

MRI acquisition parameters

Detailed MRI scanning parameters are shown in Supplementary Table S1.

Feature Selection Process and Final Feature List

Table S2-3 details the multi-step selection process (including the change in the number of features at each step) for Clinical Features, Intratumoral Heterogeneity (ITH) Features, and Radiomics Features, and lists all finally integrated features to support the feature sources for model construction in the main text. Among them:

Definition of selection steps: ① Raw (original features); ② AfterVarFilter (after variance filtering, removing low-variance non-informative features); ③ AfterCorrFilter (after correlation filtering, removing highly redundant features); ④ AfterRFECV (after Recursive Feature Elimination with Cross-Validation, selecting key features for the model); ⑤ AfterCapMaxFeatures (after capping the maximum number of features, balancing model complexity and performance);

Explanation of feature naming: GLCM (Gray-Level Co-occurrence Matrix), GLRLM (Gray-Level Run-Length Matrix), GLDM (Gray-Level Dependence Matrix), and GLSZM (Gray-Level Size Zone Matrix) are all commonly used quantitative extraction methods for imaging/ITH features;

MetaProb (meta-probability features): Fusion probabilities output after building models based on different combinations of data sources, used to improve the generalization ability of the model

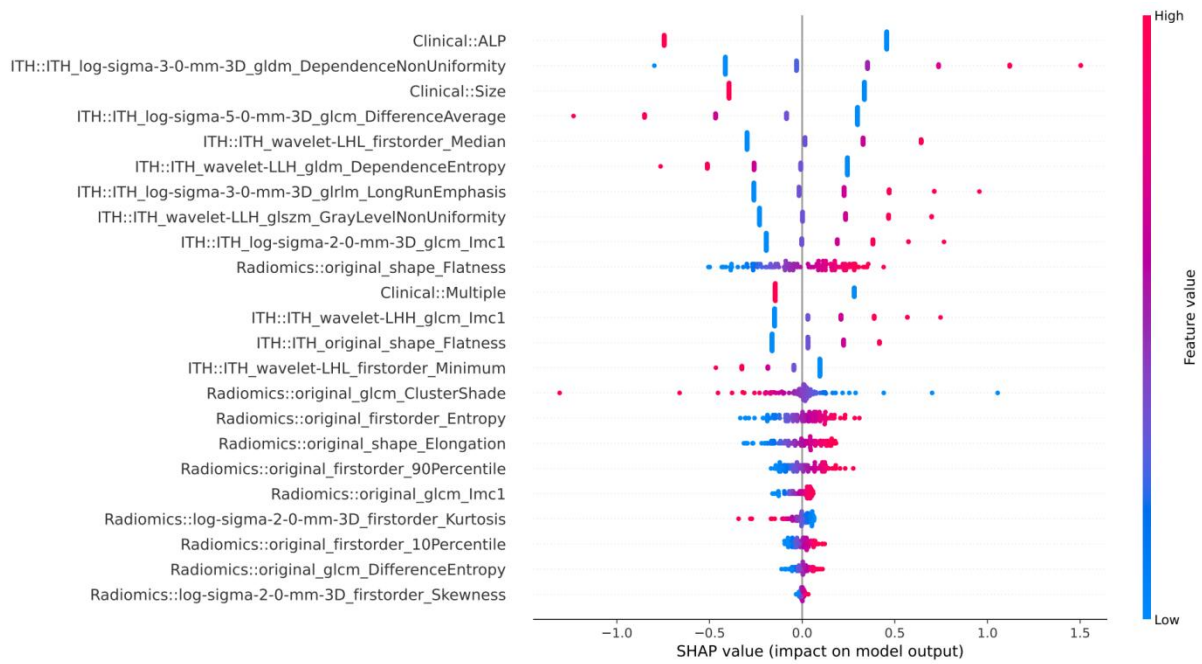
ICC calculation process

Following the ICC selection process diagram from Koo et al. [1], we used the ICC formula, defined as two-way random effects, absolute agreement, and single rater/measurement. The ICC values for each feature were calculated as follows:

