Appendix 1: methods

Imaging protocol

All patients underwent an axial baseline non-contrast CT of the head with whole-brain coverage and a maximum section thickness of 5 mm with a 5 mm section interval. Non-contrast CT parameters were set as follows: tube voltage, 120 kV; tube current, 380 mA; rotation time, 1.00 s; section thickness, 5 mm; detector coverage, 80 mm. The acquisition duration was 3.7 s. CTP was performed 5 min delay after mCTA with the following parameters: 100 mm coverage in the Z-axis; tube voltage, 80 kV; tube current, 120 mA; section thickness and slice thickness, 5 mm. 24 consecutive spiral acquisitions with 0.28 s of rotation time and 1.70 s of inter-scan delay time were performed. After CTP imaging, patients were asked to stay on the scanning table for 5 min before undergoing the required mCTA examination. The mCTA scans consisted of three phases, we scanned patients from arch to vertex coverage in the arterial phase, and skull base to vertex coverage in the venous phase and in the late venous phase. The acquisition parameters were as follows: tube voltage, 100 kV; fixed tube current, 445 mA; pitch factor 0.922, rotation time 0.6 s, display field of view 41.3 cm × 24.0 cm, and image matrix 512×512. A total of 45 mL intravenous contrast agent (iopromide, GE HealthCare, 370 mgI/mL) followed by a 30 mL saline flush at 5 mL/s. The dynamic monitoring initiated 8 s after the intravenous contrast injection. The bolus-tracking was used for image acquisition with a region of interest in the descending aorta and a trigger at 140 Hounsfield units (HU) in first phase. The second and third phases were acquired with 10- and 18-s delay. Iopromide (GE HealthCare, 370 mgI/mL) was used for mCTA and CTP examination.

Statistical analysis

The reproducibility of measurements for CSCC was evaluated by using weighted kappa (κ) analysis. Spearman correlation analyses were used to analyze the relationship between CSCC, HIR, and NIHSS at different TP. Post-hoc sample size was calculated by using power analysis with 90% power and a 5% significance level.

Appendix 2: results

Post-hoc sample size calculation

Based on the maximal coefficient of variation value of 7.8% in our reproducibility analysis, to detect 5% difference in NIHSS scores, the required sample size was 104 (90% power, 0.05 significance); to detect 10% difference, the required sample size was 26. As the difference between groups in our results were all >5% our study had sufficient sample size to draw safe conclusions.

Intra- and inter-observer reproducibility for CSCC assessment

Intra-observer reproducibility was almost perfect for CSCC (weighted κ=0.947; 95% CI: 0.905, 0.990; P<0.001). Inter-observer reproducibility was substantial for CSCC (weighted κ=0.859; 95% CI: 0.791, 0.926; P<0.001).

Predictive performance of variables for predicting clinical outcome in total cohort

In the total cohort, HIR achieved the highest AUC (AUC =0.731; 95% CI: 0.644, 0.818), with sensitivity of 0.484, specificity of 0.922, PPV of 0.861, and NPV of 0.641. The comparison of predictive performance of age, ASPECTS, CSCC, and HIR are listed in *Table S3* and *Figure S3*.

Five-fold validation of RF model

RF model and five-fold cross-validation were performed on the training set and testing set. The results showed that the

https://dx.doi.org/10.21037/qims-24-1073

average AUC of the training set was 1 (95% CI: 1, 1), and the AUC of the testing set was 0.801 (95% CI: 0.703, 0.878).

Comparison between comparison of predictive performance for clinical outcome between HIRs with different threshold

ROC analysis revealed the AUC of Tmax 10 s/6 s is significantly higher than that of Tmax 10 s/4 s (0.731 *vs*. 0.670; Z=1.996; P=0.046). The ROC curves and DeLong's test were shown in *Figure S4*.

