Article

Discovery Medicine 2024; 36(183): 765–777
https://doi.org/10.24976/Discov.Med.202436183.72

Local Tumor Progression Predictive Model Based on MRI for Colorectal Cancer Liver Metastases after Radiofrequency Ablation

Guangyuan Zhang[1,2,3,†], Pei Liu[4,†], Guangzhi Wang[5], Xinhong He[5], Lichao Xu[5], Yaohui Wang[5], Wentao Li[5,*], Weijun Peng[1,2,3,*]

[1]Department of Radiology, Shanghai Proton and Heavy Ion Center, Fudan University Cancer Hospital, 201321 Shanghai, China
[2]Department of Radiology, Shanghai Key Laboratory of Radiation Oncology, 201321 Shanghai, China
[3]Department of Radiology, Shanghai Engineering Research Center of Proton and Heavy Ion Radiation Therapy, 201321 Shanghai, China
[4]Department of Radiology, Shanghai First Rehabilitation Hospital, 200090 Shanghai, China
[5]Department of Interventional Radiology, Fudan University Cancer Center, 200032 Shanghai, China
*Correspondence: wentaoli@fudan.edu.cn (Wentao Li); pwj_06@126.com (Weijun Peng)
[†]These authors contributed equally.
Published: 20 April 2024

Purpose: To investigate the post-radiofrequency ablation (RFA) magnetic resonance imaging (MRI) characteristics in patients with liver metastases from colorectal cancer and to build a predictive model for local tumor progression based on these imaging markers.

Materials and Methods: A cohort of 73 patients with 110 colorectal cancer liver metastases (CRCLM) who underwent RFA and MRI one month post-ablation was included in image signs analysis and predictive model training. Using a newly developed MRI appearance scoring criteria, MR Image Appearance Scoring at One Month after RFA (MRIAS 1MO), the semi-quantitative analysis of MRI findings within the ablation zone were conducted independently by two radiologists. The intraclass correlation coefficient (ICC) was calculated to evaluate measurement reliability. Differences in MRIAS 1MO scores were compared using Mann-Whitney U test, focusing on local tumor response outcomes. Using local tumor progression (LTP) as the primary end point, MRIAS 1MO scores and other lesion morphological and clinical characteristics were included to establish predictive model. Predication accuracy was subsequently evaluated using calibration curve, time-dependent concordance index (C index) curve, and LTP-free survival (LTPFS) curve. Another cohort comprising 60 patients with 76 CRCLMs provided additional MRIAS 1MO scores and clinical data associated with LTP. We evaluated the performance of the established predictive model using calibration curve, time-dependent C index curve, and LTPFS curve.

Results: The MRIAS 1MO criteria exhibited strong measurement reliability. The ICC values of T1S (scores from T1WI), T2S (scores form T2WI) and NCES (scores by adding T1S to T2S) MRIS (the overall scores) were 0.825, 0.779, 0.826 and 0.873, respectively. Lesions with LTP showed significantly higher median values for the overall MRIAS 1MO score (MRIS) compared to lesions without LTP (16 vs. 12, $p < 0.001$). MRIS and lesion diameter were independent prognostic factors of LTP and were included in predictive model (hazard ratio: MRIS over 13.5:4.275, lesion diameter larger than 30 mm: 2.056). The predictive model demonstrated an overall C index of 0.721 and risk stratification using the predictive model resulted in significantly different LPTFS times. In the validation cohort, the C index were 0.825, 0.794 and 0.764 at six, twelve and twenty-four months, respectively. Patients classified as high-risk in the validation cohort had a median LTPFS time of 10.0 months, while the median LTPFS time was not reached in the low-risk group.

Conclusions: The semi-quantitative MRIAS 1MO criteria, used for post-RFA MRI appearance analysis, exhibited strong measurement reliability. Prediction models established based on overall MRIAS 1MO score (MRIS) and lesion diameter had good predictive performance for LTP in patients undergoing RFA for CRCLM treatment.

