Radiotherapy and Oncology 130 (2019) 172–179

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

A machine learning-based prediction model of H3K27M mutations in brainstem gliomas using conventional MRI and clinical features

Chang-cun Pan [a,1], Jia Liu [b,1], Jie Tang [a], Xin Chen [a], Fang Chen [b], Yu-liang Wu [a], Yi-bo Geng [a], Cheng Xu [a], Xinran Zhang [b], Zhen Wu [a], Pei-yi Gao [c], Jun-ting Zhang [a], Hai Yan [d], Hongen Liao [b,*], Li-wei Zhang [a,*]

[a] Department of Neurosurgery/China National Clinical Research Center for Neurological Diseases, Beijing Tiantan Hospital, Capital Medical University; [b] Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing; [c] Department of Radiology, Beijing Tiantan Hospital, Capital Medical University, Beijing; and [d] Department of Pathology, Duke University Medical Center, Durham, USA

ARTICLE INFO

Article history:
Received 15 January 2018
Received in revised form 3 July 2018
Accepted 12 July 2018
Available online 7 August 2018

Keywords:
Brainstem glioma
H3K27M
Machine earning
MRI
Prediction

ABSTRACT

Background: H3K27M is the most frequent mutation in brainstem gliomas (BSGs), and it has great significance in the differential diagnosis, prognostic prediction and treatment strategy selection of BSGs. There has been a lack of reliable noninvasive methods capable of accurately predicting H3K27M mutations in BSGs.

Methods: A total of 151 patients with newly diagnosed BSGs were included in this retrospective study. The H3K27M mutation status was obtained by whole-exome, whole-genome or Sanger's sequencing. A total of 1697 features, including 6 clinical parameters and 1691 imaging features, were extracted from pre- and post-contrast T1-weighted and T2-weighted images. Using a random forest algorithm, 36 selected MR image features were integrated with 3 selected clinical features to generate a model that was predictive of H3K27M mutations. Additionally, a simplified prediction model comprising the Karnofsky Performance Status (KPS) at diagnosis, symptom duration at diagnosis and edge sharpness on T2 was established for practical clinical utility using the least squares estimation method.

Results: H3K27M mutation was an independent prognostic factor that conferred a worse prognosis (p = 0.01, hazard ratio = 3.0, 95% confidence interval [CI], 1.57–5.74). The machine learning-based model achieved an accuracy of 84.44% (area under the curve [AUC] = 0.8298) in the test cohort. The simplified model achieved an AUC of 0.7839 in the test cohort.

Conclusions: Using conventional MRI and clinical features, we established a machine learning-based model with high accuracy and a simplified model with improved clinical utility to predict H3K27M mutations in BSGs.

Brainstem gliomas (BSGs) represent a group of highly heterogeneous tumors. Diffuse intrinsic pontine glioma (DIPG) is the most common type of BSG, accounting for 80% and 45–50% of BSGs in children and adults, respectively [1,2]. Over the past 5 years, the unique genetic landscape and tumorigenesis of DIPG have been elucidated [3,4]. The heterozygous somatic H3K27M mutation, which affects nearly 80% of pediatric DIPGs and is present at the original and metastatic sites throughout the entire course of the disease, initiates tumorigenesis by reprogramming histone H3K27 methylation and gene expression [5–7]. While the WHO pathological grade cannot predict prognosis [8], the H3K27M mutation portends a worse prognosis, with a mean overall survival (OS) of 0.73 years for patients with H3K27M tumors versus 4.83 years for patients with wild-type (WT) tumors [9]. Recently, inhibition of histone demethylation and deacetylation has been demonstrated to prolong survival in patient-derived orthotopic xenograft mouse models [10,11]. Taken together, these findings indicate that the H3K27M mutation can serve as a qualified biomarker for diagnosis, prognosis and therapeutic selection in DIPG patients.

Compared with DIPG, tumors arising from the midbrain or medulla oblongata are more likely to be low-grade gliomas with relatively favorable prognoses. With the advancement of intraoperative electrophysiological monitoring and multimodal neuroimaging techniques, surgical resection has become safely feasible for these tumors and can prolong high-quality OS [12–17]. The decision to pursue surgery is usually based on the

* Corresponding authors at: Tsinghua University, Haidian District, Beijing 100084, China (H. Liao), Tiantan Xili 6, Dongcheng District, Beijing 100050, China (L.-w. Zhang).

E-mail addresses: liao@tsinghua.edu.cn (H. Liao), zhangliweittyy@163.com (L.-w. Zhang).

[1] Chang-cun Pan and Jia Liu contributed equally to this work.

https://doi.org/10.1016/j.radonc.2018.07.011
0167-8140/

neurosurgeon's inference of the tumor grade from both the clinical and MRI features, as surgical resection cannot prolong OS for patients with high-grade tumors. However, there are no distinct radiological differences between low-grade and high-grade gliomas because they usually share common features; for example, both grade I pilocytic astrocytomas (PAs) and grade III–IV high-grade gliomas can demonstrate contrast enhancement, intratumoral hemorrhage and leptomeningeal dissemination [18–20]. However, an accurate differential diagnosis is of paramount importance, considering the different prognoses and treatment strategies of these tumors. Recent studies have shown that PA is predominantly driven through the activation of the MAPK pathway without the other significant mutations that are involved in grade II–IV gliomas [21]. Therefore, the presence of H3K27M mutations can be used to rule out PA and avoid an inappropriate treatment strategy.

Based on these observations, a preoperative noninvasive method for reliably predicting the H3K27M mutation status in BSGs would be extremely helpful. Radiomics is a new field of research that examines the relationship between the medical imaging data and molecular features of a disease [22]. The machine-learning algorithm has shown advantages for complex imaging and big "-omics" data [23]. However, radiomic studies in BSGs are lacking, which is likely due to the paucity of patients with available genetic information in a single institution. We collected 151 BSGs from cases with available H3K27M mutation information from November 2010 to March 2017, and using these data, we created a machine learning-based model capable of accurately predicting H3K27M mutations. Moreover, we constructed a simplified model to improve the practical utility of the model. To the best of our knowledge, this is the first study in this field.