Association between CSCC, HIR, and NIHSS at different TPs

In CSCC stratification, patients with CSCC ≥3 showed lower NIHSS in 72 hours (13±8 *vs*. 17±9 hours, P=0.02) and 7 days (12±8 *vs*. 16±9 days, P=0.02) compared with patients in CSCC <3 subgroup, but no statistical difference in NIHSS of baseline, 24 hours, 48 hours, and discharge. In addition, there were significant differences on baseline and at all TPs after EVT in HIR stratification (P=0.006 on baseline, and P<0.001 at other TPs).

ΔNIHSS was calculated as subtracting baseline NIHSS from NIHSS at discharge. There is no relationship between CSCC and baseline NIHSS (r=0.161, P=0.07), discharge NIHSS (r=0.165, P=0.06), and ΔNIHSS (r=0.096, P=0.28). HIR has moderate correlation to discharge NIHSS (r=0.421, P<0.001) and ΔNIHSS (r=0.301, P<0.001), and weak correlation to baseline NIHSS (r=0.196, P=0.04). Detailed results are outlined in *Figure S6*. The associations among CSCC, HIR, and NIHSS were shown in *Figure S7*.

https://dx.doi.org/10.21037/qims-24-1073

Table S1 The baseline characteristics of training set and testing set

Characteristics	All patients (n=128)	Training set (n=89)	Testing set (n=39)	P
Demographic characteristics				
Age (years)	64.0 [57.3, 69.0]	63.0 [56.5, 68.5]	66.0 [60.0, 70.0]	0.150
Men	97 (75.8)	69 (77.5)	28 (71.8)	0.507
Hypertension	86 (67.2)	57 (64.0)	29 (74.4)	0.309
Diabetes mellitus	25 (19.5)	14 (15.7)	11 (28.2)	0.145
Dyslipidemia	42 (32.8)	29 (32.6)	13 (33.3)	0.838
Atrial fibrillation	21 (16.4)	13 (14.6)	8 (20.5)	0.442
Smoking history	62 (48.4)	45 (50.6)	17 (43.6)	0.565
Alcohol intake	49 (38.3)	38 (42.7)	11 (28.2)	0.166
Ischemic heart disease	26 (20.3)	19 (21.3)	7 (17.9)	0.812
Previous stroke	20 (15.6)	16 (18.0)	4 (10.3)	0.303
Clinical information				
Onset time (hours)	11.0 [8.0, 16.8]	12.0 [7.5, 17.5]	10.0 [8.0, 13.0]	0.827
Thrombolysis before EVT	11 (8.6)	6 (6.7)	5 (12.8)	0.308
Modified thrombolysis in cerebral infarction score ≥ 2b	106 (82.8)	77 (86.5)	29 (74.4)	0.126
NIHSS score on admission	13.6±5.4	13.7±5.7	13.5±4.3	0.846
Symptomatic intracranial hemorrhage	40 (31.3)	37 (41.6)	13 (33.3)	0.435
Good clinical outcome (90-day mRS, 0–2)	64 (50.0)	45 (50.6)	19 (48.7)	>0.99
Imaging features				
Tandem occlusion	27 (21.1)	20 (22.5)	7 (17.9)	0.643
Clot burden	6 [6, 8]	6 [6, 8]	6 [6, 8]	0.449
CSCC				0.638
5	0 (0.0)	0 (0.0)	0 (0.0)	
4	23 (18.0)	14 (15.7)	9 (23.1)	
3	75 (58.6)	46 (51.7)	19 (48.7)	
2	30 (23.4)	22 (24.7)	7 (17.9)	
1	11 (8.6)	7 (7.9)	4 (10.3)	
0	0 (0.0)	0 (0.0)	0 (0.0)	
ASPECTS	7 [6, 8]	7 [6, 8]	7 [6, 9]	0.827
Tmax (mL)				
>4 s volume	320 [227, 456]	320 [232, 450]	316 [222, 503]	0.707
>6 s volume	157 [104, 222]	157 [107, 216]	155 [97, 223]	0.995
>8 s volume	92 [38, 142]	91 [41, 146]	96 [34, 134]	0.873
>10 s volume	50 [13, 89]	52 [14, 97]	50 [6, 82]	0.471
HIR	0.310 [0.166, 0.487]	0.304 [0.166, 0.508]	0.325 [0.163, 0.449]	0.477

