Supplementary

## Methods

### *Patients*

We retrospectively reviewed data for 258 patients hospitalized for a spontaneous intracerebral hemorrhage (ICH) at the Central Hospital of Wuhan (CHW) as either training or internal validation cohort, 87 ICH patients at the Fifth Affiliated Hospital of Nanchang University (FAHNU) were prospectively enrolled as an external test cohort. Patient selection process is shown in *Figure S1*.

## Results

### *Establishment of the rad-score*

The least absolute shrinkage and selection operator method (LASSO) is a popular method for the regression of high-dimensional predictors (34,35)_ENREF_6. The method uses an L1 penalty to shrink some regression coefficients to exactly zero (*Figure S2*). Rad-score = (∑βi*Xi) + Intercept (i=0, 1, 2, 3……), where Xi represents the $i^{th}$ selected feature and βi is its

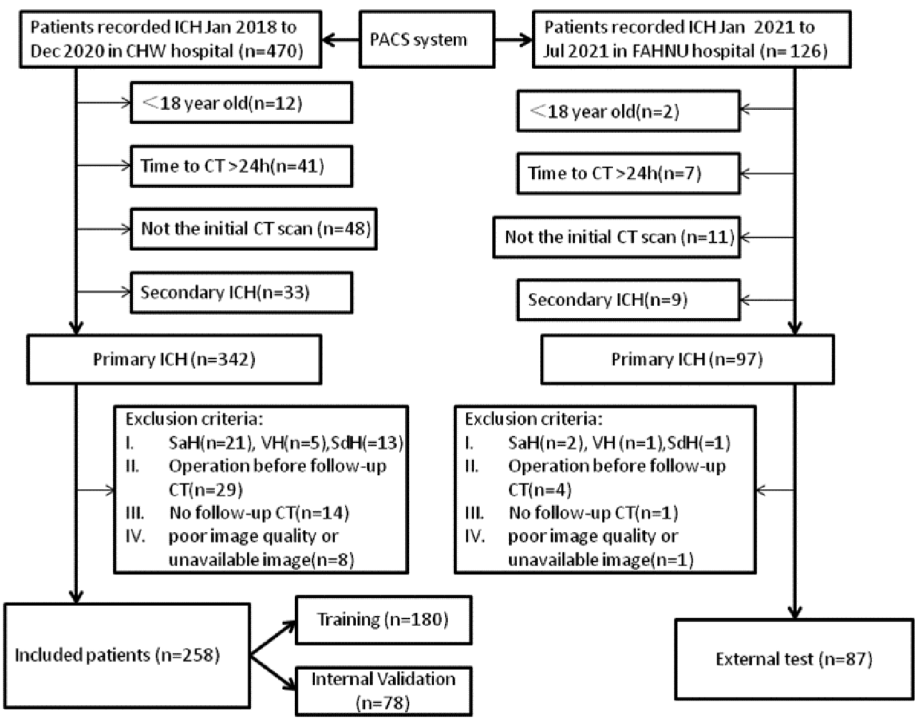


**Figure S1** Flowchart of the patient selection process. ICH, intracerebral hemorrhage; CT, computed tomography; CHW, Central Hospital of Wuhan; FAHNU, Fifth Affiliated Hospital of Nanchang University.

© Quantitative Imaging in Medicine and Surgery. All rights reserved.
https://dx.doi.org/10.21037/qims-22-128

