Supplementary Table S1 MRI acquisition parameters for the main and verification cohorts

Parameters	Main Queue	Verification Queue
Equipment Manufacturer / Model	United Imaging Healthcare uMR790	GE Healthcare Discovery 750
Magnetic field strength	3.0T	3.0T
Coil Type	16-channel body phased array coil	32-channel body phased array coil
Subject preparation	Fast for ≥ 4 hours prior to scanning and perform breathing exercises	Fast for ≥ 4 hours prior to scanning and perform breathing exercises
Posture and Breath Control	Supine position, head-first, using respiratory gating to minimize motion artifacts	Supine position, head-first, using respiratory gating to minimize motion artifacts
Contrast agent	Gadobutrol (GdBOPTA), 0.1 mmol/kg, injection rate 2.5 mL/s	Gadobutrol (GdBOPTA), 0.1 mmol/kg, injection rate 2.5 mL/s
Incorporation of radiomics analysis into the sequence	Late arterial phase (30 seconds)	Late arterial phase (30 seconds)
CE-T1-weighted sequence type	Axial Gradient Echo Sequence	Axial Phase-Shift Gradient Echo Water-Fat Separation T1-Weighted Image
Repeat Time (TR, ms)	3.0	3.8
Echo Time (TE, ms)	1.3	1.7
Field of view (FOV, cm)	40 × 30	36 × 36
Matrix size	157 × 133	224 × 192
Layer thickness / Layer spacing (mm)	1.5 / 0.0	5.2 / 0.0
T2-weighted sequence type	Fast Spin Echo	Fast Spin Echo
Repeat Time (TR, ms)	8000	6000
Echo Time (TE, ms)	116	120

Field of view (FOV, cm)	38 × 30	34 × 34
Matrix size	256 × 171	320 × 256
Layer thickness / Layer spacing (mm)	5.0 / 2.0	6.0 / 1.0
Quality Control	Two senior radiologists performed blinded image interpretation; cases with severe motion artifacts or missing critical phase images were excluded	Consistent with the main queue

Notes:All examinations were performed on 3.0 T MRI scanners using body phased-array coils with standardized patient preparation (fasting and breathing training).
Image quality was independently evaluated by two senior radiologists; cases with severe motion artifacts or missing critical phases were excluded.
Despite differences in scanner manufacturers and coil channels, consistent imaging planes and contrast phases were maintained between the two cohorts.

Supplementary Table S 2 Baseline clinical characteristics of patients with HTG-AP in the independent validation cohort

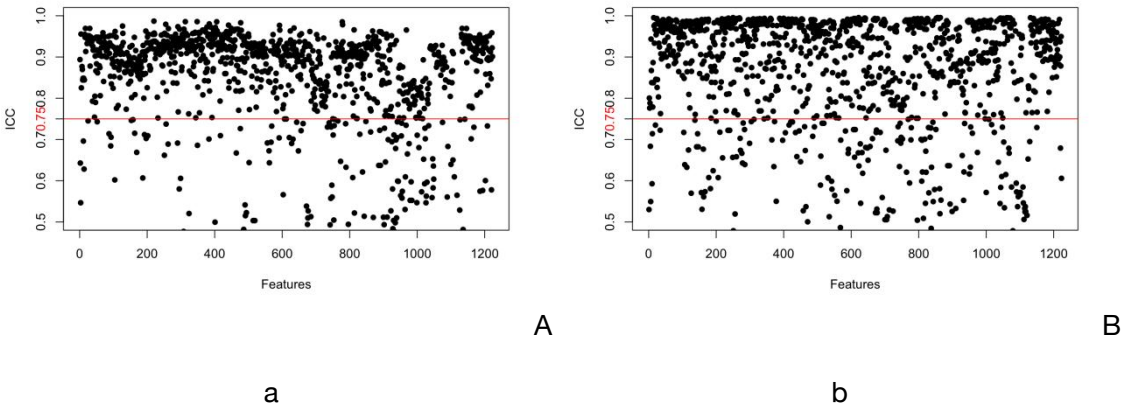
Features	HTG+RAP(n=28)	HTG+nRAP(n=24)	t / χ^2	*P* value
Age (years)	40.6±13.4	45.5±9.3	1.508[b]	0.138
Fatty liver, n (%)	15(53.6)	12(50.0)	0.066[a]	0.797
Hypertension, n (%)	6(25.0)	9(37.5)	1.626[a]	0.202
Gallstones, n (%)	9(32.1)	7(29.2)	0.054[a]	0.817
Lipase, median (IQR) [U·L^{-1}]	214.4(138.0,729.92)	493.5(102.5,2406.2)	-0.955[c]	0.340
Pancreatic amylase [U·L^{-1}]	233.6±118.7	433.0(256.0,789.5)	-3.818[c]	<0.001
Admission blood glucose (mmol·L^{-1})	7.4(6.6,9.8)	7.4±2.0	-1.267[c]	0.205
Triglyceride at readmission (mmol·L^{-1})	5.4(4.3,6.1)	5.4±1.7	-0.698[c]	0.485
Total cholesterol (mmol·L^{-1})	6.1(3.7,9.0)	2.2(0.8,5.9)	-3.030[c]	0.002
Smoking, n (%)	16(57.1)	11(45.8)	0.662[a]	0.416
HDL-C(mmol/L)	2.4±1.3	2.5±0.8	0.093[b]	0.926
WBC count ($\times 10^9$ · L^{-1})	12.5±4.8	11.0±5.5	-1.214[b]	0.231
High-sensitivity C-reactive protein (hs-CRP, mg·L^{-1})	143.4 (85.7, 339.7)	13(5.4,61.0)	-4.626[c]	<0.001
Alcohol consumption, n (%)	12(42.9)	10(41.7)	0.008[a]	0.931
Male sex, n (%)	17(60.7)	14(58.4)	0.030[a]	0.862