Materials and methods

Patient enrollment

This study was approved by the Institutional Review Board of Beijing Tiantan Hospital, Capital Medical University. The clinical, radiological and histopathological data were obtained from medical charts. The inclusion criteria were as follows: (1) primary brainstem tumor with a definitive histological diagnosis of glioma; (2) available information regarding the H3K27M mutation status; and (3) available preoperative T1-weighted images (T1), T2-weighted images (T2) and contrast-enhanced T1-weighted images (T1c). Patients who had previously received radiotherapy, stereotactic radiosurgery, anti-vascular therapy or surgery at a local hospital were excluded. Of the 183 patients, 151, including 91 with H3K27M-mutant tumors and 60 with H3K27-WT tumors, were ultimately enrolled. A flowchart is presented in Supplementary Fig. S1 to specify which patients were excluded for which reasons.

Patient characteristics and survival analysis

The patient characteristics are provided in Supplementary Table 1. Overall, 91 tumors were identified as harboring H3K27M mutations, all of which targeted H3F3A with no HIST1H3/C detected. Among these mutant tumors, 25 were grade II, 40 were grade III, and 26 were grade IV, whereas no grade I tumors were detected. Forty-two (46.2%) mutations were found in children (0–14 years old), of whom 32 (35.2%) were between 5 and 10 years old. The other 49 (53.8%) mutations were relatively evenly distributed across different age groups, ranging from 15 to 60 years old. Fifty-three (58.2%) of the H3K27M-mutant tumors were located at the ventral brainstem, accounting for 64.6% (53/82) of all ventral tumors. The H3K27M-mutant ventral tumors were predominantly (38/53, 71.7%) found in children, whereas the H3K27M-mutant dorsal tumors were predominantly (34/38, 89.5%) found in adults. The median OS of patients with H3K27M-mutant tumors was 11.1 months, which was significantly shorter than that of patients with H3K27-WT tumors (Kaplan–Meier's survival analysis, $p < 0.01$) (Fig. 2). The multivariate Cox regression analysis, which included children/adults, ventral/dorsal tumor location, pathological grade and H3K27 mutation status, showed that the H3K27M mutation was an independent prognostic factor that conferred a worse prognosis ($p = 0.01$, hazard ratio = 3.0, 95% confidence interval [CI], 1.57–5.74).

Histopathological diagnosis and H3K27M mutation detection

One experienced neuropathologist (G.L.L.) reviewed all diagnostic materials from the archived histopathological slides for each case. The H3K27M mutation status, including H3F3A and HIST1H3B/C, was obtained from whole-exome sequencing in 33 cases (previously published data) and whole-genome sequencing in 101 cases (unpublished data, performed on the Illumina X-Ten platform at Genetron Health Co., Ltd. Beijing, China) [24]. An additional 17 samples were sequenced for H3F3A K27M and HIST1H3B/C K27M by Sanger's sequencing.

MRI data acquisition and preprocessing

The standard MR imaging protocol for BSGs at our institution includes non-enhanced axial T1-weighted (TR, 2031 ms; TE, 19.536 ms; slice thickness, 5 mm), axial T2-weighted fast spin-echo (TR, 4900 ms; TE, 117 ms; slice thickness, 5 mm) and Omniscan (GE Healthcare, Little Chalfont, Buckinghamshire, United Kingdom; 0.1 mmol/kg)-based T1c (TR, 2031 ms; TE, 19.536 ms; slice thickness, 5 mm), with an FOV of 24 cm and a matrix size of 512*512. The MRI scanning device used was a GE Discovery MR750 3.0T. 3D Slicer (version 4.1), a user-driven and open-source software platform for medical image informatics and image processing, which was used to segment the BSG region in each 2D slice image [25]. The tumor contour was delineated manually by the two neurosurgeons (P.C.C. and T.J.) in source T1, T1c and T2 images. The corresponding author (Z.L.W.) and senior neuroradiologist (G.P.Y.) reviewed the results. The entire tumor region (TR) volume was labeled on T1, T2 and T1c, and the enhanced region was labeled on T1c (Fig. 1). After that, as previously reported in glioblastoma [26], the largest axial tumor slice was used to extract the region of interest (ROI) (2D-ROI and 3D-ROI, as shown in the brainstem tumor segmentation part of Fig. 1). The MRI features were extracted from the segmented TR and ROI (ROI was defined as a rectangular (for 2D slice images) or cuboid (for 3D volume data) region containing the brainstem tumor). To obtain more robust image features, we performed preprocessing of MRI sequences. First, we used an improved nonparametric nonuniform intensity normalization algorithm bias correction (N4BiasFieldCorrection) [27] to correct for inhomogeneity on T1 and T1c sequences. N4BiasFieldCorrection was performed using Slicer 3D. Then, we resampled all MR images to a uniform voxel size ($1 \times 1 \times 1$ mm) using trilinear interpolation. The final processing step involved scaling the intensity of the MRI volume data to 0–255.