Data are presented a n (%) for categorical variables and median [IQR] for continuous variables, except for NIHSS score on admission and at discharge, which is presented as mean ± SD. ASPECTS, Alberta Stroke Program Early CT Score; CSCC, Collateral Score on Color-Coded summation maps; EVT, endovascular thrombectomy; HIR, hypoperfusion intensity ratio; IQR, interquartile range; mRS, modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale; SD, standard deviation; Tmax, time to maximum.

https://dx.doi.org/10.21037/qims-24-1073

Table S2 Univariate and multivariate LR analysis of risk factors for good functional outcome in training set

Characteristics	Univariate logistic analysis			Multivariate logistic analysis		
	OR	95% CI	P value	OR	95% CI	P value
Age	1.089	1.031–1.151	0.002	1.073	1.008–1.154	0.040
HT	2.240	0.946–5.307	0.067	–	–	–
ASPECTS	0.706	0.552–0.901	0.005	0.742	0.546–0.975	0.040
Degree of recanalization after EVT	0.625	0.384–1.019	0.059	–	–	–
CSCC	0.365	0.195–0.684	0.002	0.468	0.213–0.953	0.044
Tmax						
>4 s volume	1.004	1.000–1.008	0.020	–	–	–
>6 s volume	1.010	1.004–1.016	0.001	–	–	–
>8 s volume	0.012	1.004–1.020	0.001	–	–	–
>10 s volume	0.016	1.006–1.026	0.001	–	–	–
HIR	96.207	8.839–1,047.097	<0.001	56.666	3.843–1,156.959	0.005

ASPECTS, Alberta Stroke Program Early CT Score; CI, confidence interval; CSCC, Collateral Score on Color-Coded summation maps; EVT, endovascular thrombectomy; HIR, hypoperfusion intensity ratio; HT, XXXXXXXX; LR, logistic regression; OR, odds ratio; Tmax, time to maximum;

Table S3 ROC analysis for 90-day mRS in total cohort

Characteristics	Cutoff	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	Accuracy
Variables							
Age	57.5	0.658 (0.564–0.752)	0.875	0.375	0.583	0.750	0.625
ASPECTS	6.5	0.649 (0.556–0.743)	0.500	0.719	0.640	0.590	0.609
CSCC	2.5	0.691 (0.608–0.774)	0.844	0.469	0.614	0.750	0.656
HIR	0.459	0.731 (0.644–0.818)	0.484	0.922	0.861	0.641	0.703
RF model							
Model 1	0.512	0.825 (0.754–0.896)	0.688	0.875	0.846	0.737	0.781
Model 2	0.458	0.969 (0.939–0.999)	0.922	0.953	0.952	0.924	0.938
Model 3	0.335	0.931 (0.890–0.973)	0.937	0.781	0.811	0.926	0.859

A RF model consisting of age, ASPECTS, and CSCC was defined as Model 1. Model 2 included age, ASPECTS, and HIR. The combination of age, ASPECTS, CSCC, and HIR was considered as Model 3. *, P<0.05. ASPECTS, Alberta Stroke Program Early CT Score; AUC, area under the ROC curve; CI, confidence interval; CSCC, Collateral Score on Color-Coded summation maps; HIR, hypoperfusion intensity ratio; mRS, modified Rankin Scale; NPV, negative predictive value; PPV, positive predictive value; RF, random forest; ROC, receiver operator characteristic.