Notes:Continuous variables are presented as mean ± standard deviation (SD) or median (interquartile range, IQR), as appropriate.

Categorical variables are expressed as counts and percentages, n (%).

a: χ^2 test; b: independent-samples t test; c: Mann–Whitney U test (Z value).

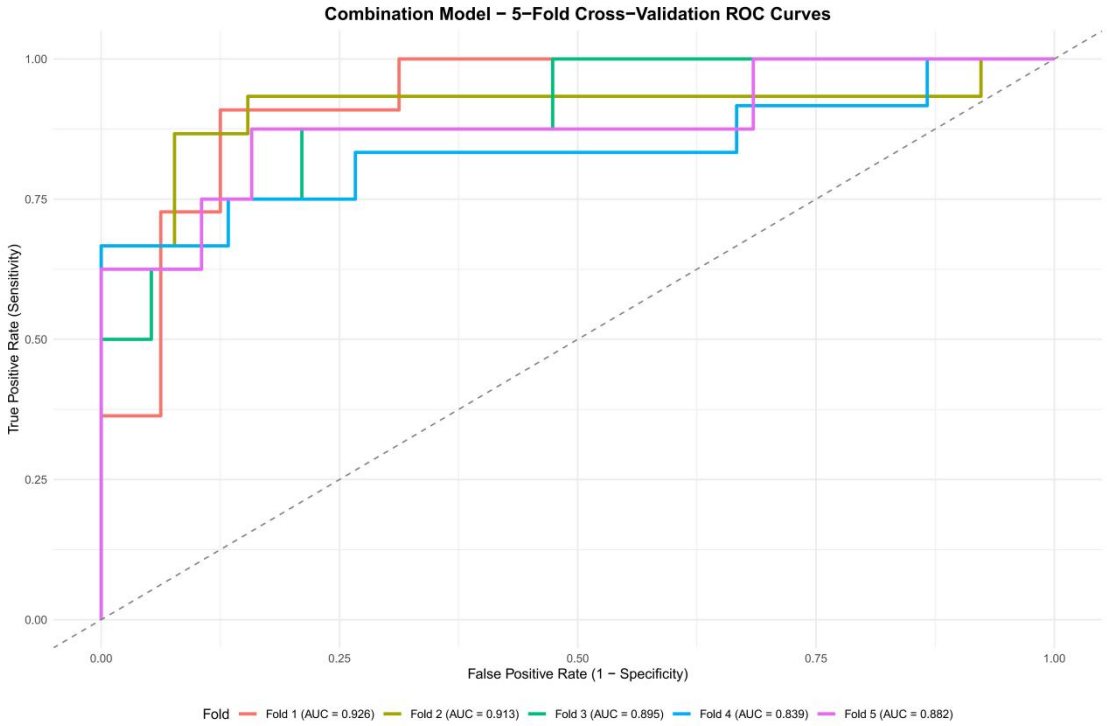
$P < 0.05$ was considered statistically significant.



