Radiomics Features on Computed Tomography Combined With Clinical-Radiological Factors Predicting Progressive Hemorrhage of Cerebral Contusion

Qingning Yang [1†], Jun Sun [1†], Yi Guo [1*‡], Ping Zeng [1*‡], Ke Jin [2], Chencui Huang [2], Jingxu Xu [2], Liran Hou [3], Chuanming Li [1] and Junbang Feng [1]

[1] Department of Radiology, Chongqing University Central Hospital, Chongqing, China, [2] Department of Research Collaboration, R&D Center, Beijing Deepwise & League of PHD Technology Co., Beijing, China, [3] Department of Radiology, Panjiang Central Hospital, Guizhou, China

Background: Traumatic brain injury (TBI) is the main cause of death and severe disability in young adults worldwide. Progressive hemorrhage (PH) worsens the disease and can cause a poor neurological prognosis. Radiomics analysis has been used for hematoma expansion of hypertensive intracerebral hemorrhage. This study attempts to develop an optimal radiomics model based on non-contrast CT to predict PH by machine learning (ML) methods and compare its prediction performance with clinical-radiological models.

Methods: We retrospectively analyzed 165 TBI patients, including 89 patients with PH and 76 patients without PH, whose data were randomized into a training set and a testing set at a ratio of 7:3. A total of 10 different machine learning methods were used to predict PH. Univariate and multivariable logistic regression analyses were implemented to screen clinical-radiological factors and to establish a clinical-radiological model. Then, a combined model combining clinical-radiological factors with the radiomics score was constructed. The area under the receiver operating characteristic curve (AUC), accuracy and F1 score, sensitivity, and specificity were used to evaluate the models.

Results: Among the 10 various ML algorithms, the support vector machine (SVM) had the best prediction performance based on 12 radiomics features, including the AUC (training set: 0.918; testing set: 0.879) and accuracy (training set: 0.872; test set: 0.834). Among the clinical and radiological factors, the onset-to-baseline CT time, the scalp hematoma, and fibrinogen were associated with PH. The radiomics model's prediction performance was better than the clinical-radiological model, while the predictive nomogram combining the radiomics features with clinical-radiological characteristics performed best.

Conclusions: The radiomics model outperformed the traditional clinical-radiological model in predicting PH. The nomogram model of the combined radiomics features and clinical-radiological factors is a helpful tool for PH.

Keywords: traumatic brain injury, progressive hemorrhage, radiomics, machine learning, nomogram

INTRODUCTION

Traumatic brain injury (TBI) is a disease with the highest mortality and disability rate in systemic trauma. It is the main cause of death and serious disability among the young adult population worldwide. The pathophysiological change in TBI is a dynamic process (1), and progressive hemorrhage (PH) is a significant cause of death and deterioration in TBI patients. It has been reported that the incidence rate of PH is approximately 18–64% (2–5). PH has an occult onset and a rapid progression, and when patients have the development of symptoms, the condition has already progressed to a certain extent, which brings difficulties for the clinical treatment. Moreover, PH directly affects the prognosis of patients, and prior cases (6, 7) have shown that the risk of developing a poor neurological outcome with PH was four times higher than that without PH. Therefore, early detection of PH is very important. At present, non-contrast CT is the first choice for brain trauma. It is very difficult to recognize PH by observing non-contrast CT images with the naked eye, and the clinical diagnosis of PH mainly depends on serial head CT scans. However, there are no unified guidelines on the CT reexamination time and frequency; blindly repeating head CT leads to unnecessary radiation exposure, and repeatedly moving the patients for examination can be harmful to them.

As a new technology, radiomics can extract many quantitative features from non-contrast CT images, can capture image information that cannot be evaluated by the naked eye, such as texture features and high-order features, and then can filter the quantitative features most relevant to the clinical findings by statistical methods. Radiomics has been widely used for evaluation of the central nervous system, such as tumors (8–11), the prediction of hematoma enlargement in hypertensive intracerebral hemorrhage (12, 13), the classification of intracranial aneurysm rupture (14), and the etiological classification of intracranial hematoma (15, 16). However, few studies have reported the application of radiomics in progressive hemorrhage of cerebral contusion.

The purpose of this retrospective study was to seek the optimal radiomics model based on non-contrast CT scanning to predict PH by applying various machine learning methods and comparing the prediction performance of this model with the traditional model based on clinical factors and radiological factors.

MATERIALS AND METHODS

Study Population

This study was approved by the medical ethics committee of the Chongqing University Central Hospital, and the requirements for informed consent were waived.

Patients with a brain contusion and hematoma on baseline head CT after trauma were recruited from June 2016 to December 2021 at the Chongqing University Central Hospital. The exclusion criteria were as follows: (1) patients <18 years old; (2) patients with a history of skull surgery; (3) patients with a baseline head CT more than 6 h after trauma; (4) patients who did not undergo repeated CT within 3 days of baseline; (5) severe artifacts on the baseline images; (6) patients with penetrating TBI; (7) patients with brain surgery or interventional therapy before the follow-up head CT; and (8) patients with anticoagulant therapy before trauma; and (9) patients with non-traumatic brain diseases, such as tumor, discovered on CT. The patient screening process is shown in **Figure 1**. After the exclusion criteria screening was applied, 165 cases with TBI were selected, including 89 cases with PH and 76 cases without PH. After admission, all patients were evaluated and treated according to the Clinical Guidelines for the Management of Head Injury (17).

The clinical-radiological data (age, sex, Glasgow Coma Scale score) were acquired from Picture Archiving and Communication Systems and the patients' medical records. PH was defined as the increase of the volume of intracranial hematoma by 25% or more when compared with the initial CT (18). The hematoma volume was calculated as ABC/2.

The head CT images were from United Imaging 760 64 Shanghai China and General Electric Company, LightSpeed 64, America. All the CT scanning parameters were similar: a tube voltage of 120 kV, a tube current of 200 mA, an axial technical section thickness of 5 mm, and a reconstruction interval of 1.25 mm.

To ensure that the labeling ratio was not affected, the data were randomly divided into training and testing sets at 7:3. Since some patients had multiple independent lesions, namely 30 patients (18.29%) had two lesions, 5 patients (3.05%) had three lesions, and 1 patient (0.61%) had four lesions, samples from the same patient were assigned to the same cohort to avoid data leakage. Finally, we randomly assigned lesions into a training set ($n = 145$) and a testing set ($n = 62$). Models were fitted on the training set and were independently tested on the testing set.

