-level discretization in both normalized MR images.

Feature Classes

Extracted features classes included *First Order* statistics, 2D/3D *Shape*, Gray Level Co-occurrence Matrix (*GLCM*), Gray Level Run Length Matrix (*GLRLM*), and Gray Level Size Zone Matrix (*GLSZM*).

Image Filters

Features were computed on original, wavelet filtered, and Laplacian of Gaussian (LoG) filtered images. Wavelet filters included high band-pass (H) and low-band pass filters (L) in the x, y, and z direction resulting in 8 different combinations of decompositions.

Feature Reduction by Least Absolute Shrinkage and Selection Operator

Training was performed with 10-fold cross validation and repeated for 1000 cycles. The mean squared error was calculated for 100 lambdas in each cycle or until a minimum was achieved. The optimal lambda was identified as the lowest mean squared error value and used for feature reduction and coefficient calculations. Features represented in ≥80% of the cycles were retained for subsequent classifier optimization.

Supplemental Appendix 3. Final hyperparameters following grid search for six classifiers evaluated binary classifiers.

Classifier	Optimal Algorithm	Parameters
Embryonal – EP	XGB	{'learning_rate': 0.5, 'max_depth': 6}
Embryonal - HGG	LR	{'C': 1, 'penalty': 'l2'}
EP - HGG	NN	{'hidden_layer_sizes': (100, 100, 50), 'learning_rate': 'constant'}