https://dx.doi.org/10.21037/qims-24-1073

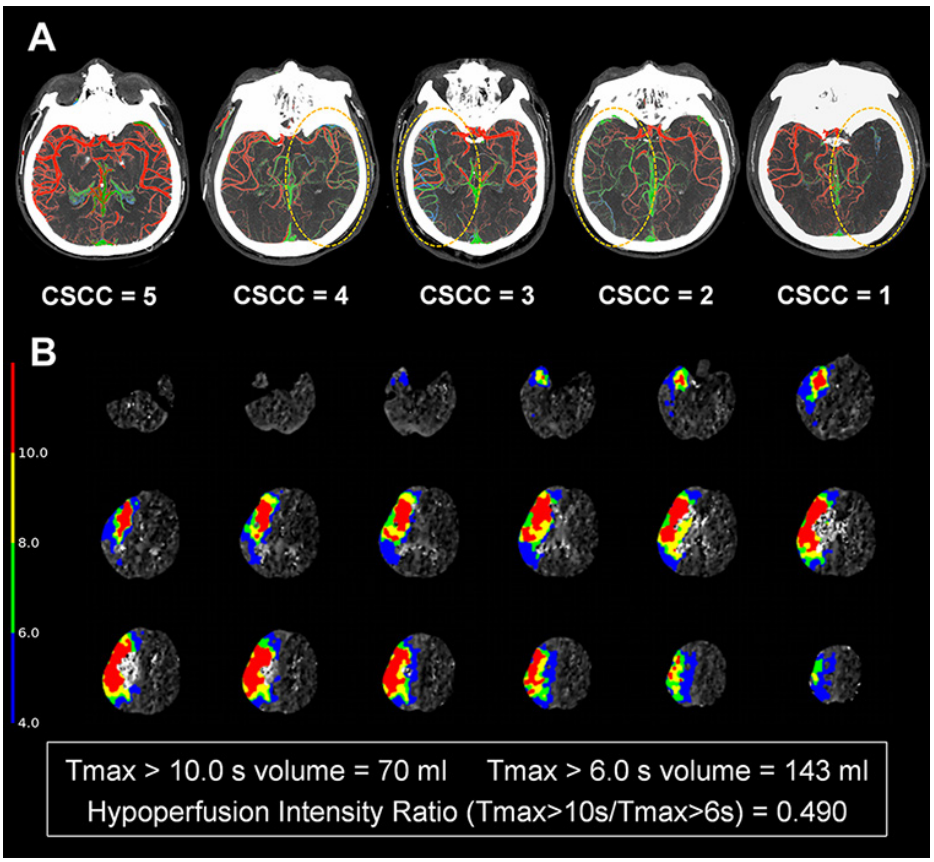

Figure S1 Example of evaluation of CSCC on mCTA and HIR on CTP imaging. CSCC, Collateral Score on Color-Coded summation maps; CTP, computed tomography perfusion; HIR, hypoperfusion intensity ratio; mCTA, multiphase computed tomography angiography; Tmax, time to maximum.