Image Segmentation and Feature Extraction

All regions of interest (ROIs) were delineated around the boundary using a postprocessing platform (Dr. Wise Multimodal Research Platform1 v1.6.3, Beijing Deepwise & League of PHD Technology Co., Ltd, Beijing, China). An open-source package of Python, called "PyRadiomics" (19), was used to generate the radiomics data from the original images. The software package could extract signal strength features, texture features, shape features, and higher-order features from the original images. Higher-order features are calculated from the above features using filters (wavelet, Laplace-Gauss, square root, logarithm, exponent, gradient transform, and local binary mode transform). All the calculation formulas and the pipeline can be found at https://pyradiomics.readthedocs.io/en/latest/. In total, 2,107 radiomics features were extracted automatically, including 414 first-order features, 14 shape features, 506 gray level co-occurrence matrix (GLCM) features, 368 gray level size zone matrix (GLSZM) features, 368 gray level run length matrix (GLRLM) features, 322 gray level dependence matrix (GLDM) features, and 115 neighboring gray-tone difference matrix (NGTDM) features.

The values of each training set feature were standardized with the Z score, and a similar normalization procedure, which
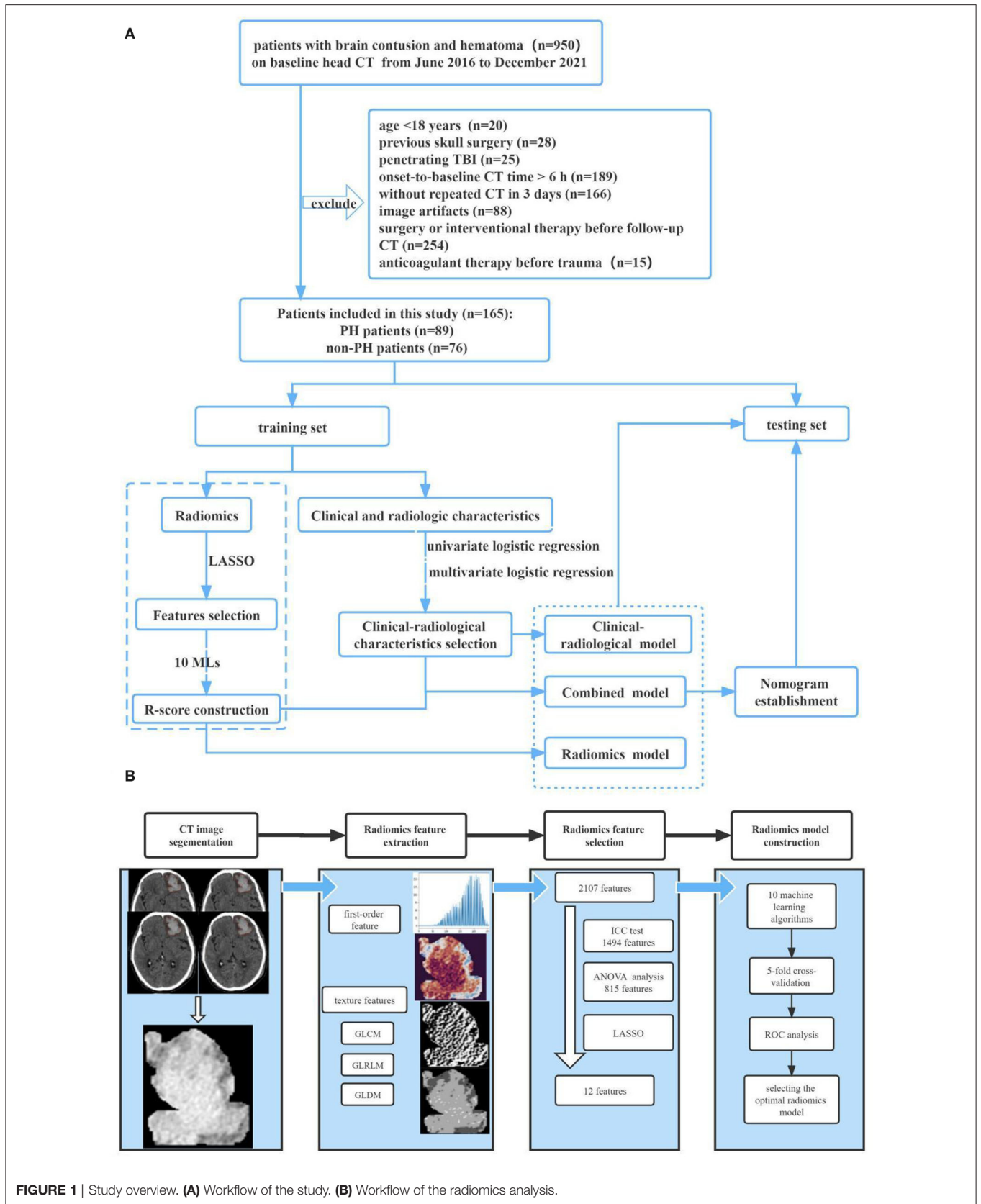
FIGURE 1 | Study overview. **(A)** Workflow of the study. **(B)** Workflow of the radiomics analysis.

utilized the mean and standard deviation values of the training set, was applied to the testing set.

Radiomics Feature Selection

First, we applied intraclass coefficient (ICC) analysis to identify features with high reproducibility and consistency of measurements. Reader 1 and Reader 2 independently redelineated the original images, and for each reader, 17 patients were randomly selected. Features with an ICC > 0.75 were selected for further analysis. After that, an analysis of variance (ANOVA) model was built, and it excluded features without significant differences between the PH group and the control group. Finally, we used the least absolute shrinkage and selection operator (LASSO) method, whose penalty parameter was tuned by a 5-fold cross-validation (CV).

The whole variable reduction process above was performed on the training set.

Machine Learning and Radiomics Signature Construction

We implemented 10 different machine learning (ML) algorithms—logistic regression (LR), support vector machine (SVM), K-nearest neighbors (KNN), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), Gaussian naive Bayes (GNB), artificial neural network (ANN), random forest (RF), XGBoost, and CatBoost—on the training dataset and determined the optimal hyperparameters of these models by the 5-fold CV method and grid search technique. For each selected model, receiver operating characteristic (ROC) curves were generated, and the average area under the curve (AUC) was calculated to assess the prediction performance. We used a learning curve to determine the overfitting of the model. A model was considered overfitted if its AUC for the training set was significantly higher than that for the validation set. Finally, with the best ML classifier, we built a radiomics model on the whole training set and constructed a radiomics signature (Rad-score) using the logit function

$$rad - score = \ln \frac{p}{1 - p}$$

where, p refers to the probability of PH predicted by our radiomics model for each patient. We also calculated the accuracy, F1 score, sensitivity, and specificity for comparison.

Clinical-Radiological Analysis

To select the clinical-radiological characteristics with significant differences between the PH group and control group, we performed the chi-square test for categorical variables and the Mann–Whitney U-test or Student's t-test for continuous variables. Then, univariate logistic regression and multivariable logistic regression were implemented to remove those features without significance and to determine the final clinical-radiological factors that were used to establish the clinical (LR) model.