Keywords: colorectal cancer liver metastases; magnetic resonance imaging; local tumor progression; radiofrequency ablation

Introduction

While the overall incidence of colorectal cancer has decreased, new cases now present in younger individuals and at more advanced stages [1]. Nearly half of the colorectal cancer cases progress to liver metastases during the illness [2]. Early observational studies revealed dismal prognosis in patients with colorectal cancer liver metastases (CRCLM), marked by a 6-month median survival time and treatment only by supportive care [3]. Although surgical resection is the sole curative care for CRCLM, only a quarter of patients identify this illness within the treatment period. For patients with unresectable CRCLM, the primary treatment involves systemic approaches such as chemother-

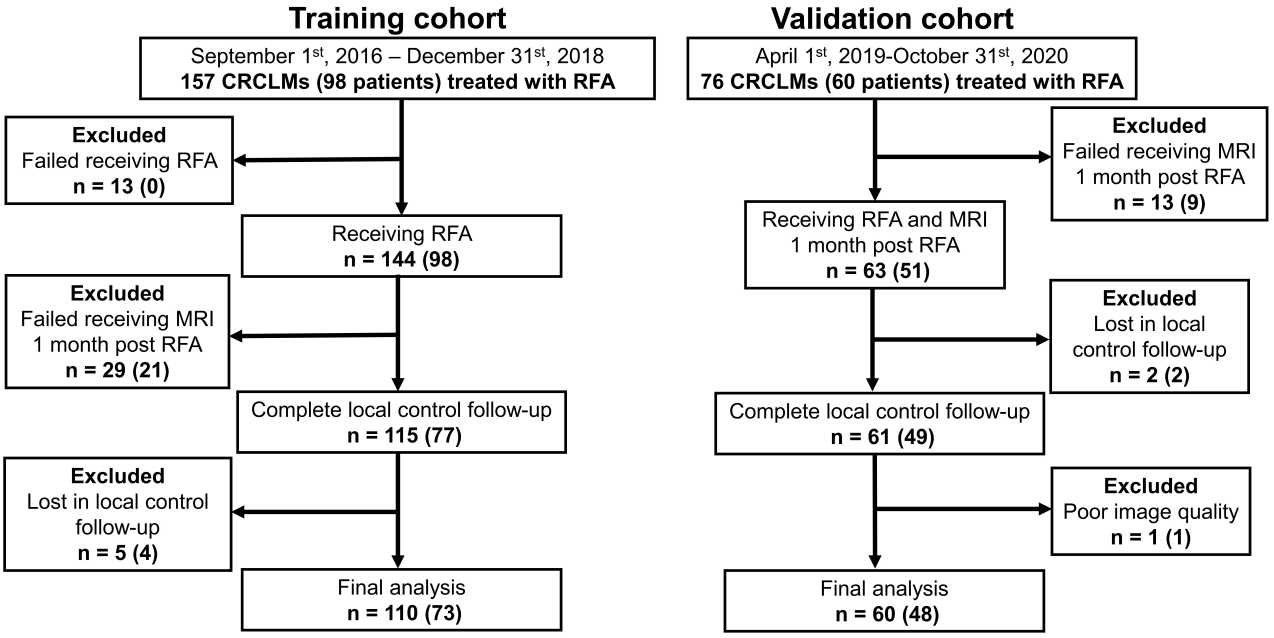

Fig. 1. The inclusion flowchart of training and validation cohorts. RFA, radiofrequency ablation; MRI, magnetic resonance imaging; CRCLM, colorectal cancer liver metastases.

apy and targeted therapy, which could extend the median survival time to 30 months [4–6]. Additionally, locoregional ablative approaches could induce long-term tumor control for CRCLM [7]. Radiofrequency ablation (RFA) has been reported as a safe and effective option for liver tumor treatment for over two decades. A recent randomized trial showed that combining RFA and chemotherapy could prolong median survival time to 45.6 months compared to chemotherapy alone [8].

Despite its efficacy, local tumor progression (LTP) remains a significant concern, limiting the widespread acceptance of thermal ablation as a treatment for CRCLM. The LTP rates were reported to range from 4.3% to 48% [9]. Furthermore, patients with early LTP (LTP at six months after ablation) exhibited shorter median survival time (15.5 months vs. 32.5 months) compared to those without [10]. Additionally, patients who received re-ablation or other treatment modalities after LTP showed extended survival time (45.5 months vs. 31.0 months) [11]. Therefore, there is potential for improving the clinical efficacy of RFA by early identification, close monitoring, and timely intervention in high-risk patients for LTP.