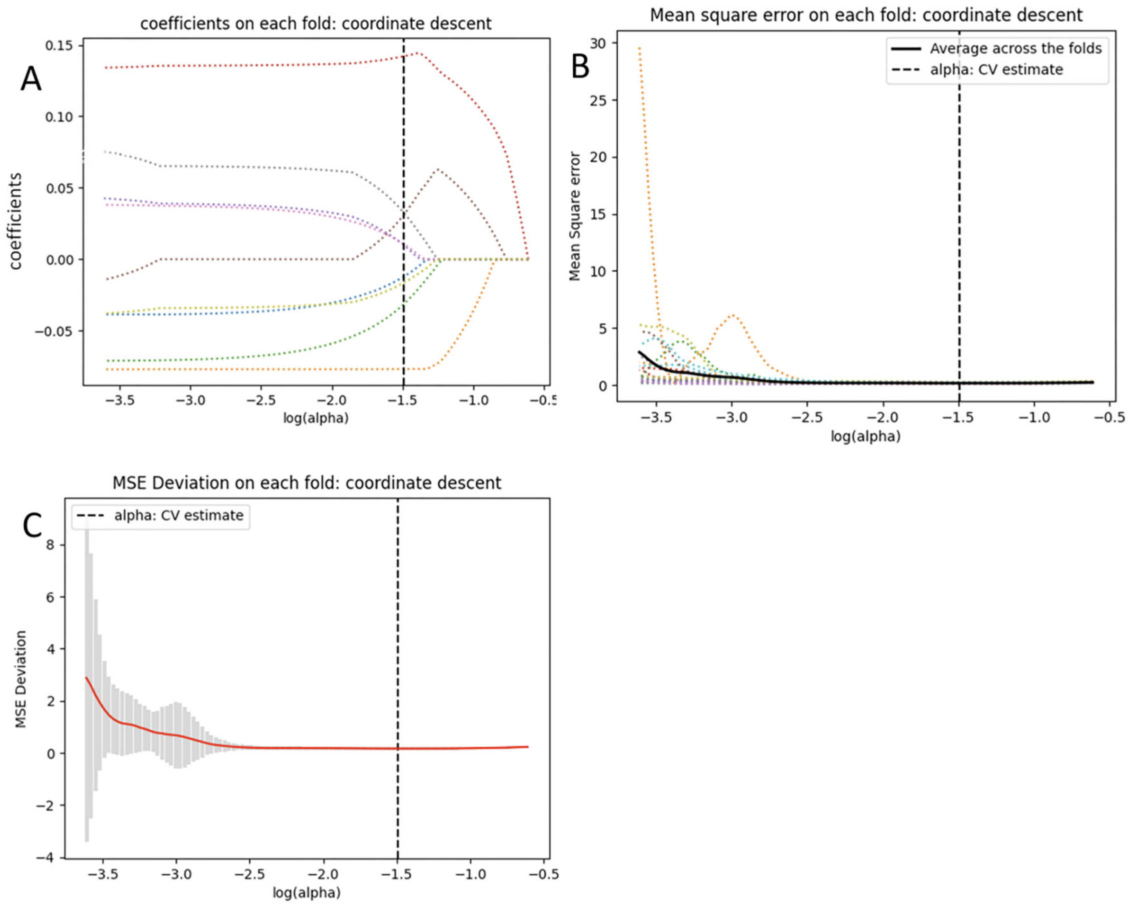


**Figure S2** Texture feature selection using the least absolute shrinkage and selection operator (LASSO) logistic regression models. (A) LASSO coefficient profiles of the selected 9 features among 1072 features. (B) Mean square error on each fold. (C) Deviation of mean square error on each fold.

coefficient. The radiomic score formula is as follows:

Radscore = −0.27896054*log-sigma-2-0-mm-3D_ firstorder_RootMeanSquared
+ −0.445686359*log-sigma-2-0-mm-3D_gldm_DependenceVariance
+ −0.421542144*log-sigma-3-0-mm-3D_glcm_ClusterTendency
+ 0.819735954*original_shape_Maximum2DDiameterColum
+ 0.528278381*original_shape_MinorAxisLength
+ 0.120289332*wavelet-HHH_gldm_LargeDependenceEmphasis
+ 0.248174582*wavelet-HHH_glrlm_LongRunHighGrayLevelEmphasis
+ 0.236439734*wavelet-LHH_gldm_LargeDependenceEmphasis
+ −0.251098979*wavelet-LLH _gldm_DependenceVariance+0.39407328

### *Model construction, calibration and validation*

All 258 SICH patients from CHW were ransomly divided into a training dataset (n=180) and an internal validation dataset (n=78) according to a 7:3 ratio for models establishment and validation, finally 87 SICH patients from FAHNU were enrolled for models test. Three models, including radimics (rad-score based), clinical (clinical factors based), hybrid (rad-score combined with clinical factors) models were established respectively. Nomogram of hybrid model was constructed.