Combined Model Establishment

A combination of the radiomics signature from the radiomics model and clinical-radiological signatures from the clinical-radiological model was used to conduct the combined model by the LR algorithm to provide a nomogram to evaluate the risk of PH. The ROC-AUC, accuracy, F1-score, sensitivity, and specificity metrics were calculated separately on the training set. The testing data was used to assess the performance of these three classifiers, and the decision curve and the calibration curve were also illustrated for comparison. A Youden index analysis was constructed to determine the optimal threshold (cutoff value) of the critical probability.

Statistical Analysis

"Scipy.stats," an open-source package of Python, was used for all statistical tests [20]. Categorical variables were compared by the chi-square test. The Shapiro–Wilk's test was performed to assess the normality of continuous variables. For variables with the null hypothesis rejected, we used the Mann–Whitney U-test to test for feature significance. For variables whose p-values were higher than 0.05, these variables followed a normal distribution, and then we performed the Levene test. If the results were significant, Welch's test was conducted. If the results were not significant, it was considered that the variances were equal, and Student's t-test was utilized. Differences in the AUC values between the different classifiers were compared using DeLong's test. A two-sided p-value < 0.05 was considered statistically significant. Hosmer–Lemeshow tests were used to calculate if the observed event rates matched the expected event rates in the population subgroups.

RESULTS

Patients' Characteristics

The clinical-radiological factors of the PH group and non-PH group are given in **Table 1**. The proportion of the PH group in the training set was 51.03% (74 of 145), and that in the validation set was 53.23% (33 of 62). In the training set, the factors including scalp hematoma, the GCS score, the onset-to-initial CT time, and fibrinogen level were significantly different between the two groups (p < 0.01). The other features were not significant (p > 0.05).

Radiomics Feature Selection

After intra- and interreader ICC analysis, a total of 1,494 radiomics features were considered robust and consistent. After that, 815 factors showed significant differences between the PH group and control groups (p < 0.05) via ANOVA. These features were finally reduced to 12 through the LASSO model (**Figure 2**).

Machine Learning and Radiomics Signature Construction

To find the optimal radiomics model to establish the Rad-score, we implemented ten machine-learning algorithms on the training set and compared their predictive ability by calculating commonly used metrics, including the ROC-AUC, accuracy, F1-score, sensitivity, and specificity (**Figure 3**). Using the 5-fold

TABLE 1 | Comparisons of patient characteristics in the training set and testing set.

	Training set (n = 145)			Testing set (n = 62)		
	PH (n = 74)	Non-PH (n = 71)	p value	PH (n = 33)	Non-PH (n = 29)	p value
Sex, male	14 (18.9)	12 (16.9)	0.999	13 (39.4)	3 (10.3)	0.147 [b]
Multiple primary brain contusions	52 (70.3)	49 (69.0)	1.000	28 (84.8)	25 (86.2)	1.000 [b]
Craniofacial fracture	67 (90.5)	59 (83.1)	0.779	33 (100.0)	28 (96.6)	0.885 [b]
Epidural hematoma	16 (21.6)	12 (16.9)	0.972	12 (36.4)	4 (13.8)	0.392 [b]
Subdural hematoma	67 (90.5)	54 (76.1)	0.239	28 (84.8)	25 (86.2)	1.000 [b]
Subarachnoid hemorrhage	67 (90.5)	58 (81.7)	0.665	33 (100.0)	27 (93.1)	0.671 [b]
Midline shift	3 (4.1)	0 (0.0)	0.568	3 (9.1)	0 (0.0)	0.597 [b]
Scalp hematoma	73 (98.6)	58 (81.7)	0.018*	33 (100.0)	26 (89.7)	0.465 [b]
Diabetes[a]	6 (8.1)	10 (14.1)	0.858	6 (18.2)	8 (27.6)	0.941 [b]
Hypertension[a]	22 (29.7)	19 (26.8)	0.997	14 (42.4)	5 (17.2)	0.330 [b]
Smoking	16 (21.6)	22 (31.0)	0.801	6 (18.2)	8 (27.6)	0.941 [b]
Alcoholism	19 (25.7)	24 (33.8)	0.887	7 (21.2)	15 (51.7)	0.179 [b]
Age, mean ± SD (years)	56.5 ± 17.742	59.155 ± 18.245	0.305	59.848 ± 18.346	57.862 ± 16.487	0.657 [c]
INR[a], mean ± SD	1.076 ± 0.187	1.061 ± 0.148	0.811	1.09 ± 0.116	1.1 ± 0.134	0.799
Admission SBP[a], mean ± SD (mmHg)	142.446 ± 21.144	148.408 ± 27.622	0.146 [c]	144.727 ± 28.627	138.69 ± 30.844	0.427 [c]
GCS score[a], mean ± SD	12.365 ± 2.356	12.958 ± 2.381	0.011*	11.242 ± 2.948	11.931 ± 3.023	0.072
Platelet, mean ± SD (109/L)	170.892 ± 57.15	180.535 ± 53.678	0.097	161.152 ± 42.854	176.897 ± 49.703	0.185 [c]
APTT[a], mean ± SD (s)	33.705 ± 4.486	33.615 ± 4.018	0.970	32.391 ± 2.567	35.379 ± 4.631	0.004 [d]
Prealbumin, mean ± SD (mg/L)	241.689 ± 59.41	239.38 ± 56.184	1.000	237.394 ± 67.875	232.931 ± 61.04	0.832
Onset-to-CT time[a], mean ± SD (h)	2.313 ± 1.345	2.984 ± 1.764	0.029*	2.326 ± 1.219	3.441 ± 1.921	0.062
First-to-second CT time, mean±SD (h)	22.53 ± 20.38	21.66 ± 18.57	0.182	20.11 ± 18.13	20.11 ± 20.79	0.554
Admission DBP[a], mean ± SD (mmHg)	84.541 ± 12.456	86.704 ± 14.395	0.147	84.485 ± 17.566	83.759 ± 13.543	0.761
Fibrinogen, mean ± SD (g/L)	2.483 ± 0.788	5.708 ± 13.488	0.003*	2.419 ± 0.623	2.897 ± 0.993	0.025 [c]

[a] Diabetes, Diabetes History; Hypertension, Hypertension history; INR, International normalized ratio; Admission SBP, Systolic blood pressure on admission; GCS score, Glasgow coma scale score; APTT, Activated partial thrombin time; Onset-to-CT time, Interval from trauma to first CT; Admission DBP, Diastolic blood pressure on admission.

[b] For categorical factors, the chi-squared test was used, data in parentheses are percentages. [c] Student's t-test. [d] Welch's t-test. For other continuous factors, the Mann–Whitney U-test was used, and for categorical factors, the chi-squared test was used. * Predictors included for further analysis (p < 0.05).

cross-validation (CV) method, we evaluated the performance of several classifiers in the training phase (CV training) and validation phase (CV validation). In addition to GNB, the other models performed well on the CV training. However, the two tree models of RF and Xgboost overfitted the CV training. The SVM model had the best predictive ability on the CV validation (AUC = 0.918), and there was no overfitting compared to the result on the CV training (AUC = 0.879). Therefore, the SVM algorithm was determined as our radiomics model and was used to generate the radiomics scores (Rad-score) on the training set and testing set.