Presently, there are no generally accepted methods for local response evaluation and prediction. Existing assessment criteria, like modified computed tomography (CT) criteria, modified Response Evaluation Criteria in Solid Tumor (mRECIST), and the European Association for the Study of Liver (EASL) guidelines, have been reported to offer limited value in stratifying survival risks [12,13]. Although few studies have developed LTP prediction models following thermal ablation, their efficacy in clinical practice requires further external validation.

With its superior soft tissue resolution, magnetic resonance imaging (MRI) has been accepted as a significant modality for CRCLM diagnosis. MRI demonstrates better identification of ablative margins and higher sensitivity in diagnosing incomplete ablations compared to computed tomography (CT) [14]. Typical MRI imaging findings post-RFA include hyper- and hypo-intensity on T1 weighted images (T1WI) and T2 weighted image (T2WI) with an enhanced thin ring surrounding the ablation zone. However, several cases showed varied image appearances. For instance, a discontinuous ring peripheral to the ablation zone was found to be related to intrahepatic distant recurrence in patients with hepatocellular carcinoma (HCC) treated by RFA [15]. The clinical significance of other atypical image signs has been rarely investigated.

Therefore, this study aimed to develop a semi-quantitative scoring criteria of MR image findings post-RFA with the hypothesis that atypical post-ablation image signs were related to LTP. Using combined scores from this criterion with other clinicopathological factors, we tried to build a prediction model of LTP and further verify its potential in a validation cohort.

Materials and Methods

Participants

The study cohort consisted of a training group and a validation one. From September 1st, 2016, to March 1st, 2019, 98 consecutive patients with 157 colorectal cancer liver metastases intended to receive CT-guided RFA. A total of 25 patients with 47 liver metastases were excluded

Table 1. MR imaging parameters.

Sequence name	TR (ms)	TE (ms)	FOV	Matrix	Thickness (mm)	Breath control	Slice number
T1WI in/out phase	120	2.74/1.4	380	512	6	breath hold	32
T2WI	5300	80	380	512	6.5	free	30
Golden VIBE	3.77	1.72	440	512	3	free	52

TR, time repetition; TE, time echo; FOV, field of view; T1WI, T1 weighted images; T2WI, T2 weighted images; VIBE, Volumetric Interpolated Breath-hold Examination.

Table 2. Image Findings and Corresponding Scores of MRIAS 1MO.

Sequence name	Inner area			Marginal area		
	Factors	Image Findings	Scores	Factors	Image Findings	Scores
T1WI	SI	Hyper	1	Ring Completeness	Continuous	1
		Iso	2		Discontinuous	2
		Hypo	3		No Ring	4
Distribution of SI		Homogeneous	1	Ring Thickness	Thin	1
		Heterogeneous	2		Thick	2
T2WI	SI	Hyper	3	Ring Completeness	Continuous	1
		Iso	2		Discontinuous	2
		Hypo	1		No Ring	4
Distribution of SI		Homogeneous	1	Ring Thickness	Thin	1
		Heterogeneous	2		Thick	2
T1CEVP				Ring Thickness	Thin	1
					Thick	2
					Irregular	3

MRIAS 1MO, MR Image Appearance Scoring at One Month after RFA; SI, signal intensity; T1CEVP, T1WI with contrast enchanted of venous phase.

for not receiving RFA or not completing scheduled imaging follow-ups. Resultantly, 73 patients with 110 liver metastases were included in the training set.

From April 1st, 2019, to October 31st, 2020, 60 patients with 76 colorectal cancer liver metastases were initially included in validation group. Due to the lack of complete follow up images and poor lesion visibility, 12 patients with 16 lesions were excluded. Finally, a total of 48 patients with 60 lesions were included in the validation group.

The inclusion procedure is shown in Fig. 1.

The diagnosis of colorectal cancer liver metastases was made by MRI or PET/CT image signs. For those lesions with indeterminate diagnosis, percutaneous biopsy should be considered within the multidisciplinary team discussion.

This study was approved by the institutional review board. Signed informed consent was obtained from all patients. The study complied with Declaration of Helsinki.