© Quantitative Imaging in Medicine and Surgery. All rights reserved. https://dx.doi.org/10.21037/qims-22-128

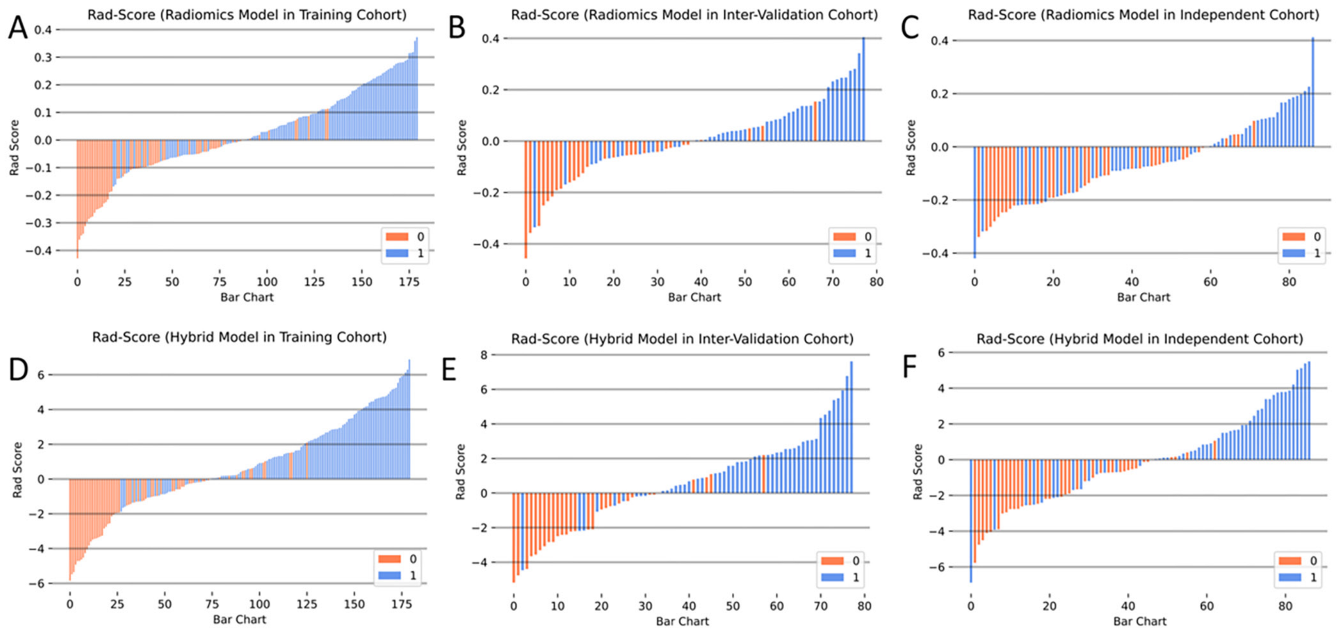


**Figure S3** Bar chart of Rad-Score in the training (left), internal validation (middle) and external testing (right) cohorts. (A-C) Rad-Score model; (D-F) hybrid model. The y-axis refers to the Rad-Score minus the optimal cutoff value. Upper and lower bars refer to the predicted positive and negative poor outcomes, respectively. Blue and orange bars refer to actual positive and negative poor outcomes, respectively.

## Models construction

*Rad-score model:* According to the dichotomy criterion of mRS 0-3 or mRS 4-6, the Rad-score model was established in the training cohort by the following equations: logit $\pi_{\text{Rad-Score}}$=1.002+0.884×Rad-Score (*Figure S3*).

*Clinical model:* We firstly used univariate logistic regression analysis for screening clinical independent risk factors. Then, multivariate logistic regression model was applied for clinical-based model construction according to those independent risk factors in training cohort.

*Hybrid nomogram:* Based on Rad-Score and independent clinical risk factors, a hybrid model nomogram was established for poor outcome prediction using multivariate logistic regression (*Figure S3*).

## Models calibration

*Discrimination:* The AUCs under ROCs were used to assess the predictive performances of Rad-Score based, clinical-based and hybrid models in discriminating SICH patients with poor outcome(mRS 4-6) from those with mRS 0-3.