Clinical-Radiological Model

The results of the univariate and multivariable LR are listed in **Table 2**. After that, scalp hematoma, the onset-to-CT time, and fibrinogen level, whose coefficients were significant in both univariate LR (p = 0.008, p = 0.013, and p = 0.014, respectively) and multivariable LR (p < 0.0001, p = 0.007, and p = 0.010, respectively), were determined as the variables to be used in the clinical-radiological model.

Combined Model Establishment

We then added the Rad-score to the clinical-radiological model; that is, radiomics and clinical-radiological signatures were used to develop a combined signature. On the training set, we implemented an LR algorithm and built a nomogram as the combined classifier, made predictions on the testing set, and assessed the performance of the clinical-radiological, radiomics, and combined models by calculating metrics including the ROC-AUC, accuracy, F1-score, sensitivity, and specificity (**Figure 4**; **Table 3**).

The combined model fit best on the training set (AUC = 0.937), and the results of DeLong's test showed significant differences between this and the other two models. In addition, the radiomics model was significantly better than the clinical-radiological model (p = 0.0002). In the other metrics, the performance of the combined model was the best. On the testing set, the score attained by the combined model (AUC = 0.853) was still significantly higher than that of the other two classifiers. Except for the specificity, the other indicators of the nomogram all ranked first. From the perspective of the calibration curve, the combined model was closer to the ideal on both the training
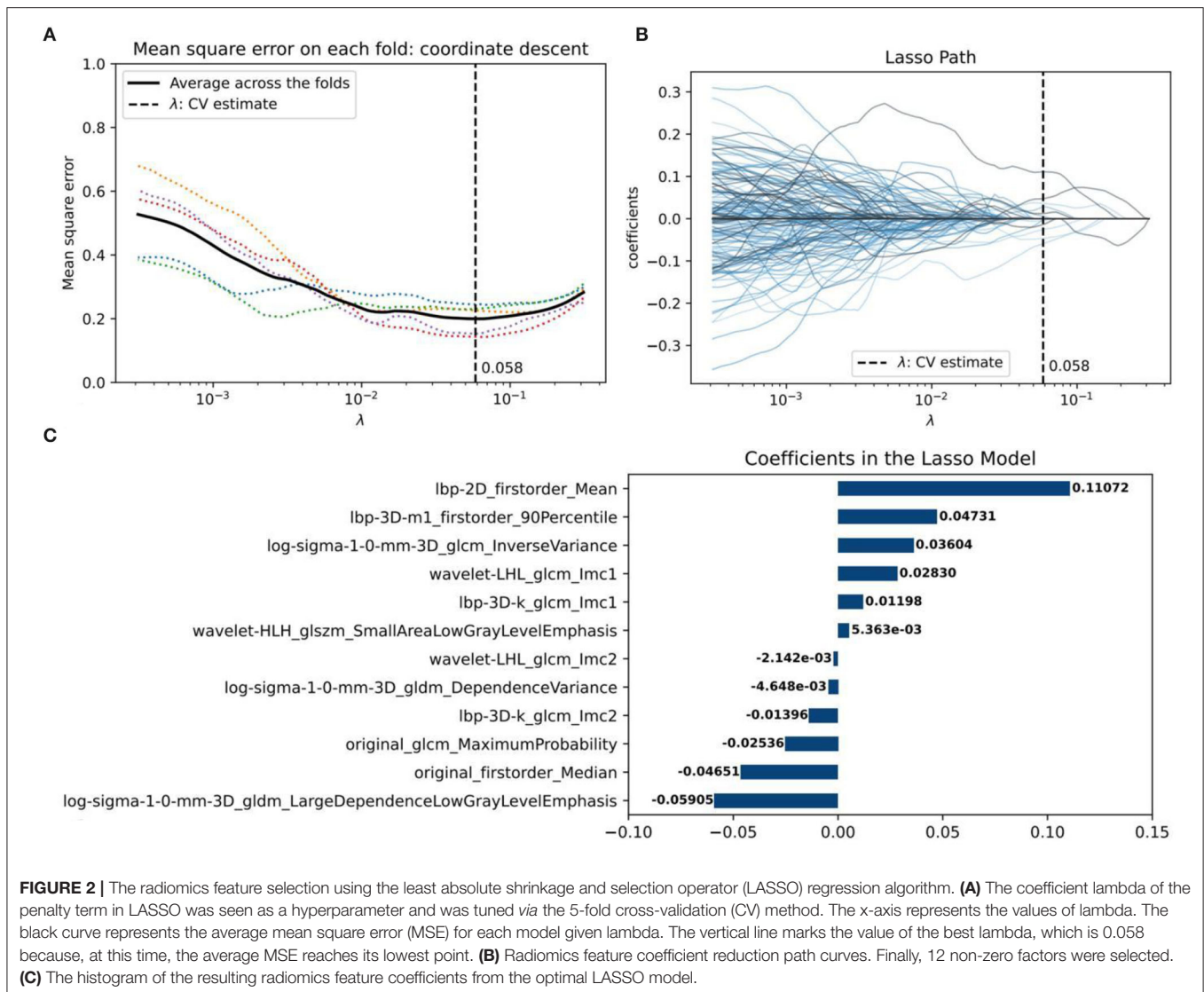
FIGURE 2 | The radiomics feature selection using the least absolute shrinkage and selection operator (LASSO) regression algorithm. **(A)** The coefficient lambda of the penalty term in LASSO was seen as a hyperparameter and was tuned *via* the 5-fold cross-validation (CV) method. The x-axis represents the values of lambda. The black curve represents the average mean square error (MSE) for each model given lambda. The vertical line marks the value of the best lambda, which is 0.058 because, at this time, the average MSE reaches its lowest point. **(B)** Radiomics feature coefficient reduction path curves. Finally, 12 non-zero factors were selected. **(C)** The histogram of the resulting radiomics feature coefficients from the optimal LASSO model.
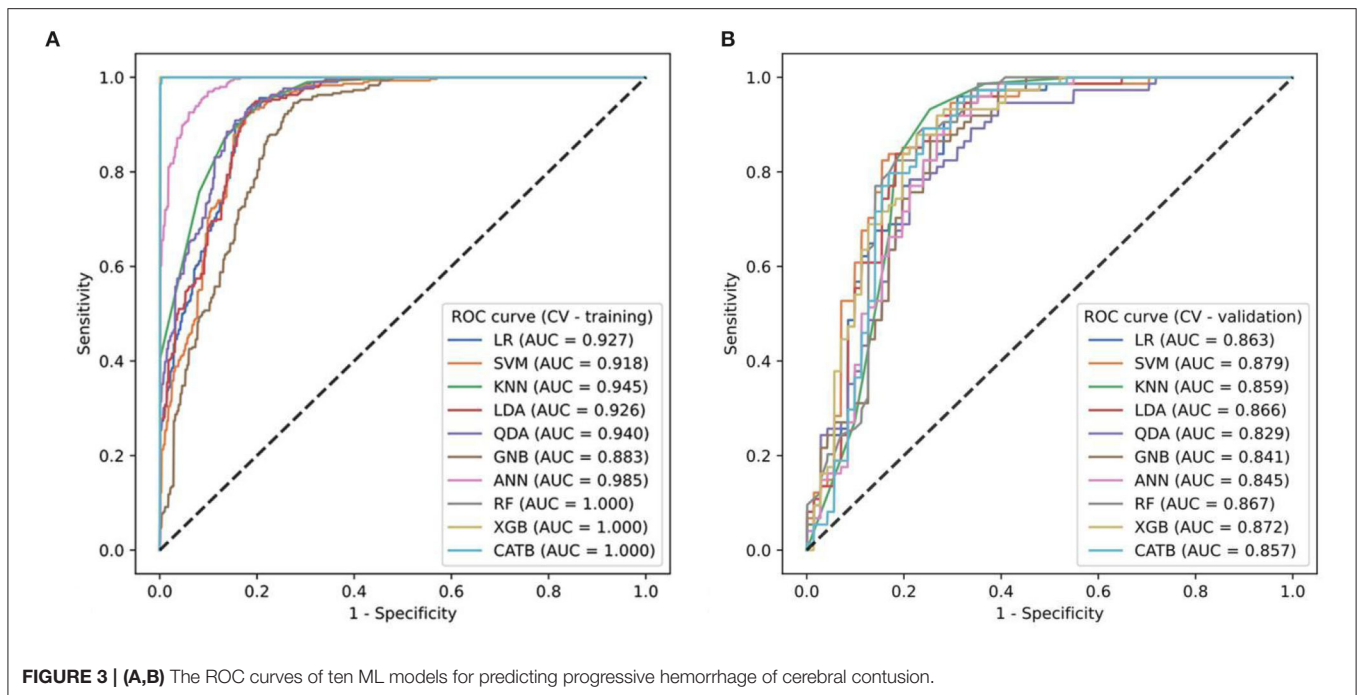
and testing sets; from the view of the decision curve, using the nomogram to predict the PH would be more beneficial than that without the Rad-score or the Radiomics model. Furthermore, only the combined model showed good results ($p = 0.4208 > 0.2$) when the HL test was conducted in the training set, and the p-values of the clinical-radiological and radiomics models were both < 0.05. In the testing set, the HL statistics indicated a poor fit for the clinical-radiological model ($p = 0.0158 < 0.05$). The radiomics and the combined models adequately fitted the data ($p = 0.1686$ and $p = 0.1834 > 0.05$).