RFA Procedure

All included patients received percutaneous RFA with CT guidance. MedSphere (MedSphere International, Inc., Shanghai, China) electrodes were used. The ablations were performed to achieve an ablative margin of at least 5 mm. An immediate post-procedure CT scan was performed to evaluate the ablative margin and complications for each patient. The minimal ablative margin was measured with the approach introduced by previous literature [16]. Namely, distances from tumor and ablative zone to intrahepatic markers like cyst or large vessels were measured. The minimal difference of these two distances was considered as minimal ablative margin. Up to three interventional radiologists with at least 5 years of experience in CT-guided percutaneous RFA performed the ablation procedures.

Follow-up and Local Tumor Progression (LTP) Evaluation

All included patients received MRI four weeks post-RFA to evaluate clinical efficacy. Subsequently, patients underwent contrast-enhanced CT or MRI follow-ups at three-month intervals in the first year and then at six-month intervals in subsequent years post-RFA to assess local response. LTP was determined based on the appearance of tumor foci within 1 cm of the ablation zone's edge on contrast-enhanced CT or MRI, assuming adequate ablation confirmation of ablation efficacy within one month of image follow-up [17]. LTP-free survival time was the endpoint in current study, and was defined as the duration between the ablation procedure date and the LTP diagnosis date via follow-up imaging.