MRI feature and clinical parameter extraction

Four categories of features, namely, texture, shape, edge sharpness and pixel intensity, were extracted from the conventional MR images of each tumor. For 2D image features, only the largest axial tumor slice was used. The textural features included two cohorts of global image features: gray-level co-occurrence matrix (GLCM),

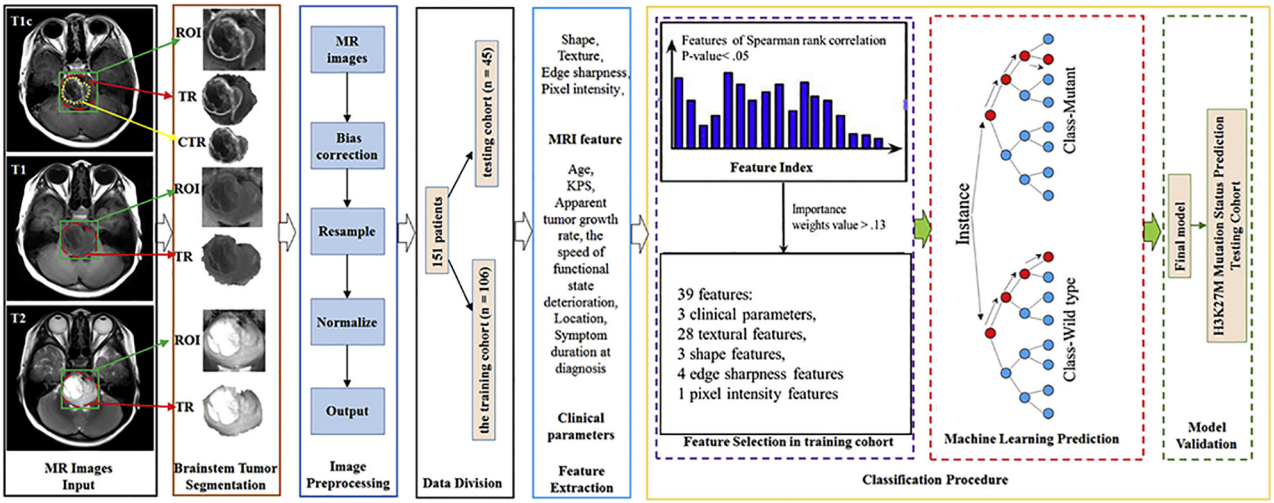

Fig. 1. MRI feature extraction and machine-learning pipeline. With manually segmented tumor, we first extracted 1691 quantitative imaging features, including texture, shape, edge sharpness and pixel intensity from masked pre-surgical T1, T1c and T2 MRIs. Then, feature selection was applied on the extracted features with Spearman's correlation and relief algorithm and 39 features were selected. Finally, selected 36 radiomics signature and 3 clinical characteristics were incorporated into a random forest based prediction model for individually prediction.

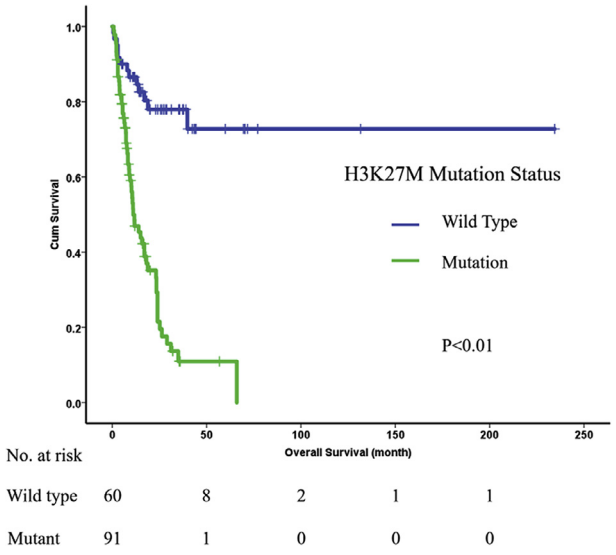

Fig. 2. Kaplan–Meier's survival curves of H3K27-WT versus H3K27M-mutant brainstem gliomas.

local binary patterns (LBPs), histogram of oriented gradients (HOGs), and Haralick textural features. For each MRI sequence, we extracted 144 3D-GLCM features, 80 3D-HOG features, 10 2D-LBP features and 14 2D-Haralick features for ROI and TR. The 15 shape features included the compactness index, surface area, volume, 10 Fourier descriptors and the fractal dimension for TR in each sequence. The 30 edge sharpness features included the median, mean and variance of the pixel intensity on, inner and outer of the tumor edge (the details were shown in Fig. S2 of Supplementary Material) and the difference in the mean pixel intensity values across the tumor edge in each sequence. We also extracted 20 pixel intensity statistical features for each TR and ROI, including the mean value, median value, variance value, Shannon's entropy and quartile value in each sequence. In addition, we obtained the volume and surface area of contrast-enhanced TR. Six different values (intensity average and variance value between three sequences) were calculated as MRI features. Therefore, 1691 MRI features were extracted from three MRI sequences (detail definition in Supplementary material).

Four traditional clinical parameters were used, including the Karnofsky Performance Status (KPS) at diagnosis, symptom duration at diagnosis (SD), age, and tumor location. Additionally, we established two previously undefined clinical parameters: the apparent tumor growth rate (ATGR, which was defined as the ratio of tumor volume to SD) and the speed of functional state deterioration (SFSD, calculated as KPS/SD). Finally, 1697 features were extracted for each patient, including a feature vector for 1691 quantitative image features and 6 clinical features.

Construction of machine learning-based prediction model

The proposed prediction model aimed to predict the H3K27M mutation status using MRI features and clinical parameters based on machine-learning methods (shown in Fig. 1). To achieve this goal, three parts of this study were executed using a pattern-recognition procedure.

Part 1, division of samples: In this study, all 151 patients were randomly assigned to either the training cohort ($n = 106$) or the test cohort ($n = 45$). The training cohort included 63 H3K27M-mutant tumors and 43 H3K27-WT tumors. Because the sample class of training cohort is not balanced, 43 H3K27-WT tumors' data were added to 61 data s by directly copying the sample data. Therefore the new training cohort included 63 H3K27M-mutant tumors and 61 H3K27-WT tumors. The test cohort included 28 H3K27M-mutant tumors and 17 H3K27-WT tumors.