DISCUSSION

In this study, we used 10 different MLs to construct the radiomics models and to select the best model. We established the other two models to predict PH, including the traditional clinical-radiological model and the combined model (integrating clinical-radiological factors with radiomics features) and we compared

the prediction performance of the three models. Our results suggest that the radiomics model is better than the clinical-radiological model, and the combined model is the best of all models. Finally, we established a convenient prediction nomogram based on the combined model to assist neurosurgeons and radiologists in identifying PH.

Many clinical and radiological factors have been demonstrated to be predictors of PH. Some studies (2, 3, 18, 21–24) have suggested that hypertension, the triglyceride level, D-dimer level, fibrin monomers, the onset-to-baseline CT time, the initial volume of contusion, the coexistence of SAH or SDH on baseline CT, skull fracture, and the "spot sign" formed by contrast extravasation on baseline CTA were independent predictors of PH. In this study, the first result we found was that fibrinogen was an independent predictor of PH, which might be attributed to the hyperfibrinolysis that is induced by a TBI. In the early stage of TBI, the blood is in a hypercoagulable state due to the activation of the coagulation pathway. However, the activation of

FIGURE 3 | (A,B) The ROC curves of ten ML models for predicting progressive hemorrhage of cerebral contusion.

TABLE 2 | Risk factors for progressive hemorrhage.

	Univariate logistic regression[b]		Multivariable logistic regression	
	Odds ratio	**p value**	**Odds ratio**	**p value**
Scalp hematoma	16.362 (2.079, 128.757)	0.008	10.382 (3.595, 29.978)	<0.0001*
GCS score, mean ± SD	1.0 (0.975, 1.025)	0.974	/	/
Onset-to-CT time[a], mean ± SD (h)	0.759 (0.611, 0.943)	0.013	0.728 (0.578, 0.915)	0.007*
Fibrinogen, mean ± SD (g/L)	0.607 (0.407, 0.905)	0.014	0.638 (0.453, 0.898)	0.010*

[a]Onset-to-CT time, Interval from trauma to first CT.
[b]Univariate LRs of the clinical factors on the labels were implemented one by one to check whether there were insignificant variables.
*The predictors included for the further analysis (p < 0.05).

the protein C system, the consumption of platelets, the decline of platelet function, the depletion of coagulation factors, and the activation of endothelial cells lead to hyperfibrinolysis, and fibrinogen decreases over time. Previous studies have confirmed that fibrinolysis is related to the progression of bleeding in TBI patients. As a retrospective study, the clinical data of some patients were incomplete (D-dimer, blood lipid, etc.). Therefore, the impact of these clinical data on PH could not be assessed, which is a limitation of this study. The second factor in the clinical-radiological model was time. The shorter the onset-to-baseline CT time was, the higher the incidence of PH, and it was an independent predictor of PH, which was consistent with previous studies (25). In addition, we also found that scalp hematoma was another independent predictor of PH. Some patients without a skull fracture but who had a scalp hematoma developed PH in our study. Scalp hematoma was also a manifestation of the severity of trauma and was a radiological factor that cannot be ignored. We did not find any literature report that scalp hematoma was an independent predictor of

PH, which might be because the previous relevant literature (2, 3, 5, 7, 26) did not include scalp hematoma in its studies. In the past, researchers had established some models to predict PH based on clinical and radiological factors, but the accuracy of these models was relatively low, ranging from 0.69 to 0.77 (2, 18, 27). Like our clinical-radiological model, a low accuracy meant that patients at risk of PH could not be accurately selected.