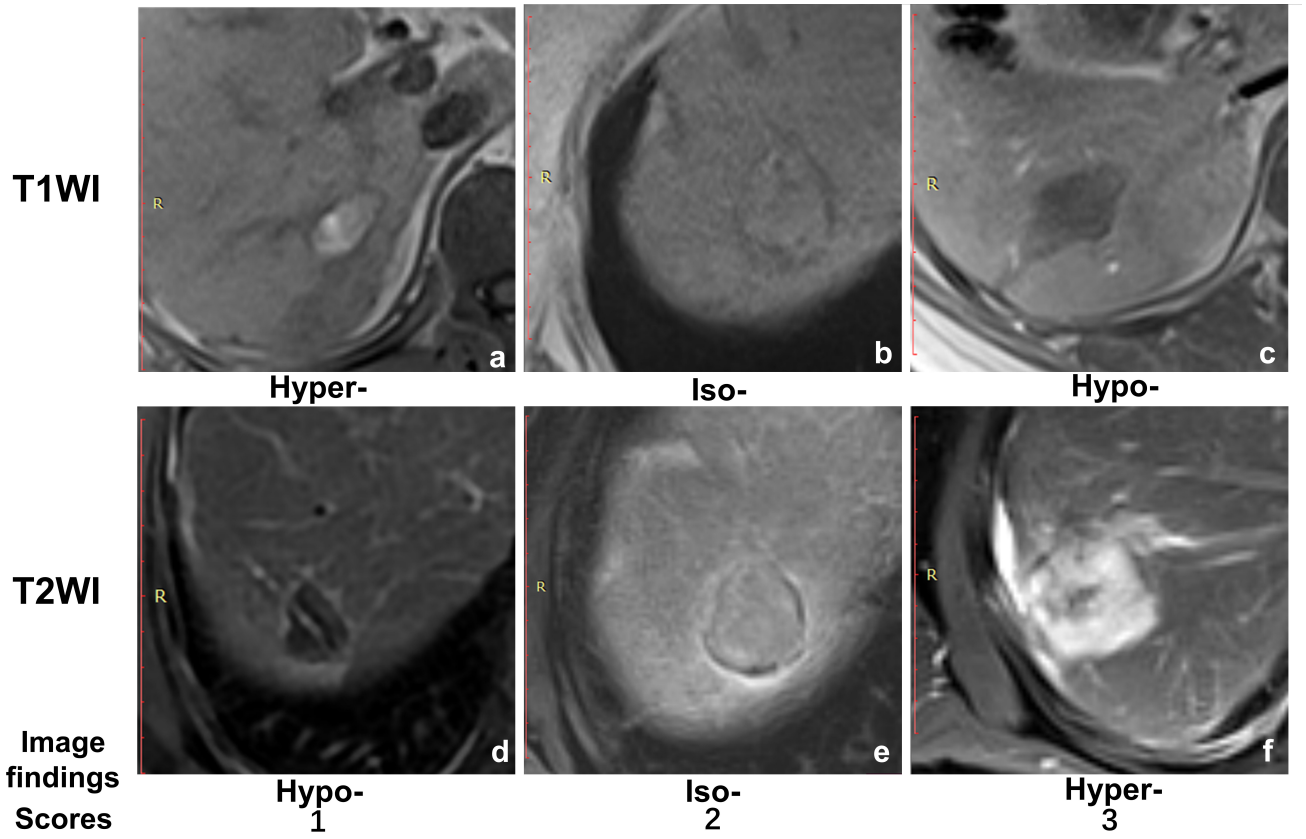

Fig. 2. The typical image signs of inner area signal intensity evaluation and scoring of MR Image Appearance Scoring at One Month after RFA (MRIAS 1MO). (a–c) The signal intensity with either hyper-, iso-, or hypo- on T1 weighted image (T1WI). (d–f) The different signal intensity scales and scoring on T2 weighted image (T2WI). The hypo-, iso-, and hyper-signal intensity were assigned to 1, 2, and 3 points respectively on T2WI.

MR Imaging

The MR imaging at one month post-RFA utilized a 3-T MRI scanner (Skyra, Siemens healthcare, Bavaria, Germany) with a standardized imaging protocol for all included patients. The imaging series included T1-weighted images (T1WI), T2-weighted images (T2WI), and three-phase dynamic contrast enhanced images. The detailed imaging parameters are listed in Table 1. The imaging process involved a bolus injection of gadolinium contrast agent (Magnevist, Bayer Healthcare Pharmaceuticals, Leverkusen, Germany) at a rate of 0.1 mmol/kg body weight intravenously at a rate of 2 mL/s, followed by a 20 mL saline flush. T1-weighted volumetric interpolated breath-hold series were repeated at 20 to 30 seconds, 70 to 80 seconds, and 180 seconds to obtain arterial, portal venous, and venous phase images.

MR Image Appearance Scoring at One Month after RFA (MRIAS 1MO)

To evaluate the imaging appearance one month post-RFA in a semi-quantitative fashion, we established a criterion based on MR image findings from T1WI, T2WI, and T1 contrast-enhanced images of venous phase (T1CEVP).

Initially, we divided the ablation zone into two areas: the inner area and marginal area. The central portion of the ablation zone was considered the inner area, while the peripheral section adjacent to the liver parenchyma interface was labeled the marginal area, typically appearing as a ring.

Then, the signal intensity strength in the inner area was evaluated in three scales: hyper, iso, and hypo. The distribution of signal intensity in the inner area was rated as either heterogeneous or homogeneous. The morphological features of the marginal area ring completeness and thickness were assessed. The ring completeness was rated as continuous, discontinuous, or absent. Using 2 mm as cutoff point, the ring was rated either as thick or thin. Additionally, on T1 contrast enhanced images of the venous phase, the ring was either rated as thick, thin, or irregular.

The typical imaging signs of a fully ablated zone included homogeneous hyper signal intensity on T1WI, hypo signal intensity on T2WI, and a lack of enhancement. Additionally, the ring appeared continuous, thin, and regular. We hypothesized that the more image appearance deviated from typical signs, the higher chance of subtotal ablation, increasing the risk of local tumor progression. Therefore, we assigned higher scores to image signs that ap-

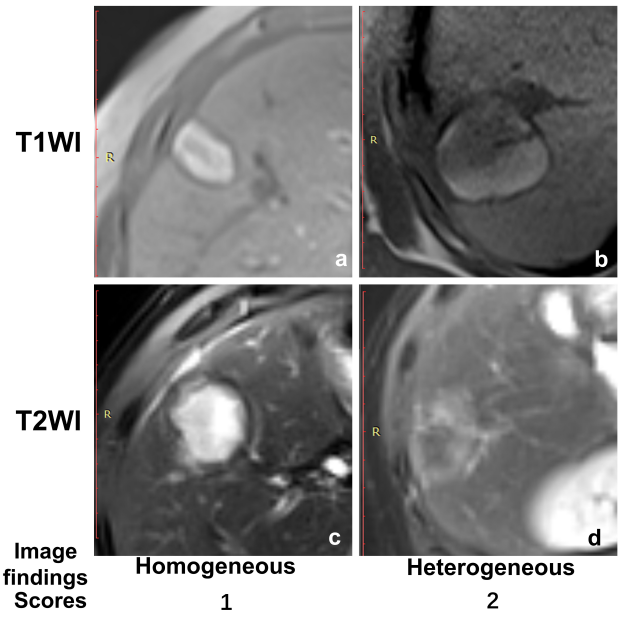

Fig. 3. The typical image signs of inner area signal intensity distribution evaluation and scoring of MR Image Appearance Scoring at One Month after RFA (MRIAS 1MO). (a,b) The signal intensity evaluated either as homogeneous or heterogeneous on the T1 weighted image (T1WI). (c,d) The signal intensity distribution on the T2 weighted image (T2WI). The homogeneous distribution was assigned as 1 point. The heterogeneous distribution was considered as 2 points.