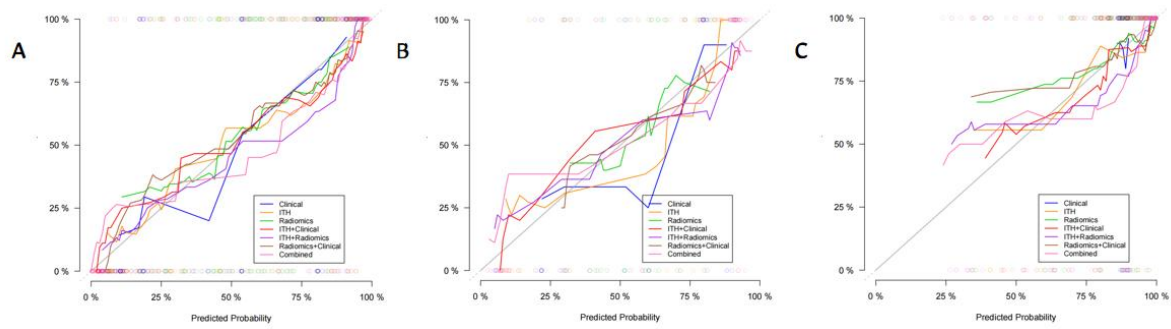
$$ICC = \frac{MS_R - MS_E}{MS_R + (k-1)MS_E + \frac{k}{n}(MS_C - MS_E)}$$

References:

[1]. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*15(2) (2016) 155-63.



Supplementary Figure 1 SHAP summary plot illustrating the contribution of clinical, intratumoral heterogeneity (ITH) habitat imaging, and radiomics features to the combined model. The x-axis (SHAP value) reflects each feature’s impact direction/magnitude on model output. The y-axis lists included features, and the blue-to-red gradient denotes feature values (low to high). The SHAP summary plot demonstrates that ITH-related features contribute substantially to the prediction, ranking among the top predictors.



Supplementary Figure 2 Calibration curves of different models in the training cohort (A), internal validation cohort (B), and external validation cohort (C). The calibration curve of the Combined model closely approximates the ideal 45° diagonal line in all three cohorts, indicating a high degree of agreement between the predicted probability of TACE non-response and the actual probability. This demonstrates that the Combined model exhibits excellent calibration performance across internal and external cohorts, thereby supporting its reliability in clinical practice.

Supplementary Table 1 Parameter Settings of MRI Sequences

MRI Sequence Name	Slice Thickness	Slice Gap	FOV (Field of View)	Number of Slices	TR (Repetition Time)	TE (Echo Time)	Flip Angle	Matrix	NEX (Number of Excitations)	Echo Length	Train	Acquisition Time
LAVA-flex Sequence	5 mm	-	38 cm	46	5.5 ms	TE1: minimum; TE2: 2.2 ms	10°	320×224	1	-		16 s
Fat-Suppressed T2WI	6 mm	0.5 mm	38 cm	36	-	85 ms	-	320×320	1.5	30		-
DWI-EPI	6 mm	0.5 mm	38 cm	36	-	minimum	-	128×130	1	-		-
LAVA	5 mm	-	38 cm	46	3.7 ms	TE1: 1.5 ms	12°	320×256	1	-		13 s per phase

Notes: 1. "-" indicates that the corresponding parameter is not mentioned for the sequence; 2. TR =Repetition Time, TE =Echo Time, FOV =Field of View, NEX =Number of Excitations , DWI-EPI=Diffusion-Weighted Imaging with Echo Planar Imaging , LAVA=Liver Acquisition with Volume Acceleration

3. All radiomic features in this study were exclusively extracted from arterial-phase images of the dynamic LAVA sequence. The LAVA-flex sequence listed herein served as an auxiliary imaging module and was not involved in feature extraction.

