Supplementary Table 1 Baseline information for the training population, stratified by malignant and inflammatory hepatic lesions

Variable	Total	Malignant lesions (*n* = 315)	Inflammatory lesions (*n* = 306)	Statistic	P
Sex				1.226	0.268
Male	417 (67.15%)	218 (69.21%)	199 (65.03%)		
Female	204 (32.85%)	97 (30.79%)	107 (34.97%)		
Diabetes history				0.587	0.444
Yes	69 (11.11%)	32 (10.16%)	37 (12.09%)		
No	552 (88.89%)	283 (89.84%)	269 (87.91%)		
Hypertension history				0.354	0.552
Yes	100 (16.10%)	48 (15.24%)	52 (16.99%)		
No	521 (83.90%)	267 (84.76%)	254 (83.01%)		
Echocardiographic features				0.557	0.455
Mixed	358 (57.65%)	177 (56.19%)	181 (59.15%)		
Other	263 (42.35%)	138 (43.81%)	125 (40.85%)		
Lesion location				0.339	0.561
Left lobe	497 (80.03%)	255 (80.95%)	242 (79.08%)		
Right lobe	124 (19.97%)	60 (19.05%)	64 (20.92%)		
Border clarity				0.728	0.394
Ill-defined	456 (73.43%)	236 (74.92%)	220 (71.90%)		
Well-defined	165 (26.57%)	79 (25.08%)	86 (28.10%)		
Morphology				16.391	<0.001
Irregular	361 (58.13%)	208 (66.03%)	153 (50.00%)		

Regular	260 (41.87%)	107 (33.97%)	153 (50.00%)		
Homogeneous enhancement				90.149	<0.001
Yes	390 (62.80%)	255 (80.95%)	135 (44.12%)		
No	231 (37.20%)	60 (19.05%)	171 (55.88%)		
Cirrhosis				118.903	<0.001
Yes	226 (36.39%)	180 (57.14%)	46 (15.03%)		
No	395 (63.61%)	135 (42.86%)	260 (84.97%)		
Hepatitis				73.388	<0.001
Yes	477 (76.81%)	287 (91.11%)	190 (62.09%)		
No	144 (23.19%)	28 (8.89%)	116 (37.91%)		
Blood flow signals				186.196	<0.001
0	221 (35.59%)	34 (10.79%)	187 (61.11%)		
1	153 (24.64%)	89 (28.25%)	64 (20.92%)		
2	157 (25.28%)	120 (38.10%)	37 (12.09%)		
3	90 (14.49%)	72 (22.86%)	18 (5.88%)		
Calcification				67.871	<0.001
Yes	303 (48.79%)	205 (65.08%)	98 (32.03%)		
No	318 (51.21%)	110 (34.92%)	208 (67.97%)		
Age	57.41±8.26	59.34±9.16	55.41±6.68	4.718	<0.001
Body mass index	22.96±2.01	22.97±1.98	22.96±2.04	0.032	0.974
Lesion size	3.80 [2.10, 5.90]	4.95 [2.50, 7.38]	3.20 [1.90, 4.70]	5.105	<0.001
Wash-in time	22.99±5.25	21.24±5.27	24.81±4.59	6.963	<0.001
Wash-out time	137.00 [72.00, 210.00]	117.00 [64.00, 162.75]	183.00 [95.50, 252.00]	5.512	<0.001
AFP	9.70 [6.20,	17.07 [10.87,	6.50 [4.38,	13.616	<0.001

	17.27]	23.52]	8.82]		
MELD score	18.85 [15.98, 22.33]	21.88 [18.77, 25.53]	16.34 [13.64, 18.98]	11.503	<0.001

Descriptive statistics of continuous variables are expressed using the mean ± SD (normal distribution) or median (quartile; skewed distribution); categorical variables are shown as *n* (%). AFP: Alpha-fetoprotein; MELD: Model for End-Stage Liver Disease.

Supplementary Table 2 **DeLong test for area under the curve comparison between models (test set)**

Comparison	**AUC-1**	**AUC-2**	**Difference**	**Z-statistic**	**P-value**	**Significant**
Logistic *vs* DT	0.965	0.886	0.079	2.926	0.003	Yes
Logistic *vs* RF	0.965	0.967	-0.002	-0.325	0.745	No
Logistic *vs* XGBoost	0.965	0.978	-0.013	-1.063	0.288	No
Logistic *vs* SVM	0.965	0.964	0.001	0.255	0.799	No
RF *vs* XGBoost	0.967	0.978	-0.010	-0.943	0.346	No
RF *vs* SVM	0.967	0.964	0.004	0.503	0.615	No

Significant = $P < 0.05$. DeLong test was used to compare paired ROC curves. AUC: Area under the curve; SVM: Support vector machine; RF: Random forest; DT: Decision tree.

Supplementary Table 3 Mcnemar test for classification agreement between models (test set)

Comparison	**M1 Corr.**	**M2 Corr.**	**Both Corr.**	**Both Wrong**	**M1 Only**	**M2 Only**	χ^2	***P* value**	**Sig.**
Logistic *vs* DT	116	104	99	3	17	5	5.500	0.019	Yes
Logistic *vs* RF	116	115	113	6	3	2	0.000	1.000	No
Logistic *vs*	116	113	110	5	6	3	0.444	0.505	No

XGBoost								
Logistic *vs* SVM	116	116	115	7	1	1	0.000	1.000	No
RF *vs* XGBoost	115	113	109	5	6	4	0.100	0.752	No
RF *vs* SVM	115	116	114	7	1	2	0.000	1.000	No

M1 = Model 1; M2 = Model 2; Corr. = Correct; Sig. = Significant ($P < 0.05$). SVM: Support vector machine; RF: Random forest; DT: Decision tree.

Supplementary Table 4 Feature importance comparison across models

Rank	Logistic Feature	Imp.	RF Feature	Imp.	XGBoost Feature	Imp.
1	Blood_Flow_Signal	1.000	Blood_Flow_Signal	1.000	Regression_Time	1.000
2	Hepatitis	0.695	Lesion_Size	0.803	Lesion_Size	0.957
3	Liver_Cirrhosis	0.633	Regression_Time	0.748	Liver_Cirrhosis	0.717
4	Morphology	0.458	Liver_Cirrhosis	0.702	Start_Increase_Time	0.464
5	Calcification_Status	0.405	Start_Increase_Time	0.523	Morphology	0.451
6	Boundary	0.252	Morphology	0.478	Blood_Flow_Signal	0.393
7	Lesion_Size	0.073	Hepatitis	0.387	Age	0.290
8	Start_Increase_Time	0.057	Calcification_Status	0.264	Hepatitis	0.250
9	Age	0.016	Age	0.250	Calcification_Status	0.235
10	Regression_Time	0.000	Boundary	0.000	Boundary	0.000

Imp. = Importance (min-max scaled to 0-1 within each model).

Supplementary Table 5 Decision curve analysis - net benefit summary (test set)

Model	Max NB	Threshold at Max	Mean NB (0-50%)	Mean NB (50-100%)	Effective Range
Logistic	0.476	0%	0.416	0.350	0%-99%
DT	0.476	0%	0.379	0.176	0%-99%

RF	0.476	0%	0.409	0.291	0%-99%
XGBoost	0.476	0%	0.428	0.329	0%-99%
SVM	0.476	0%	0.425	0.336	0%-99%

Effective range indicates threshold range where model provides positive net benefit.NB: Net benefit; SVM: Support vector machine; RF: Random forest; DT: Decision tree.