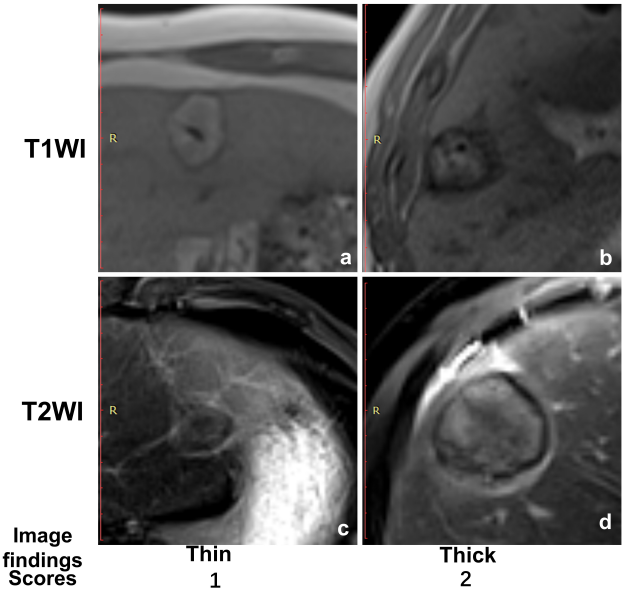

Fig. 4. The typical image signs of marginal area ring thickness evaluation and scoring of MR Image Appearance Scoring at One Month after RFA (MRIAS 1MO). (a,b) The ring thickness with evaluation of either thin or thick, which assigned to 1 or 2 points on the T1 weighted image (T1WI) and the T2 weighted image (T2WI) (c,d).

peared more different from typical signs. The detailed image findings and corresponding scores were summarized and demonstrated in Table 2 and Figs. 2,3,4,5,6.

Finally, we scored the post-ablation image findings on T1WI, T2WI and T1CEVP. Accordingly, we obtained four kinds of scores for each ablation zone, namely scores from T1WI (T1S), scores from T2WI (T2S), scores from non-contrast-enhanced images (NES) obtained by adding T1S to T2S, and the overall scores (MRIS) derived from the sum of NES and scores from T1CEVP.

Two radiologists, each with a minimum of five years of experience in abdominal MRI diagnosis, independently scored all included cases using MRIAS 1MO, unaware of the local response outcome. If the scores from two radiologists varied, the higher scores were used for data analysis. Before formal scoring, these two radiologists underwent a thirty-minute training session on MRIAS 1MO criteria and completed a trial of scoring for ten cases.

The Measurement Reliability Evaluation of MRIAS 1MO

The four types of scores (T1S, T2S, NES, and MRIS) derived from the MRIAS 1MO criteria by both radiologists were documented. The intraclass correlation coefficient (ICC) was calculated to evaluate measurement reliability, considering an ICC over 0.75 as indicative of good reliability.

Statistics

The comparison of clinical characteristics between training and validation cohorts was completed using Chi-square test. Then, the inter-rater reliability of MRIAS 1MO was evaluated by calculating the intraclass correlation coefficient (ICC), using an ICC value of 0.75 as the cutoff point for good reliability. Mann-Whitney U tests were performed to compare MRIAS 1MO scores among different LTP outcomes. Receiver operating characteristic curves (ROC) were used to evaluate the discriminative value of MRIAS 1MO scores on LTP and to identify an optimal cut-off point. Then, calculation and comparison of the area under the curve (AUC) for T1S, T2S, NES, and MRIS were conducted using the Delong test to determine the score type with the best performance. Clinical characteristics and MRIAS 1MO scores were screened by univariate analysis and enrolled to build a predictive LTP model by multivariate analysis using Cox regression. Evaluation of the predictive model's performance in both the training and validation sets was carried out using the concordance index (C index), calibration curve, and time-dependent C index curve. Furthermore, patients in both cohorts were classified into either high or low-risk groups based on the predictive model. The Kaplan-Meier survival analysis and log-rank test were used to test the discrimination value of the predictive model.