Part 2, feature selection: To remove irrelevant or less discriminative features, which could skew the classifier's performance, features with a Spearman rank correlation coefficient (p-value > 0.05) were removed from the training cohort. The remaining 39 features with importance weights >0.013 were used as the result of feature selection by relief algorithm [28] on the training cohort. We then performed a linear discriminant analysis (LDA) to map the features to a dimensionality feature space that could be more easily classified.

Part 3, classification: We used a random forest algorithm to create the prediction model using selected features. All cases in the training cohort were used to train the classifier, while cases in the validation cohort were used to independently evaluate the performance of the final model. A key problem with constructing

random forests is selection of the best parameters. The solution is based on the calculation of the out-of-bag classification error. In implementing the random forest algorithm, we specified 2 parameters: (i) the number of decision trees in the ensemble and (ii) the minimum number of observations per tree leaf. Two other parameters—the number of trees and the minimum number of leaf nodes in the random forest—were determined through hyperparameter tuning, and we selected the best model based on the minimum out-of-bag classification error in the training cohort. The out-of-bag classification error is an unbiased estimate of the generalization error of a random forest, and its results are similar to those of K-fold cross-validation, which requires a large number of calculations. Moreover, the estimated error rate is as accurate as the error rate obtained by a test set that is in accordance with the size of the training set [29]. The number of trees was set to 50, and the minimum number of leaf nodes was 2. The performance of the final model was validated in the test cohort by calculating the prediction accuracy and performing an area under the receiver operating characteristic curve (AUC) analysis.

The prediction method was developed in the MATLAB 2017a programming language with the MATLAB Statistics and Machine-Learning Toolbox.

Simplified model

Despite the high accuracy of a machine learning-based prediction model of IDH1 status [30], many of the adopted features are difficult to interpret and use in clinical practice. Therefore, we attempted to create a simplified version of the machine learning-based model to improve its clinical utility. Considering the demand for utility, we chose two clinical parameters after a comprehensive consideration of clinical practicality and one MRI feature to construct a simplified. To create the simplified model, we generated a logistic model of the probability of H3K27M mutation as a function of four selected features in the training cohort. The mutation probability of a sample was designated *P*, and the two clinical parameters and one MRI feature was represented by X1, X2 and X3.

Results

Image feature robustness analysis

We assessed the feature stability for delineation inaccuracies using the multiple delineation data generated by two neurosurgeons (P.C.C. and T.J.). All radiomic features were computed for two delineations per patient, and a stability value was calculated for each feature using Student's *t*-test. The *t*-test can be used to determine whether two sets of data are significantly different from each other. We tested the null hypothesis that the pairwise difference between data vectors x and y has a mean equal to zero. A returned p-value < 0.05 indicated that *t*-test rejected the null hypothesis at the default 5% significance level. Therefore, a p-value < 0.05 indicated that the feature was not robust for delineation inaccuracies. Of the 1697 analyzed features, 499 had a p-value < 0.05, indicating that they were the least stable features. These 499 radiomic features included 68 of 480 HOG features, 303 of 864 GLCM features, 3 of 45 shape features, 82 of 90 edge features and 43 of 60 intensity features. In addition, our selected 39 features used to generate the prediction model were all robust, with p-values < 0.05. The results indicated that all the 2D-LBP features and 2D-Haralick features were robust for multiple delineation data.

Feature selection results and univariate analysis

Finally, 39 features were selected, including 3 clinical parameters, 28 textural features, 3 shape features, 4 edge sharpness features and 1 pixel intensity features. The three clinical parameters selected were KPS, ATGR and SFSD. The selected MRI textural features included 14 3D-HOG features, 4 2D-LBP features,

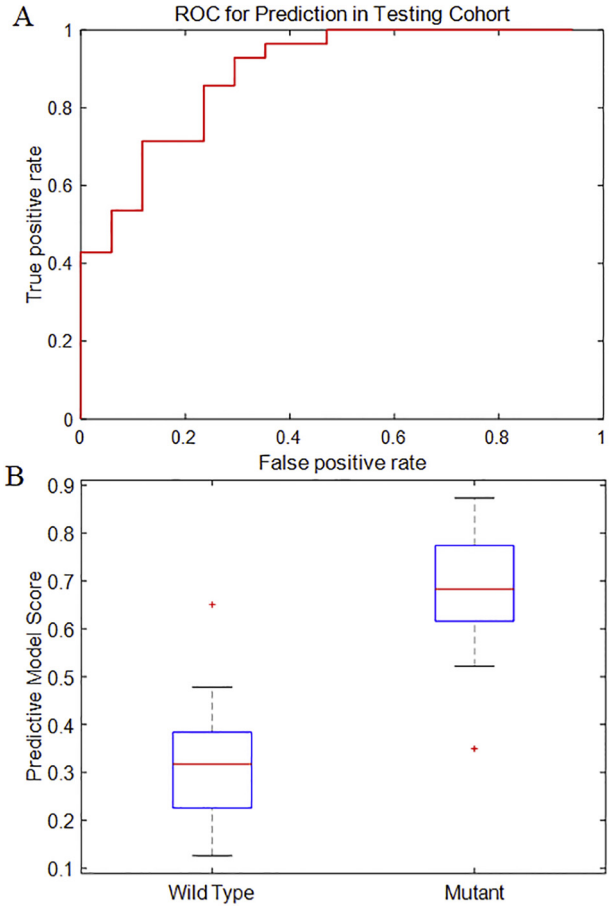

Fig. 3. (A) Random forest classifier scores for H3K27M-wildtype and H3K27M-mutant brainstem gliomas in the validation cohort; (B) Receiver operating characteristic (ROC) curve for prediction in the validation cohort.

Table 1
Top-performing non-redundant features in univariate prediction of H3K27M mutation based on area under the receiver operating characteristic curve (AUC) values in the training and test cohorts.