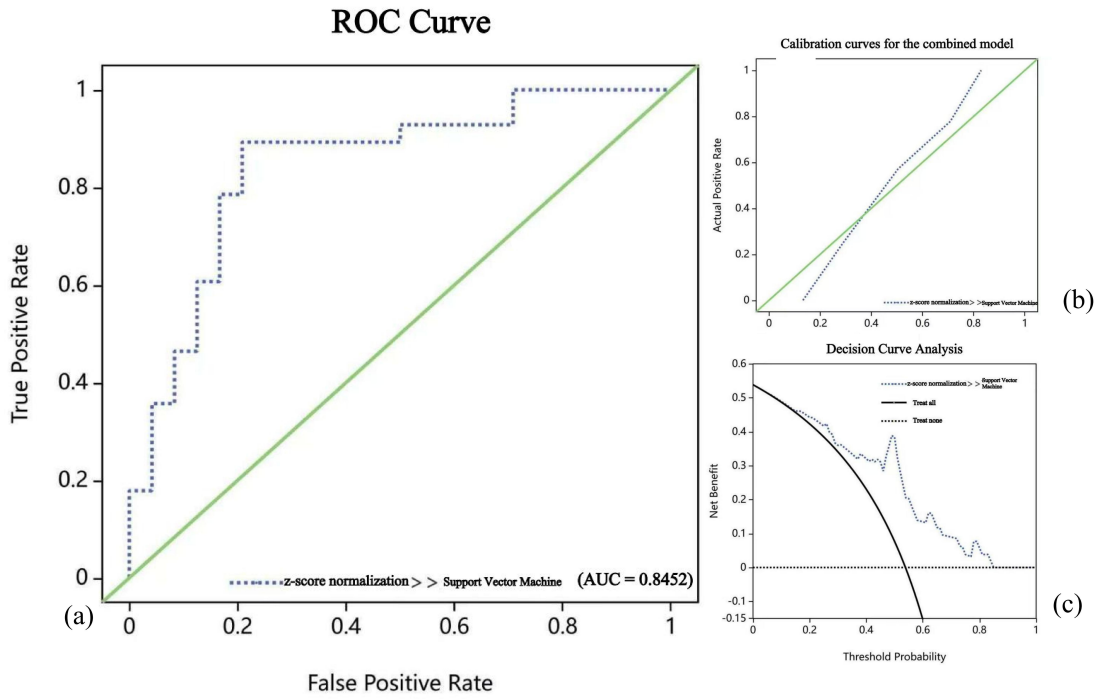
Supplementary Figure 1 Illustration of inter-group consistency analysis

(a) Consistency assessment on the contrast-enhanced arterial phase T1-weighted sequence.

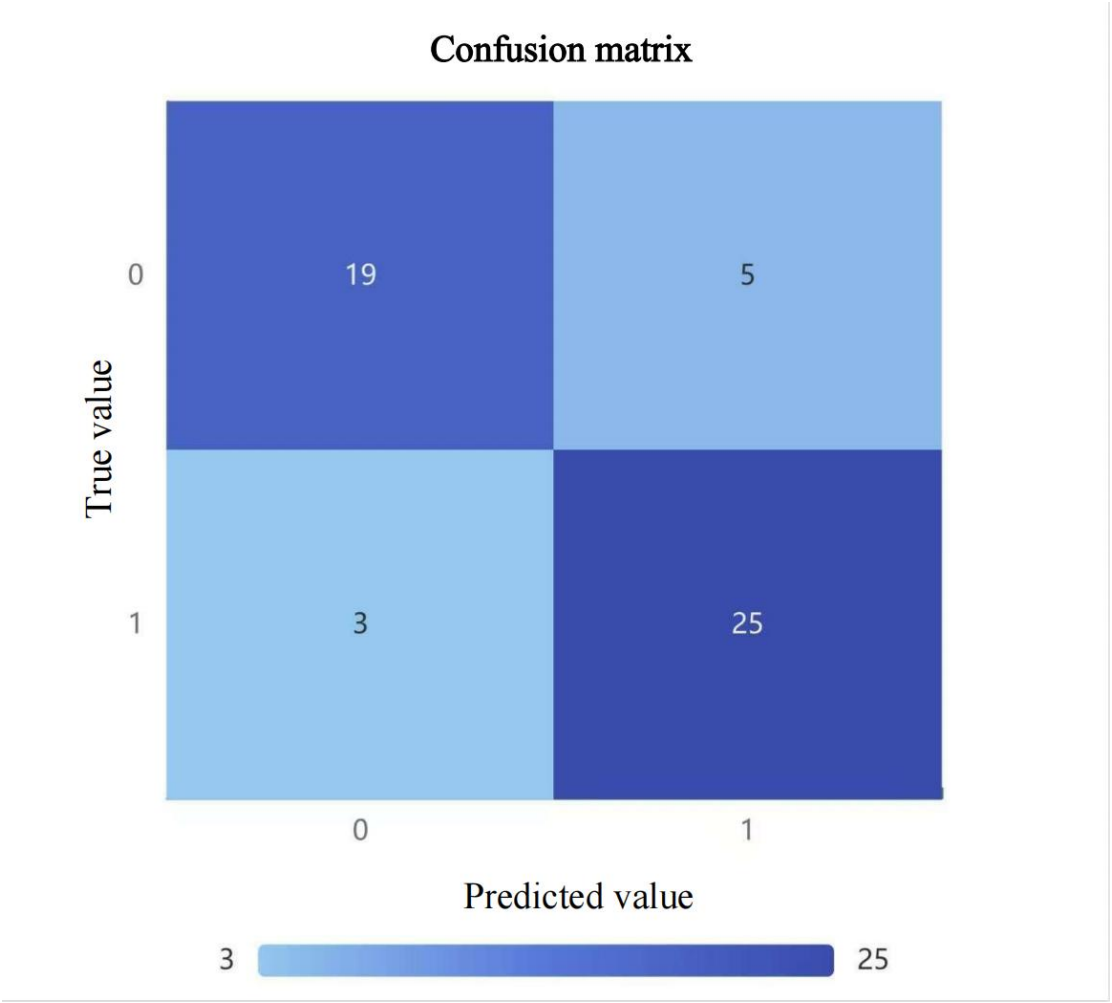
(b) Consistency assessment on the T2-weighted sequence.