The patients' age, lesion diameter, and ablative margin were shown as median (interquartile range, IQR) due to their non-normal distribution. Additionally, Mann-Whitney U tests were used to compare these variables, with

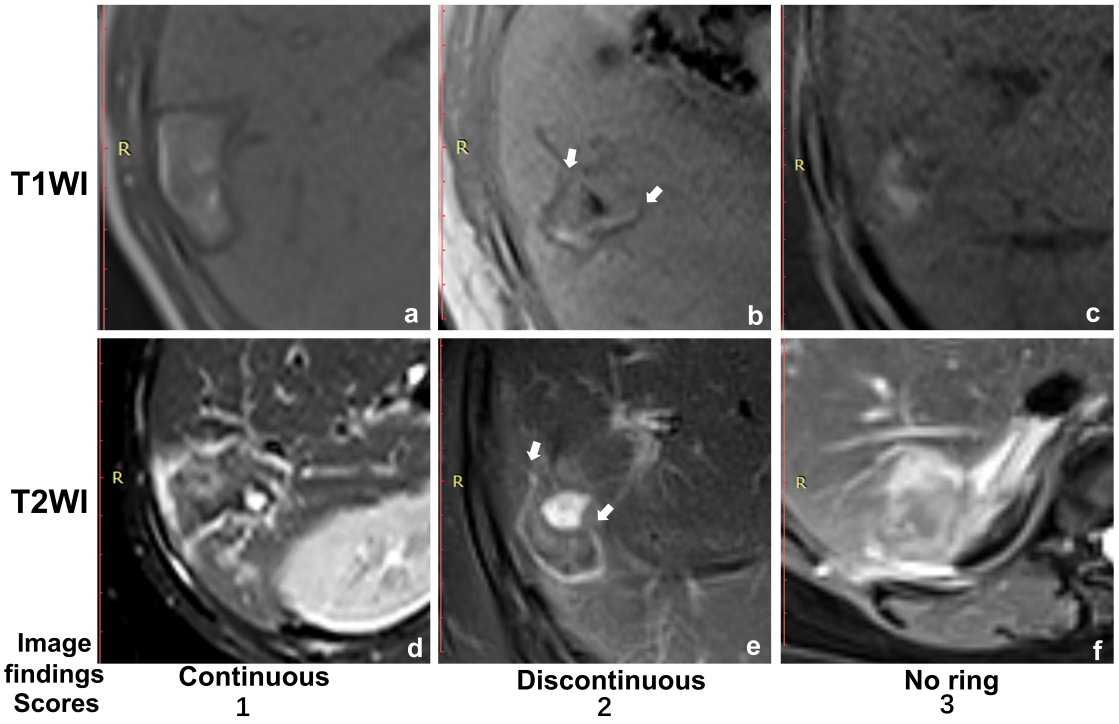

Fig. 5. The typical image signs of marginal area ring completeness evaluation and scoring of MR Image Appearance Scoring at One Month after RFA (MRIAS 1MO). (a–c) The ring completeness with evaluation of either as continuous, discontinuous, or absent, which assigned to 1, 2, or 3 points respectively on the T1 weighted image (T1WI). (d–f) The morphology of ring and corresponding points on the T2 weighted image (T2WI). The white arrows on (b) and (e) demonstrated the signs of discontinuous ring.