Feature Type	Image	Feature	Training AUC	Test AUC
Edge sharpness	T2-3D	Difference in average intensity between outer edge and inner edge	0.7218	0.6744
Edge sharpness	T2-3D	Difference in average intensity between edge and inside edge	0.7039	0.7311
Texture GLCM	T2-3D-ROI	MaxProbability	0.6780	0.5861
Texture GLCM	T2-3D-ROI	Energy	0.6664	0.6176
Texture GLCM	T2-3D-ROI	Correlation	0.6651	0.6449

T2, T2-weighted images; ROI, region of interest.

3 2D-Haralick features, and 7 3D-GLCM features. To assess the association between the selected features and H3K27M mutation status, the AUC was calculated by generating a logistic model of the probability of H3K27M mutation as a function of univariate features in the training cohort and the test cohort. The performance of the top 5 features is presented in Table 1. From the results in Table 1, the edge sharpness on T2 feature achieved the highest performance for predicting H3K27M mutation, and MRI textural features on T2 were more important than those on T1 and T1c.

Model prediction performance

The out-of-bag classification error of the prediction model in the training cohort was 0.1935. A prediction accuracy of 84.44% (AUC = 0.8298) was achieved in the test cohort (Fig. 3A). In a random forest classifier, a matrix of classification scores (score) indicates the likelihood that a label comes from a particular class. In our prediction classifier model, a sample with a mutation label in the test cohort had a mean probability (score) of 0.68 (95% CI: 0.64–0.72) for being predicted to be mutant and 0.32 (95% CI: 0.28–0.36) for being predicted to be WT (Fig. 3B).

To assess the impact of MRI features alone, we generated a model excluding selected 3 clinical parameters (KPS, ATGR and SFSD). This MRI feature-only model achieved a prediction accuracy of 80.00% (AUC = 0.7899) in the test cohort and an out-of-bag classification error of 0.2177 in the training cohort. To assess the impact of clinical parameters alone, we generated a model including three clinical parameters and excluding 36 MRI features. This clinical parameters-only model achieved a prediction accuracy of 66.67% (AUC = 0.6071) in the test cohort and an out-of-bag classification error of 0.3065 in the training cohort. Therefore, our model combining clinical parameters with image features could predict H3K27M mutation status better than models relying solely on clinical parameters or MRI features.

When selecting features for the simplified model, we first chose two clinical parameters (X1: KPS and X2: SD) that are easy for a physician to obtain and with a Spearman rank correlation coefficient $p < 0.05$. Then, one MRI feature was selected from the top 5 non-redundant features in the univariate prediction of H3K27M mutation (see Table 1). The one selected features were X3: edge sharpness on T2 (top 1). Example values of these features in different patients are presented in Fig. 4. Edge sharpness mainly reflects the invasiveness of the tumor cells. The selected MRI feature was assigned a score of −1, 0, 1 according to the distribution of their values in the training, which can be used to neuroradiologist. In addition, this one MRI features were chosen because they are easily understood by physicians and as a subjective qualitative judgment based on the clinical experience of the senior author. Therefore, two clinical parameters and one MRI features were used to create the simplified prediction model of H3K27M mutation.

To create the simplified model, we generated a logistic model of the probability of H3K27M mutation as a function of the four selected features in the training cohort. The performance of the simplified model was assessed by determining its AUC in the training cohort (0.7608) and its prediction accuracy in the test cohort (AUC = 0.7836).

Through fitting the regression equation, each feature's weight was determined and used to calculate the mutation probability (P) as follows:

$$\mathrm{Logit}(P) = 3.17 - 0.04 * X1 - 0.04 * X2 + 0.75 * X3 \quad (1)$$

In addition, we calculated the mutation probability and created a nomogram, as shown in Fig. 5. The above results demonstrated that the simplified prediction model can serve as a qualified and convenient tool for daily use in the clinic.

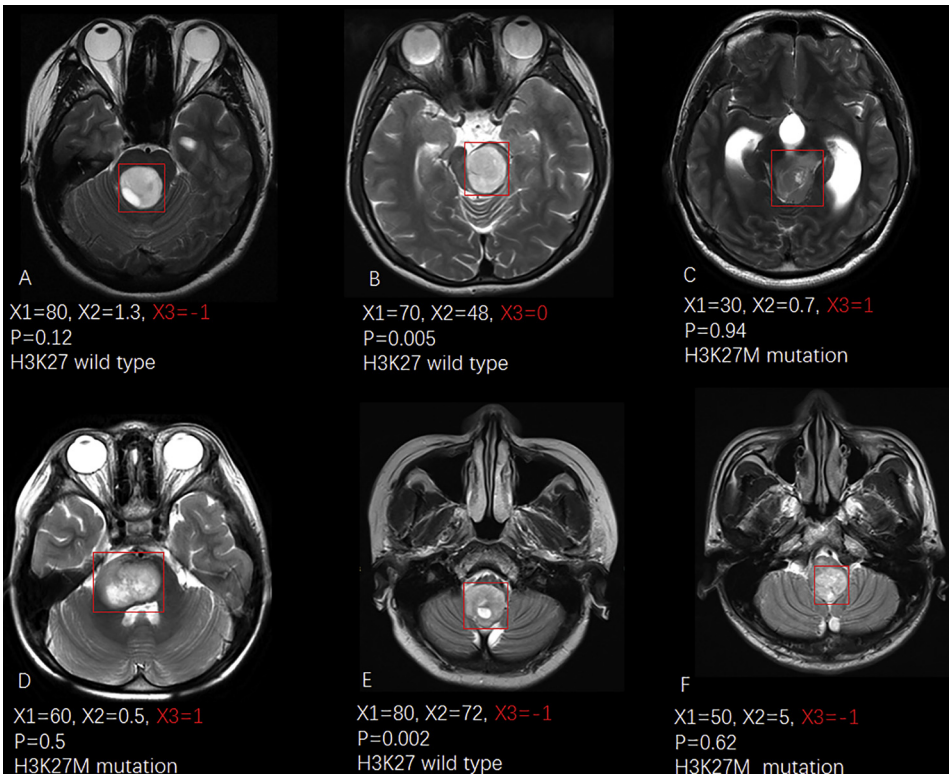

Fig. 4. Illustrative cases of the simplified prediction model. The upper and low panel demonstrates illustrative T2-weighted images from 3 BSGs with difference value of X3 (edge sharpness on T2); these images shows the variety of the intensity difference between the inner of the tumor and the outer of the tumor.