			LLH_gldm_DependenceEntropy, ITH_wavelet-LHL_firstorder_Median, ITH_wavelet-LHL_firstorder_Minimum, ITH_wavelet-LHH_glcm_Imc1
	AfterCapMaxFeatures	10	-
Radiomics Features	Raw	1132	-
	AfterVarFilter	817	-
	AfterCorrFilter	147	-
	AfterRFECV	50	original_shape_Elongation, original_shape_Flatness, original_firstorder_10Percentile, original_firstorder_90Percentile, original_firstorder_Entropy, original_glcm_ClusterShade, original_glcm_DifferenceEntropy, original_glcm_Imc1, log-sigma-2-0- mm-3D_firstorder_Kurtosis, log-sigma-2-0-mm-3D_firstorder_Skewness
	AfterCapMaxFeatures	10	-

Notes: In the stage of feature selection and model development, this study designed a multi-level dimensionality reduction process to ensure the stability, independence, and predictive relevance of features in the training dataset.

Supplementary Table 3 Final Combined Feature List

Feature Category	Feature Name
Clinical Features	Clinical::Size
	Clinical::ALP
	Clinical::Multiple
ITH Features	ITH::ITH_original_shape_Flatness
	ITH::ITH_log-sigma-2-0-mm-3D_glcm_Imc1
	ITH::ITH_log-sigma-3-0-mm-3D_glrIm_LongRunEmphasis
	ITH::ITH_log-sigma-3-0-mm-3D_gldm_DependenceNonUniformity
	ITH::ITH_log-sigma-5-0-mm-3D_glcm_DifferenceAverage
	ITH::ITH_wavelet-LLH_glszm_GrayLevelNonUniformity
	ITH::ITH_wavelet-LLH_gldm_DependenceEntropy
	ITH::ITH_wavelet-LHL_firstorder_Median
	ITH::ITH_wavelet-LHL_firstorder_Minimum
	ITH::ITH_wavelet-LHH_glcm_Imc1
Radiomics Features	Radiomics::original_shape_Elongation
	Radiomics::original_shape_Flatness
	Radiomics::original_firstorder_10Percentile
	Radiomics::original_firstorder_90Percentile
	Radiomics::original_firstorder_Entropy

Radiomics::original_glcM_ClusterShade

Radiomics::original_glcM_DifferenceEntropy

Radiomics::original_glcM_Imc1

Radiomics::log-sigma-2-0-mm-3D_firstorder_Kurtosis

Radiomics::log-sigma-2-0-mm-3D_firstorder_Skewness

Supplementary Table 4 DeLong Test Results

Dataset	Model Compared with the Combined Model	AUC (Combined Model)	AUC (Compared Model)	p value	- Significantly Different?
Training Set	Clinical	0.966	0.864	0.020	Yes
Training Set	ITH	0.966	0.893	0.046	Yes
Training Set	Radiomics	0.966	0.711	< 0.001	Yes
Training Set	ITH + Clinical	0.966	0.949	0.579	No
Training Set	ITH + Radiomics	0.966	0.929	0.281	No
Training Set	Radiomics + Clinical	0.966	0.896	0.036	Yes
Validation Set	Clinical	0.907	0.778	0.288	No
Validation Set	ITH	0.907	0.898	0.924	No
Validation Set	Radiomics	0.907	0.657	< 0.001	Yes
Validation Set	ITH + Clinical	0.907	0.880	0.784	No
Validation Set	ITH + Radiomics	0.907	0.907	1.000	No
Validation Set	Radiomics + Clinical	0.907	0.676	< 0.001	Yes
External Validation Set	Clinical	0.933	0.541	< 0.001	Yes
External Validation Set	ITH	0.933	0.873	0.030	Yes

External Validation Set	Radiomics	0.933	0.780	< 0.001	Yes
External Validation Set	ITH + Clinical	0.933	0.888	0.041	Yes
External Validation Set	ITH + Radiomics	0.933	0.938	0.928	No
External Validation Set	Radiomics + Clinical	0.933	0.782	< 0.001	Yes