Radiomics analysis has great potential in hemorrhagic diseases and is helpful for the etiological diagnosis of intracerebral hemorrhage and for detecting hematoma expansion in cases of spontaneous intracerebral hemorrhage (28). However, it is unclear whether radiomics features can predict the development of traumatic PH. Our study indicated the value of radiomics signatures in detecting PH. We screened 12 out of 2,107 extracted radiomics features for modeling, including 3 first-order features, 6 GLCM features, 1 GLSZM feature, and 2 GLDM features. Among these radiomics features, the first-order features reflected the heterogeneity of the hematoma density by describing the voxel intensity distribution and internal variation degree. The
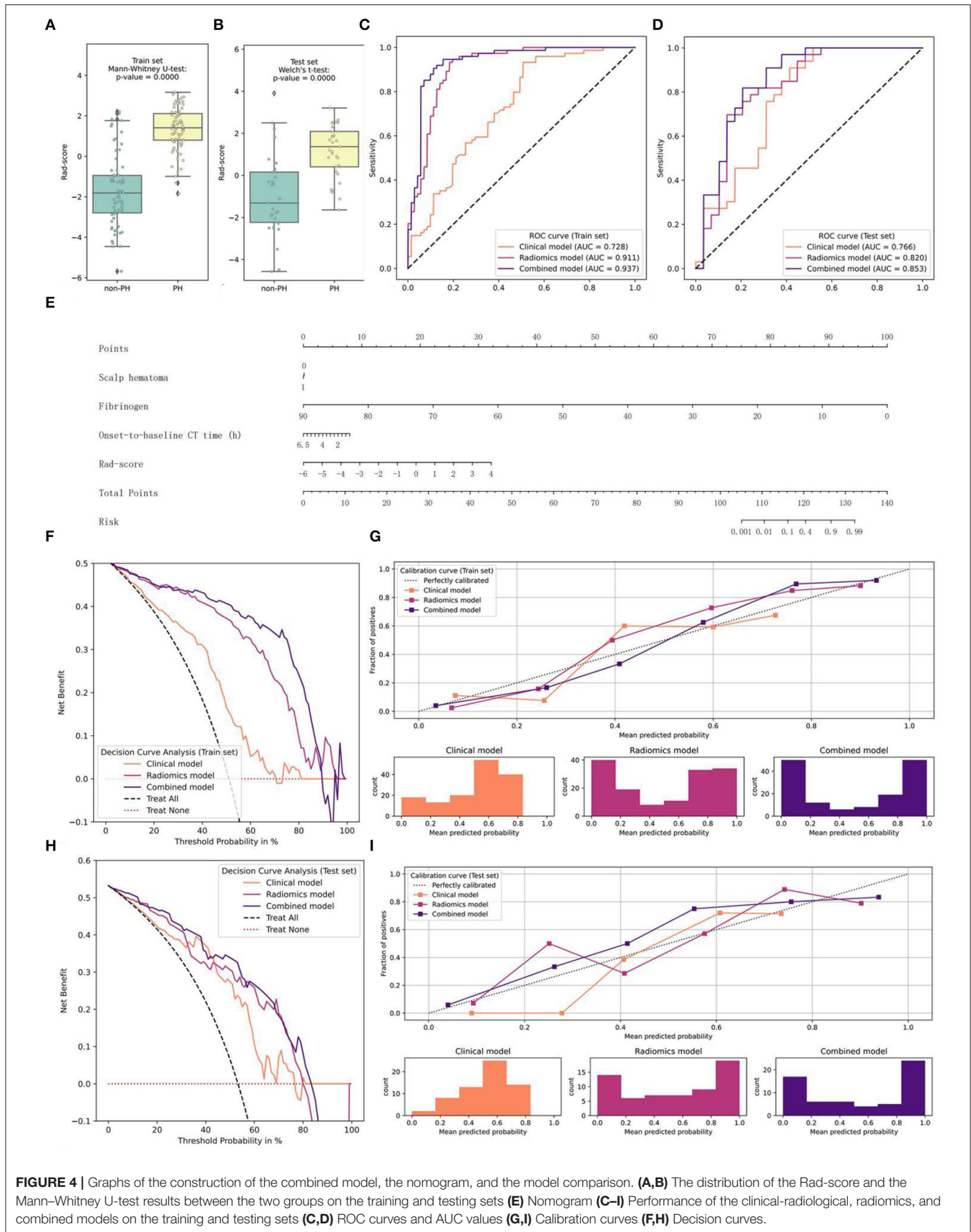
FIGURE 4 | Graphs of the construction of the combined model, the nomogram, and the model comparison. (A,B) The distribution of the Rad-score and the Mann–Whitney U-test results between the two groups on the training and testing sets (E) Nomogram (C–I) Performance of the clinical-radiological, radiomics, and combined models on the training and testing sets (C,D) ROC curves and AUC values (G,I) Calibration curves (F,H) Decision curves.
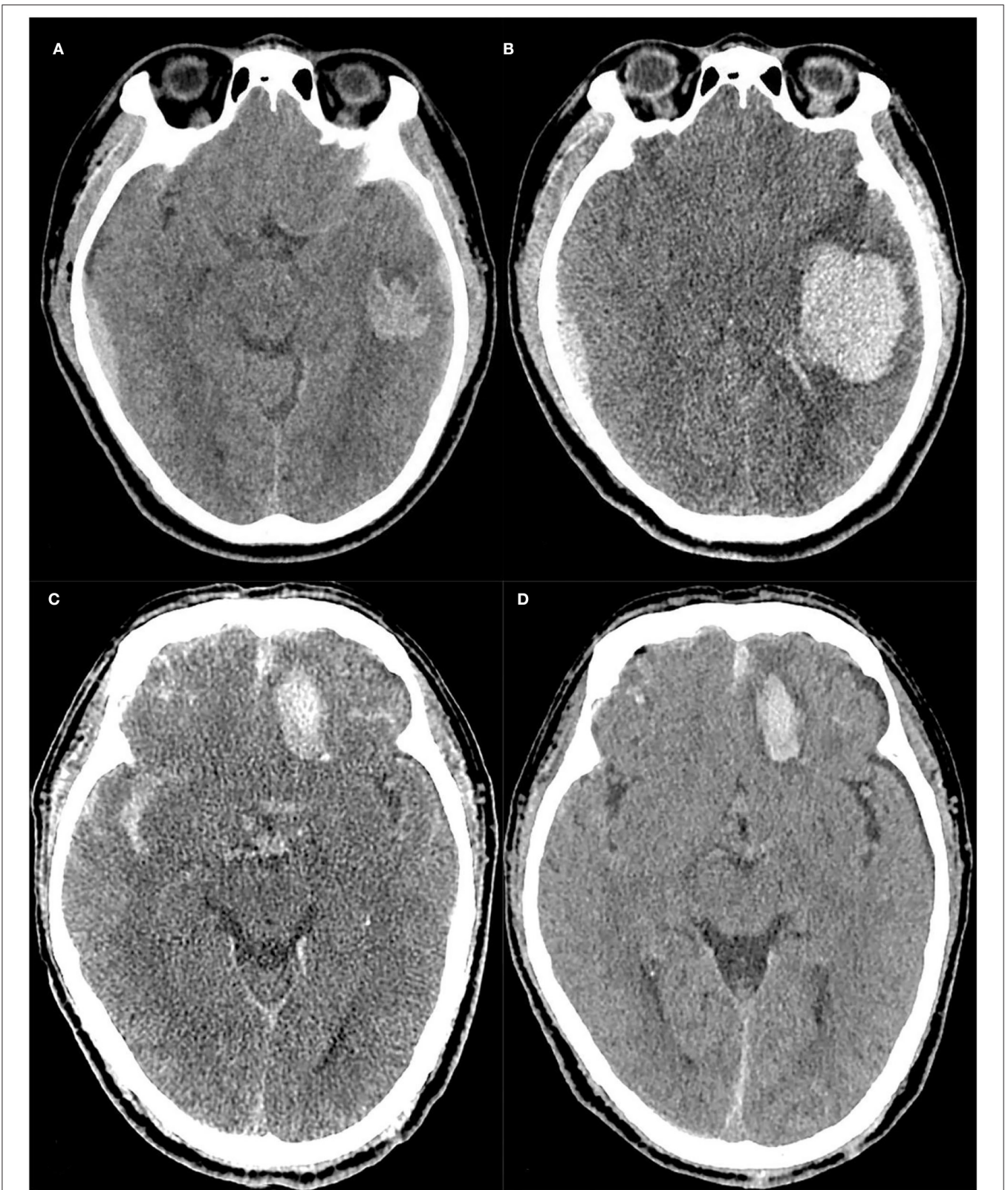
FIGURE 5 | Case 1: A case with progressive hemorrhage, The baseline and follow-up CT images of a TBI patient **(A,B)**, Scalp hematoma = 1, Fibrinogen = 2.68, Onset-to-CT time = 2 h, R-score = 1.925. The risk of PH calculated by the nomogram was approximately 94%, and the hematoma volume was expanded to approximately 42.9 ml on the follow-up 10-h CT compared to 10.5 ml on the baseline CT. Case 2: A case without progressive hemorrhage, The baseline and follow-up CT images of another TBI patient **(C,D)**, Scalp hematoma = 1, Fibrinogen = 3.01, Onset-to-CT time = 6 h, R-score =-1.144. The risk of PH calculated by the nomogram was approximately 0.044%. At 17 h after trauma, the hematoma volume had not expanded.