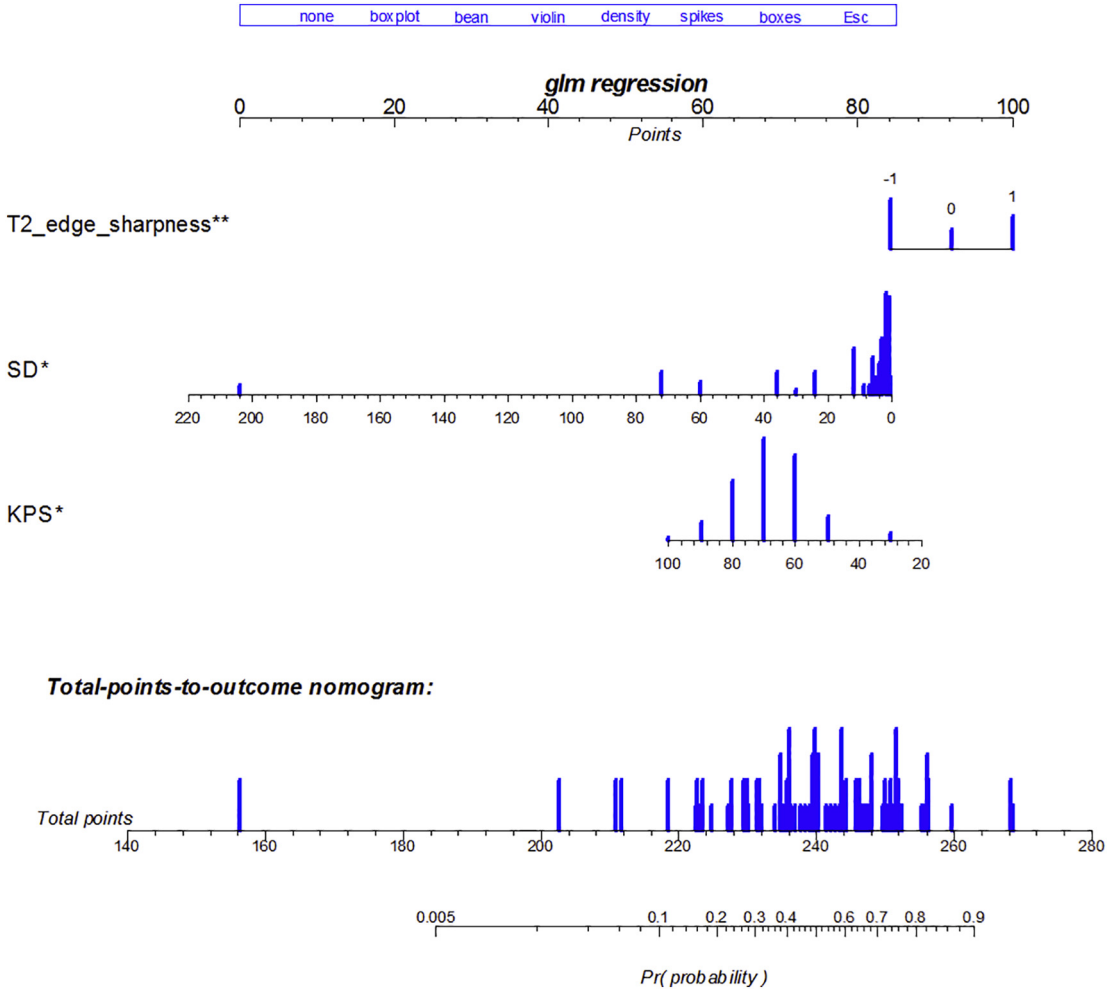

Fig. 5. Nomogram for the estimation of H3K27M mutation risk in brainstem glioma patient using simplified model.

Discussion

The emergence of radiomics has made routinely obtained MR images a promising noninvasive surrogate for important genetic features [31–33]. H3K27M is the most frequent mutation in BSGs and a perfect marker for differential diagnosis, prognosis prediction and treatment strategy selection [7,10,11,34]. In the present study, using conventional MRI and clinical features, we created a machine learning-based prediction model for H3K27M mutations, which achieved a prediction accuracy of 84.44% in the test cohort. Moreover, to improve its practical utility, we constructed a simplified model comprising 4 radiological and clinical features and achieved an accuracy of 75.55% in the test cohort. To the best of our knowledge, this represents the first radiomic study of BSGs.

In previous retrospective studies involving only tumors of a specific WHO grade or pathological type, qualitative analyses exhibited good performance. Sonoda et al. reported that the IDH mutation was strongly associated with the frontal location of grade III anaplastic gliomas [35]. Using the tumor volume and non-contrast-enhanced tumors, Carrillo, JA et al. achieved 97.5% accuracy in predicting IDH1 mutations in glioblastoma multiforme [36]. In an expanded cohort that included WHO grade III–IV tumors, a random forest algorithm-based prediction model of IDH1 achieved an accuracy of 89%, which demonstrated the capacity of machine learning-based quantitative analysis in addressing complex medical data [30]. However, these preset conditions limit the practical application of these studies' findings because tumor grade and pathological type cannot be obtained preoperatively. The current study included both children and adults with BSGs across all WHO grades, making our model more clinically applicable.