*Calibration:* A calibration curve was plotted in the training, internal validation and the independent test cohorts for the purpose of examining the agreement between the observed outcomes and predicted probabilities. Hosmer-leme show test was performed to test the calibration, and decision curve analysis (DCA) to evaluate the clinical net benefit of the models (*Figure S4*).

## Models validation

Rad-Score model, clinical model and hybrid model nomogram constructed in training cohort (CHW cohort1) were introduced into internal validation cohort (CHW cohort2), external validation cohort (FAHNU cohort) to validate respectively.

© Quantitative Imaging in Medicine and Surgery. All rights reserved.

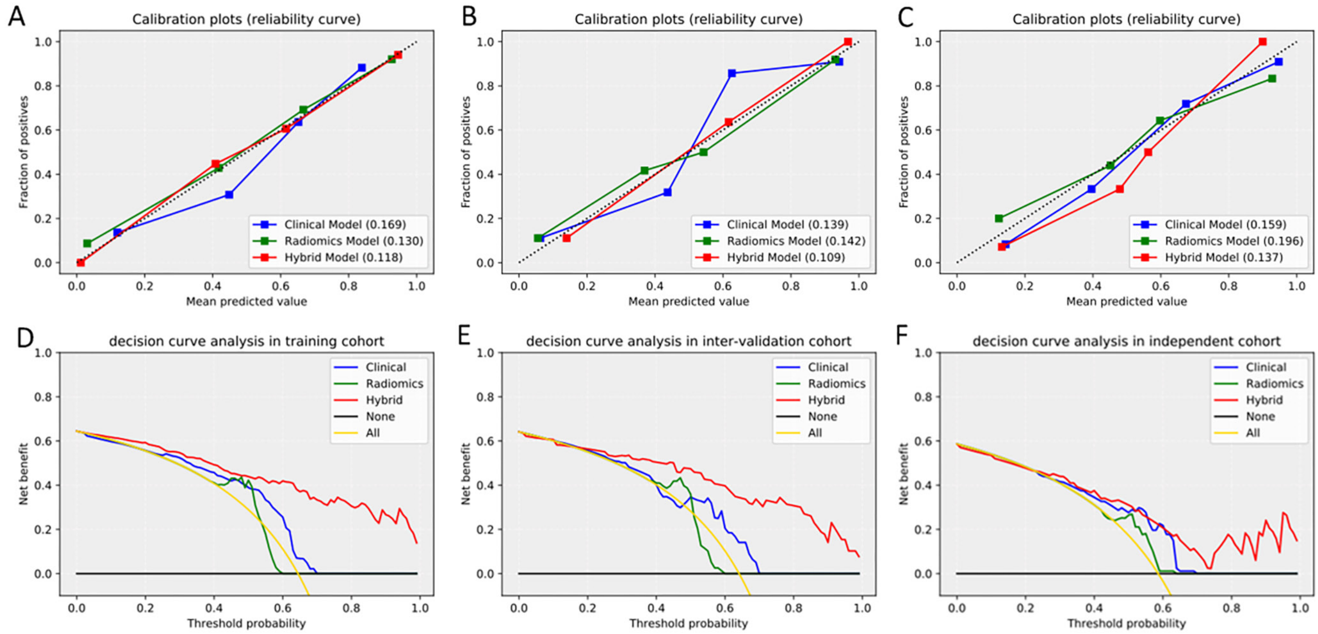


**Figure S4** Calibration curves and decision curve analysis (DCA) of three models in the training (left), internal validation (middle) and external testing (right) cohorts. Calibration curves (A-C) depict agreement between the predicted risks of a poor outcome and actual observed poor outcomes following ICH. The solid blue, green and red lines represent good predictive ability of a poor outcome using the clinical, Rad-Score and hybrid models, respectively. DCA (D-F) for clinical, Rad-Score and hybrid model in the training (left), internal validation (middle) and external testing (right) cohorts. The yellow curve represents the assumption that all will have a poor outcome. The black line represents the assumption that none will have a poor outcome.

## References


34. Jiang Y, He Y, Zhang H. Variable Selection with Prior Information for Generalized Linear Models via the Prior LASSO Method. J Am Stat Assoc 2016;111:355-76.
35. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med 1997;16:385-95.


© Quantitative Imaging in Medicine and Surgery. All rights reserved. https://dx.doi.org/10.21037/qims-22-128