https://dx.doi.org/10.21037/qims-24-1073

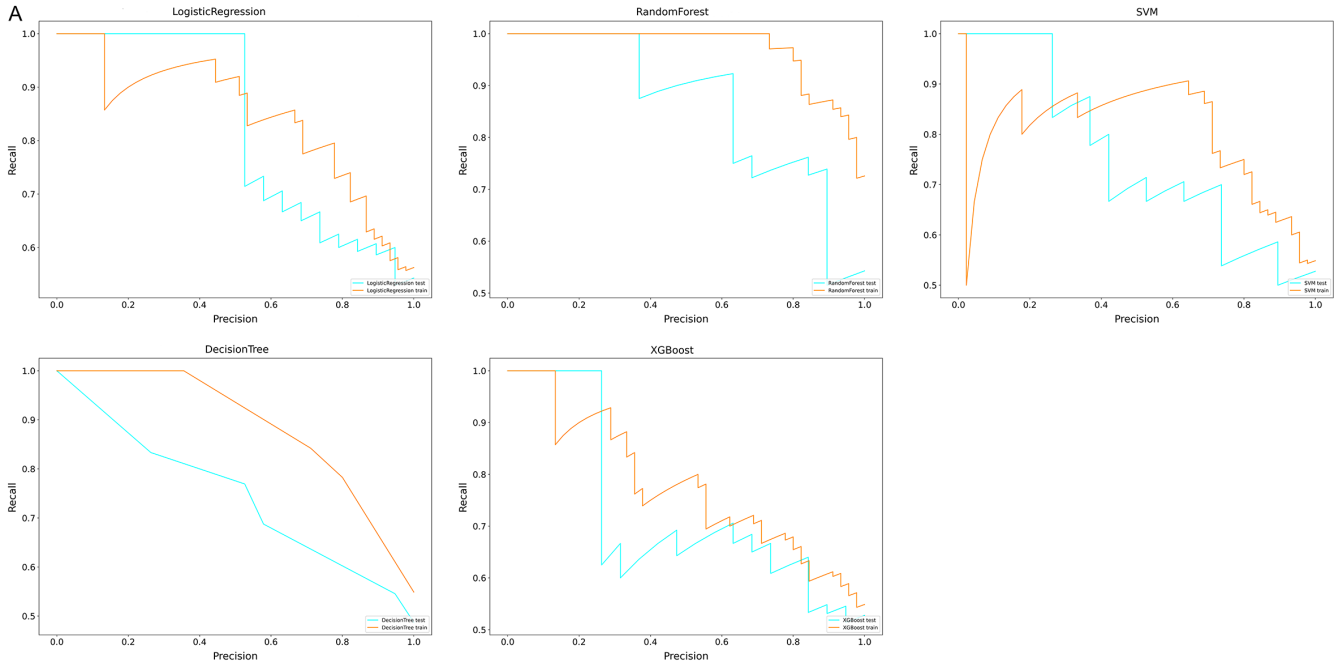

Figure S2 PR curves of training set and testing set in ML models. The Y-axis is precision and the X-axis is recall. The AUCs of LR model, RF model, SVM model, DT model, and XGBoost model were 0.844 (95% CI: 0.707–0.924), 0.968 (95% CI: 0.852–0.994), 0.799 (95% CI: 0.657–0.892), 0.876 (95% CI: 0.744–0.945), 0.776 (95% CI: 0.632–0.875) in training set, and 0.827 (95% CI: 0.593–0.940), 0.858 (95% CI: 0624–0.956), 0.755 (95% CI: 0.520–0.898), 0.714 (95% CI: 0.480–0.871), 0.727 (95% CI: 0.492–0.880) in testing set, respectively. Among them, RF model had the highest AUC in the PR curve in both training set and testing set. AUC, area under the ROC curve; CI, confidence interval; DT, decision tree; LR, logistic regression; ML, machine learning; PR, precision-recall; RF, random forest; ROC, receiver operator characteristic; SVM, support vector machine; XGBoost, eXtreme gradient boosting.

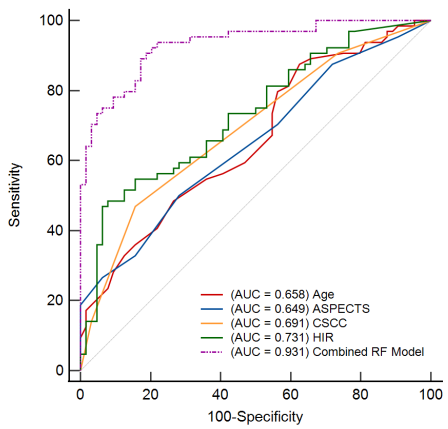

Figure S3 ROC curves for age, ASPECTS, CSCC, HIR, and RF model for prediction of clinical outcomes of EVT patients. ASPECTS, Alberta Stroke Program Early CT Score; AUC, area under the ROC curve; CSCC, Collateral Score on Color-Coded summation maps; HIR, hypoperfusion intensity ratio; RF, random forest; ROC, receiver operator characteristic.