We did not use advanced MR images, such as DTI, MRS or PWI [30,37,38], because they are more susceptible to the effects of variations in scan parameters, head motion, and homogeneity of the magnetic field. However, unlike previous studies, we extracted 4 categories (i.e., texture, shape, edge and pixel intensity) of features to take full advantage of routinely obtained conventional MR images, making our model accurate and easy to implement. Notably, each of the 4 categories yielded features with strong contributions to the model, especially our newly defined edge sharpness on T2 (represented by the intensity difference across the edge).

The edge sharpness mainly reflects the invasiveness of the tumor cells. Autopsy studies have demonstrated that 50–56% of DIPGs had invasion of adjacent structures (cerebellum, spinal cord and thalamus), 25–38.6% had leptomeningeal dissemination and 62.5% had subventricular spread [8,39–41]. Despite a lack of H3K27M information, these observations confirmed the high migration ability of DIPG tumor cells. Since the discovery of the H3K27M mutation, 7 H3K27M-mutant DIPGs have been reported to exhibit distant spread to the hippocampus, frontal lobe or occipital lobe, even in the absence of clinical and histopathological evidence of a disseminated tumor in 2 cases [7,8,42]. Furthermore, the expression of H3.3K27M can increase the migration and invasion of neural progenitor cells in vitro [6]. Given the high frequency of

H3K27M mutations in DIPG, these findings may explain why H3K27M-mutant tumors are more likely to have a less sharp edge.

In the simplified model, SD and KPS are the two important contributors because they are faithful indicators of tumor biology [43]. H3K27M-mutant tumors follow a more rapid and progressive course with destructive neurological deficits, whereas H3K27-WT tumors follow an indolent course with minor neurological dysfunctions that are mainly due to tumor compression.

We also acknowledge several limitations of this work. First, these models were created based on data from a single institution; therefore, their generalizability needs to be further validated using independent data. Second, 2D-ROIs were manually annotated in a single section that contained the largest cross-section of the tumor, and manually annotating ROIs is labor intensive. In addition, manual segmentation of the tumor may introduce inter-performer variations; a future study will focus on optimizing manual segmentation using a convolution neural network method, which can reduce subjective factors and extract high-level image features. Additionally, to achieve high prediction accuracy, doctors need to be trained to use the simplified model.

In summary, we created 2 prediction models for H3K27M mutation in BSGs. One is a machine learning-based model with a high prediction accuracy of 84.44%. The other is a simplified model with improved practical utility and a modest prediction accuracy of 75.55%. These models use conventional MR images and clinical features and are based on BSGs from patients of all ages and all WHO grades, which make them more applicable and easy to implement than traditional models. These models may serve as qualified and convenient tools for differential diagnosis, prognostic prediction and treatment strategies for BSG patients.

Funding

Financial support was provided by the National Key Technology Research and Development Program of the Ministry of Science and Technology of China (Grant nos. 2014BAI04B01, 2015BAI12B04 and 2017YFC0108000); the National Natural Science Foundation of China (Grant nos. 81427803, 81528018 and 81771940); Beijing Municipal Administration of Hospitals Clinical Medicine Development of Special Funding Support (Grant no. ZYLX201608); the Beijing Municipal Science & Technology Commission (Grant no. Z151100003915079); and the Beijing Municipal Natural Science Foundation (Grant no. 7172122 and L172003).

Conflict of interest

The authors have no conflicts of interest to report.

Acknowledgments

We would like to thank Professor Gui-lin Li (Department of Neuropathology, Beijing Neurosurgical Institute) for her assistance in re-evaluating the histopathological slices.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.radonc.2018.07.011.