TABLE 3 | Comparisons of predictive models on the training and testing set.

Model	vs. the Combined model[b]	AUC (95% CI)	Accuracy	F1-score	Sensitivity	Specificity	Threshold[a]
Training set							
Clinical model	<0.00001****	0.728 (0.660–0.799)	0.717	0.771	0.905	0.493	0.427
Radiomics model	0.0618	0.911 (0.865–0.951)	0.876	0.885	0.905	0.817	0.376
Combined model	/	0.937 (0.897–0.970)	0.897	0.899	0.932	0.887	0.545
Testing set							
Clinical model	0.0179 *	0.766 (0.663–0.871)	0.726	0.779	0.909	0.517	/
Radiomics model	0.0456*	0.820 (0.724–0.906)	0.758	0.805	0.939	0.552	/
Combined model	/	0.853 (0.765–0.930)	0.806	0.818	0.818	0.793	/

[a]On the training set, a Youden index analysis was constructed to determine the optimal threshold (cut-off value) of the classification probability and was used to predict the outcomes based on the testing set.
[b]p-values of DeLong's test; * $p < 0.05$; **** $p < 0.0001$.

GLCM, GLSZM, and GLDM features reflected the heterogeneity of the hematoma density by describing the irregular texture features of the hematoma. The change in the hematoma density depended on the time course of bleeding. When progressive hemorrhage occurs and when there is a greater change in the hematoma density, there is a greater heterogeneity of the hematoma density. Therefore, brain trauma with PH has greater heterogeneity than stable brain trauma. A previous study (29) has suggested that the attenuation of a heterogeneous hematoma is a predictor of hematoma expansion and might reflect active bleeding. Then, we applied 10 ML algorithms to establish the omics model. Compared with the traditional models such as LR, the ML algorithms are better at processing multidimensional features, because, from the data, they can identify some potential patterns that in most scenarios are not linear or polynomial, so LR is not always the most reasonable choice. Our study supports these data (**Figure 3**). The SVM classifier we finally chose used a sigmoid kernel function to map our data to higher dimensions and then distinguished it with a hyperplane that allows the model to learn richer non-linear patterns of the relationship between the PH group and the non-PH group. Thus, our results revealed that the radiomics model had a better discrimination ability than the traditional clinical-radiological model.

To find the best predictive PH model, we constructed a combined model integrating the Rad-score with clinical-radiological factors. As we expected, compared with the clinical-radiological and radiomics models, the combined model showed the best discrimination and sensitivity in detecting PH. The nomogram was a more objective, convenient, and rapid personalized tool to predict PH based on the combined model (**Figure 5**). It helps neurosurgeons to identify PH patients and to revise treatment schemes in time, reduces mortality, and improves the prognosis of patients. At the same time, it solves the problem of blind repeat CT and reduces the waste of medical resources and unnecessary radiation exposure. At present, our research has some limitations. First, it is a single-center retrospective study, and it lacks external validation of data from other centers. Second, because some patients had multiple lesions, we assigned data belonging to the same patient to the same set to avoid data leakage, but the fly in the ointment is that

these samples may cause a local autocorrelation in the model. Finally, when a hematoma is in the cerebral cortex, there may be an inaccurate separation on CT due to the partial solvent effect.

CONCLUSIONS

In this study, the radiomics model outperformed the traditional clinical-radiological model in predicting PH. A nomogram model of combined radiomics features and clinical-radiological factors could offer an efficient and convenient tool for PH.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Medical Ethics Committee of Chongqing University Central Hospital. Written informed consent from the patients/participants or patients/participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

YG, QY, PZ, and JS designed and managed. QY, JS, LH, and JF collected the clinical and CT data. KJ, CH, and JX analyzed all data and technical support. YG and QY drafted the manuscript. CL, PZ, and JS revised the manuscript. YG, QY, PZ, JS, and CL approved for the final manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

REFERENCES

1. Kaur P, Sharma S. Recent advances in pathophysiology of traumatic brain injury. *Curr Neuropharmacol.* (2018) 16:1224–38. doi: 10.2174/1570159X15666170613083606

2. Allison RZ, Nakagawa K, Hayashi M, Donovan DJ, Koenig MA. Derivation of a predictive score for hemorrhagic progression of cerebral contusions in moderate and severe traumatic brain injury. *Neurocrit Care.* (2017) 26:80–6. doi: 10.1007/s12028-016-0303-5

3. Wan X, Fan T, Wang S, Zhang S, Liu S, Yang H, et al. Progressive hemorrhagic injury in patients with traumatic intracerebral hemorrhage: characteristics, risk factors and impact on management. *Acta Neurochir (Wien).* (2017) 159:227–35. doi: 10.1007/s00701-016-3043-6

4. X Wan, Gan C, You C, Fan T, Zhang S, Zhang H, et al. Association of APOE ε4 with progressive hemorrhagic injury in patients with traumatic intracerebral hemorrhage. *J Neurosurg.* (2019) 19:1–8. doi: 10.3171/2019.4.JNS183472

5. Carnevale JA, Segar DJ, Powers AY, Shah M, Doberstein C, Drapcho B, et al. Blossoming contusions: identifying factors contributing to the expansion of traumatic intracerebral hemorrhage. *J Neurosurg.* (2018) 129:1305–16. doi: 10.3171/2017.7.JNS17988