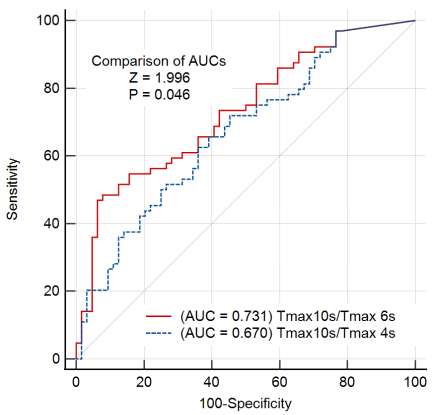

Figure S4 Comparison of predictive performance for clinical outcome between HIRs with different thresholds. AUC, area under the ROC curve; ROC, receiver operator characteristic; Tmax, time to maximum.

https://dx.doi.org/10.21037/qims-24-1073

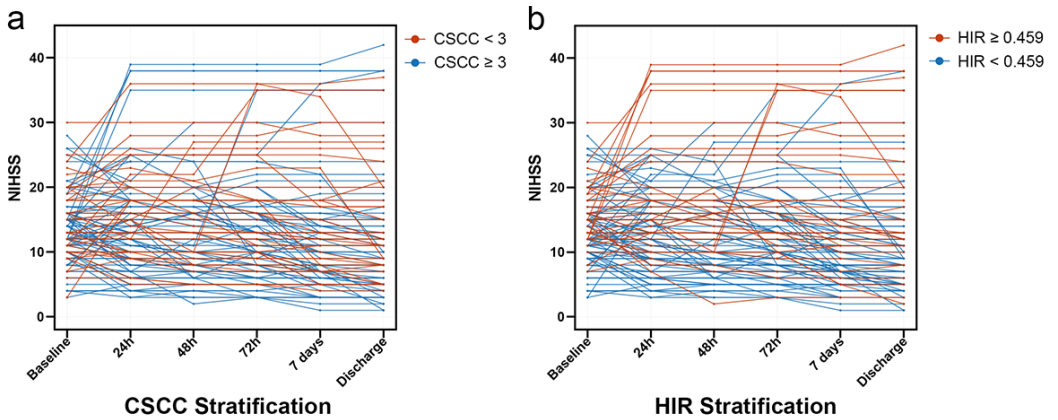

Figure S5 The spaghetti plots traced the trend of postoperative NIHSS during hospitalization in CSCC stratification and HIR stratification. CSCC, Collateral Score on Color-Coded summation maps; HIR, hypoperfusion intensity ratio; NIHSS, National Institutes of Health Stroke Scale.