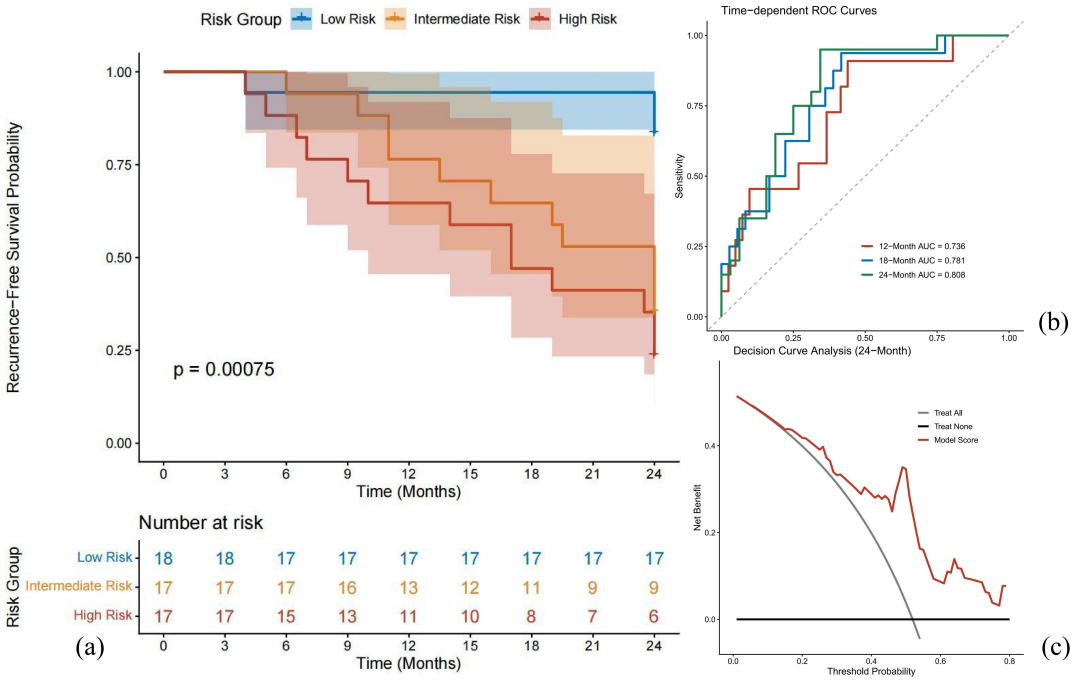
Supplementary Figure 2 Five-fold cross-validation performance of the combination model. Receiver operating characteristic (ROC) curves and fold-wise performance of the combination model in five-fold cross-validation within the training cohort (n=137), showing stable discrimination across folds.



Supplementary Figure 3 (a) Model performance evaluation combining ROC; (b) calibration, and; (c) decision curve analyses



Supplementary Figure 4 Confusion matrix of the fusion model at the optimal Youden cutoff (0.42). The model achieved an accuracy of 84.6%, sensitivity 89.3%, specificity 79.2%, and F1-score 0.862. Five false-positive and three false-negative cases were observed, mostly near borderline predicted probabilities (~0.35–0.55).



Supplementary Figure 5 (a) Kaplan–Meier curves for recurrence-free survival among three risk groups in the external validation cohort. The high-risk group showed significantly worse survival than the intermediate- and low-risk groups ($P = 0.00075$); (b);(c)