6. Juratli TA, Zang B, Litz RJ, Sitoci KH, Aschenbrenner U, Gottschlich B, et al. Sobottka, Early hemorrhagic progression of traumatic brain contusions: frequency, correlation with coagulation disorders, and patient outcome: a prospective study. *J Neurotrauma.* (2014) 31:1521–7. doi: 10.1089/neu.2013.3241

7. Cepeda S, Gómez PA, Castaño-Leon AM, Martínez-Pérez R, Munarriz PM, Lagares A. Traumatic intracerebral hemorrhage: risk factors associated with progression. *J Neurotrauma.* (2015) 32:1246–53. doi: 10.1089/neu.2014.3808

8. Prasanna P, Patel J, Partovi S, Madabhushi A, Tiwari P. Radiomic features from the peritumoral brain parenchyma on treatment-naïve multi-parametric MR imaging predict long versus short-term survival in glioblastoma multiforme: preliminary findings. *Eur Radiol.* (2017) 27:4188–97. doi: 10.1007/s00330-016-4637-3

9. Kang D, Park JE, Kim YH, Kim JH, Oh JY, Kim J, et al. Diffusion radiomics as a diagnostic model for atypical manifestation of primary central nervous system lymphoma: development and multicenter external validation. *Neuro Oncol.* (2018) 20:1251–61. doi: 10.1093/neuonc/noy021

10. Li Y, Liu X, Qian Z, Sun Z, Xu K, Wang K. X. Genotype prediction of ATRX mutation in lower-grade gliomas using an MRI radiomics signature. *Eur Radiol.* (2018) 28:2960–8. doi: 10.1007/s00330-017-5267-0

11. Kim JY, Park JE, Jo Y, Shim WH, Nam SJ, Kim JH, et al. Incorporating diffusion- and perfusion-weighted MRI into a radiomics model improves diagnostic performance for pseudoprogression in glioblastoma patients. *Neuro Oncol.* (2019) 21:404–14. doi: 10.1093/neuonc/noy133

12. Liu J, Xu H, Chen Q, Zhang T, Sheng W, Huang Q, et al. Prediction of hematoma expansion in spontaneous intracerebral hemorrhage using support vector machine. *EBioMedicine.* (2019) 43:454–9. doi: 10.1016/j.ebiom.2019.04.040

13. Zhang D, Chen J, Xue Q, Du B, Li Y, Chen T, et al. Heterogeneity signs on noncontrast computed tomography predict hematoma expansion after intracerebral hemorrhage: a meta-analysis. *Biomed Res Int.* (2018) 2018:6038193. doi: 10.1155/2018/6038193

14. Tong X, Feng X, Peng F, Niu H, Zhang B, Yuan F, et al. Morphology-based radiomics signature: a novel determinant to identify multiple intracranial aneurysms rupture. *Aging (Albany NY).* (2021) 13:13195–210. doi: 10.18632/aging.203001

15. Nawabi J, Kniep H, Kabiri R, Broocks G, Faizy TD, Thaler C, et al. Neoplastic and non-neoplastic acute intracerebral hemorrhage in CT brain scans: machine learning-based prediction using radiomic image features. *Front Neurol.* (2020) 11:285. doi: 10.3389/fneur.2020.00285

16. Zhang J, Zhang F, Dong JF. Coagulopathy induced by traumatic brain injury: systemic manifestation of a localized injury. *Blood.* (2018) 131:2001–6. doi: 10.1182/blood-2017-11-784108

17. Jiang JY, Zhu C, Luo QZ. *Clinical Guidelines for the Management of Head Injury.* 3rd ed 353 Shanghai: Second Military Medical University Publisher. (2008).

18. Esnault P, Mathais Q, D'Aranda E, Montcriol A, Cardinale M, Cungi PJ. P, et al. Ability of fibrin monomers to predict progressive hemorrhagic injury in patients with severe traumatic brain injury. *Neurocrit Care.* (2020) 33:182–95. doi: 10.1007/s12028-019-00882-6

19. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* (2017) 77:e104–7. doi: 10.1158/0008-5472.CAN-17-0339

20. P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods.* (2020) 17:261–72. doi: 10.1038/s41592-020-0772-5

21. Chen S, Zhao B, Wang W, Shi L, Reis C, Zhang J. Predictors of hematoma expansion predictors after intracerebral hemorrhage. *Oncotarget.* (2017) 8:89348–63. doi: 10.18632/oncotarget.19366

22. Orito K, Hirohata M, Nakamura Y, Yamamoto M, Takeshige N, Aoki T. G, et al. Predictive value of leakage signs for pure brain contusional hematoma expansion. *J Neurotrauma.* (2018) 35:760–6. doi: 10.1089/neu.2017.5247

23. Adatia K, Newcombe VFJ, Menon DK. Contusion progression following traumatic brain injury: a review of clinical and radiological predictors, and influence on outcome. *Neurocrit Care.* (2021) 34:312–24. doi: 10.1007/s12028-020-00994-4

24. Yuan F, Ding J, Chen H, Guo Y, Wang G, Gao WW, et al. Tian, Predicting progressive hemorrhagic injury after traumatic brain injury: derivation and validation of a risk score based on admission characteristics. *J Neurotrauma.* (2012) 29:2137–42. doi: 10.1089/neu.2011.2233

25. Narayan RK, Maas AI, Servadei F, Skolnick BE, Tillinger MN, Marshall LF. Progression of traumatic intracerebral hemorrhage: a prospective observational study. *J Neurotrauma.* (2008) 25:629–39. doi: 10.1089/neu.2007.0385

26. Rehman L, Afzal A, Aziz HF, Akbar S, Abbas A, Rizvi R. Radiological parameters to predict hemorrhagic progression of traumatic contusional brain injury. *J Neurosci Rural Pract.* (2019) 10:212–7. doi: 10.4103/jnrp.jnrp_335_18

27. Fair KA, Farrell DH, McCully BH, Rick EA, Dewey EN, Hilliard C, et al. Fibrinolytic activation in patients with progressive intracranial hemorrhage after traumatic brain injury. *J Neurotrauma.* (2021) 38:960–6. doi: 10.1089/neu.2018.6234

28. Morotti, G. Boulouis, D. Dowlatshahi, Q. Li, C.D. Barras, C. Delcourt, Z, et al. Standards for detecting, interpreting, and reporting noncontrast computed tomographic markers of intracerebral hemorrhage expansion. *Ann Neurol.* 86 (2019) 480–92. doi: 10.1002/ana.25563

29. Barras CD, Tress BM, Christensen S, Collins M, Desmond PM, Skolnick BE, et al. Quantitative CT densitometry for predicting intracerebral hemorrhage growth. *AJNR Am J Neuroradiol.* (2013) 34:1139–44. doi: 10.3174/ajnr.A3375