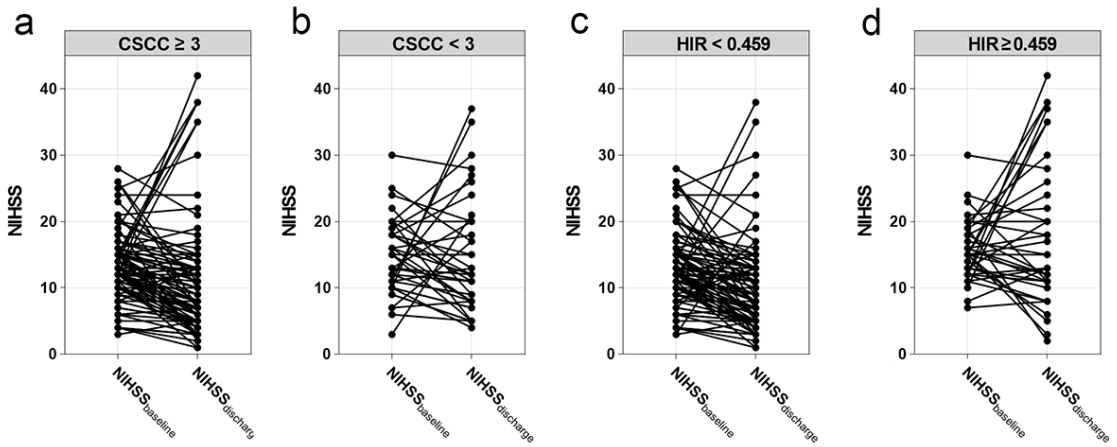

Figure S6 Graphs show changes in NIHSS scores between baseline and discharge for patients in subgroups in CSCC stratification (A,B) and HIR stratification (C,D). CSCC, Collateral Score on Color-Coded summation maps; HIR, hypoperfusion intensity ratio; NIHSS, National Institutes of Health Stroke Scale.

https://dx.doi.org/10.21037/qims-24-1073

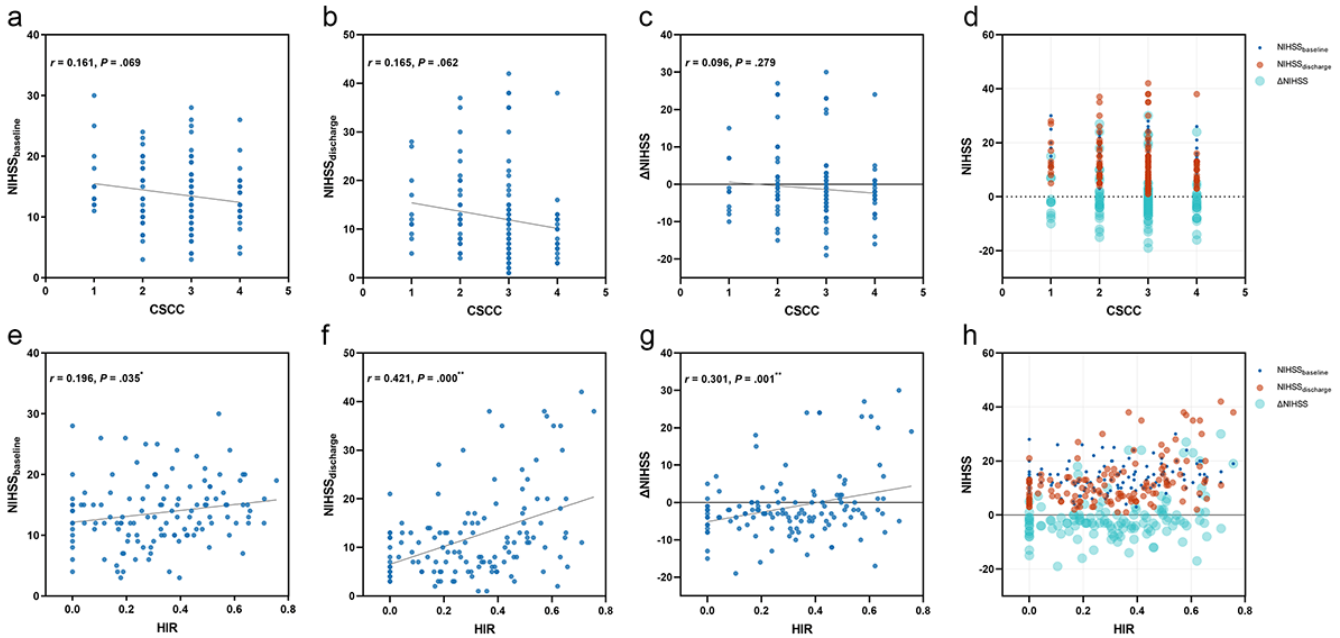

Figure S7 The association between CSCC, HIR, and NIHSS. *, P<0.05; **, P<0.01. CSCC, Collateral Score on Color-Coded summation maps; HIR, hypoperfusion intensity ratio; NIHSS, National Institutes of Health Stroke Scale.

https://dx.doi.org/10.21037/qims-24-1073