References

[1] Hargrove D, Bartels U, Bouffet E. Diffuse brainstem glioma in children: critical review of clinical trials. Lancet Oncol 2006;7:241–8.
[2] Laigle-Donadey F, Doz F, Delattre JY. Brainstem gliomas in children and adults. Curr Opin Oncol 2008;20:662–7.
[3] Puget S, Beccaria K, Blauwblomme T, et al. Biopsy in a series of 130 pediatric diffuse intrinsic Pontine gliomas. Childs Nerv Syst 2015;31:1773–80.
[4] Bartels U, Hawkins C, Vezina G, Kun L, Souweidane M, Bouffet E. Proceedings of the diffuse intrinsic pontine glioma (DIPG) Toronto Think Tank: advancing basic and translational research and cooperation in DIPG. J Neurooncol 2011;105:119–25.
[5] Chan KM, Fang D, Gan HY, et al. The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. Genes Dev 2013;27:985–90.
[6] Funato K, Major T, Lewis PW, Allis CD, Tabar V. Use of human embryonic stem cells to model pediatric gliomas with H3.3K27M histone mutation. Science 2014;346:1529–33.
[7] Nikbakht H, Panditharatna E, Mikael LG, et al. Spatial and temporal homogeneity of driver mutations in diffuse intrinsic pontine glioma. Nat Commun 2016;7:11185.
[8] Buczkowicz P, Bartels U, Bouffet E, Becher O, Hawkins C. Histopathological spectrum of paediatric diffuse intrinsic pontine glioma: diagnostic and therapeutic implications. Acta Neuropathol 2014;128:573–81.
[9] Khuong-Quang DA, Buczkowicz P, Rakopoulos P, et al. K27M mutation in histone H3.3 defines clinically and biologically distinct subgroups of pediatric diffuse intrinsic pontine gliomas. Acta Neuropathol 2012;124:439–47.
[10] Hashizume R, Andor N, Ihara Y, et al. Pharmacologic inhibition of histone demethylation as a therapy for pediatric brainstem glioma. Nat Med 2014;20:1394–6.
[11] Grasso CS, Tang Y, Truffaux N, et al. Functionally defined therapeutic targets in diffuse intrinsic pontine glioma. Nat Med 2015.
[12] Epstein F, McCleary EL. Intrinsic brain-stem tumors of childhood: surgical indications. J Neurosurg 1986;64:11–5.
[13] Epstein F, Wisoff J. Intra-axial tumors of the cervicomedullary junction. J Neurosurg 1987;67:483–7.
[14] Pollack IF, Hoffman HJ, Humphreys RP, Becker L. The long-term outcome after surgical treatment of dorsally exophytic brain-stem gliomas. J Neurosurg 1993;78:859–63.
[15] Constantini S, Epstein F. Surgical indication and technical considerations in the management of benign brain stem gliomas. J Neuro-oncology 1996;28:193–205.
[16] Bricolo A. Surgical management of intrinsic brain stem gliomas. Operative Tech Neurosurg 2000;3:137–54.
[17] Wang CC, Zhang JT, Liu A, Sun B, Zhao YL. Surgical treatment of primary midbrain gliomas. Surg Neurol 2000;53:41–51.
[18] Kumar AJ, Leeds NE, Kumar VA, et al. Magnetic resonance imaging features of pilocytic astrocytoma of the brain mimicking high-grade gliomas. J Comput Assist Tomogr 2010;34:601–11.
[19] White JB, Piepgras DG, Scheithauer BW, Parisi JE. Rate of spontaneous hemorrhage in histologically proven cases of pilocytic astrocytoma. J Neurosurg 2008;108:223–6.
[20] Bian SX, McAleer MF, Vats TS, Mahajan A, Grosshans DR. Pilocytic astrocytoma with leptomeningeal dissemination. Childs Nerv Syst 2013;29:441–50.
[21] Jones DT, Hutter B, Jager N, et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat Genet 2013;45:927–32.
[22] Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
[23] Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. Nat Rev Genet 2015;16:321–32.
[24] Zhang L, Chen LH, Wan H, et al. Exome sequencing identifies somatic gain-of-function PPM1D mutations in brainstem gliomas. Nat Genet 2014;46:726–30.
[25] Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D slicer as an image computing platform for the quantitative imaging network. Magn Reson Imaging 2012;30:1323–41. PMID: 22770690.
[26] Gevaert O, Mitchell LA, Achrol AS, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features. Radiology 2014;273:168–74.
[27] Tustison NJ, Avants BB, Cook PA, et al. N4ITK: improved N3 bias correction. IEEE Trans Med Imag 2010;29:1310–20.
[28] Robnikšikonja M, Kononenko I. Theoretical and empirical analysis of ReliefF and RReliefF. Machine Learn 2003;53:23–69.
[29] Breiman L. Random forests. Machine Learn 2001;45:5–32.
[30] Zhang B, Chang K, Ramkissoon S, et al. Multimodal MRI features predict isocitrate dehydrogenase genotype in high-grade gliomas. Neuro-oncology 2016;19:109–17.
[31] Yamashita K, Hiwatashi A, Togao O, et al. MR Imaging-based analysis of glioblastoma multiforme: estimation of IDH1 mutation status. AJNR Am J Neuroradiol 2016;37:58–65.
[32] Brown R, Zlatescu M, Sijben A, et al. The use of magnetic resonance imaging to noninvasively detect genetic signatures in oligodendroglioma. Clin Cancer Res 2008;14:2357–62.
[33] Ellingson BM. Radiogenomics and imaging phenotypes in glioblastoma: novel observations and correlation with molecular characteristics. Curr Neurol Neurosci 2015;15:506.
[34] Nagaraja S, Vitanza NA, Woo PJ, et al. Transcriptional dependencies in diffuse intrinsic pontine glioma. Cancer Cell 2017;31. pp. 635–652 e636.
[35] Sonoda Y, Shibahara I, Kawaguchi T, et al. Association between molecular alterations and tumor location and MRI characteristics in anaplastic gliomas. Brain Tumor Pathol 2015;32:99–104.

[36] Carrillo J, Lai A, Nghiemphu P, et al. Relationship between tumor enhancement, edema, IDH1 mutational status, MGMT promoter methylation, and survival in glioblastoma. Am J Neuroradiol 2012;33:1349–55.
[37] Xiong J, Tan WL, Pan JW, et al. Detecting isocitrate dehydrogenase gene mutations in oligodendroglial tumors using diffusion tensor imaging metrics and their correlations with proliferation and microvascular density. J Magn Reson Imaging 2016;43:45–54.
[38] Choi C, Ganji SK, DeBerardinis RJ, et al. 2-Hydroxyglutarate detection by magnetic resonance spectroscopy in IDH-mutated patients with gliomas. Nat Med 2012;18:624–9.
[39] Caretti V, Bugiani M, Freret M, et al. Subventricular spread of diffuse intrinsic pontine glioma. Acta Neuropathol 2014;128:605–7.
[40] Yoshimura J, Onda K, Tanaka R, Takahashi H. Clinicopathological study of diffuse type brainstem gliomas: analysis of 40 autopsy cases. Neurol Medico-chirurgica 2003;43:375–82.
[41] Gururangan S, McLaughlin CA, Brashears J, et al. Incidence and patterns of neuraxis metastases in children with diffuse pontine glioma. J Neurooncol 2006;77:207–12.
[42] Buczkowicz P, Hoeman C, Rakopoulos P, et al. Genomic analysis of diffuse intrinsic pontine gliomas identifies three molecular subgroups and recurrent activating ACVR1 mutations. Nat Genet 2014;46:451–6.
[43] Tate MC, Lindquist RA, Nguyen T, et al. Postnatal growth of the human pons: a morphometric and immunohistochemical analysis. J Comp Neurol 2015;523:449–62.