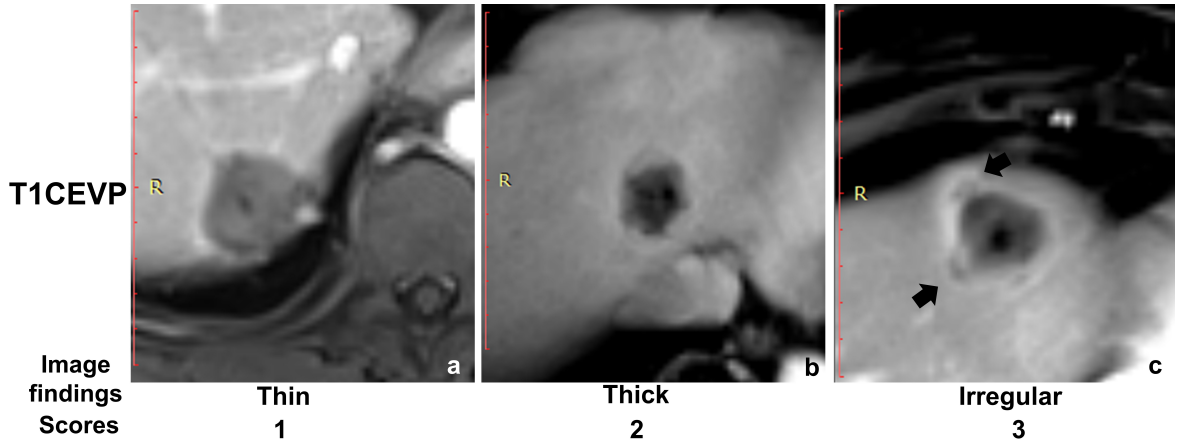

Fig. 6. The typical image signs of ring thickness evaluation and scoring of MR Image Appearance Scoring at One Month after RFA (MRIAS 1MO) on T1 contrast-enhanced images of venous phase (T1CEVP). (a–c) The ring morphology on contrast agent-enhanced images. The ring was evaluated either as thin, thick, or irregular which assigned to 1, 2, or 3 points respectively. The black arrows on (c) indicated irregular ring enhancement.

a significance level set at $p < 0.05$. All statistical analyses were performed using R software (V4.0.3, The R Foundation, Vienna, Austria).

Results

The median age was 58 (IQR: 14.0) years old and 60 (IQR: 16.8) years old in the training and validation groups, respectively ($p = 0.191$). The median lesion diameter was 23.0 (IQR: 11.85) mm in the training cohort and 23.5 (IQR: 13.2) mm in the invalidation cohort ($p = 0.645$). The median ablative margin was 4.0 (IQR: 4.5) mm in the training group and 5.0 (IQR: 5.9) mm in the invalidation group ($p = 0.203$).

The median follow-up time was 47.5 months (95% confidence interval (CI), 42.1–51.0 months) for the training cohort and 18.9 months (95% CI, 17.5–24.2 months)

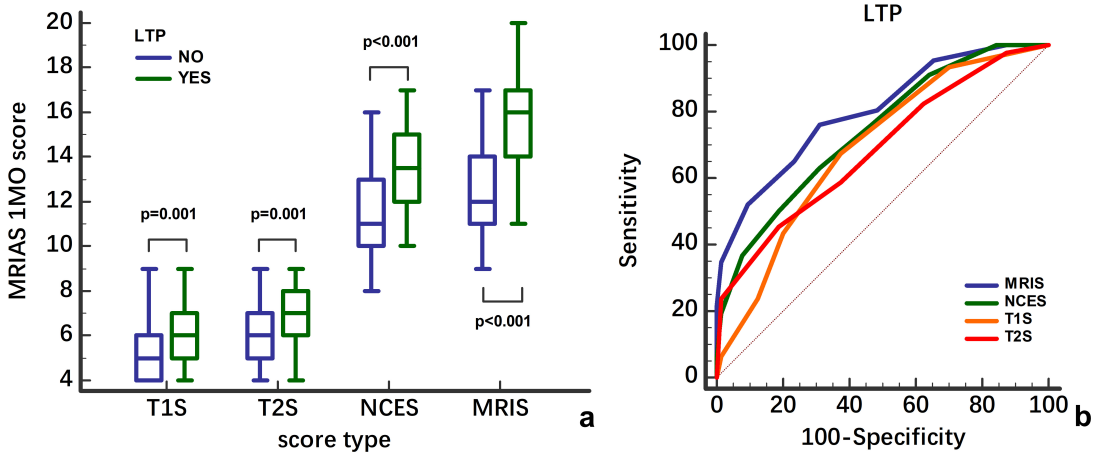

Fig. 7. The MR Image Appearance Scoring at One Month after RFA (MRIAS 1MO) scores comparison and receiver operating characteristic (ROC) analysis based on local tumor progression (LTP). The scores of four types were significantly higher in patients with LTP (a). The four types of scores were able to discriminate local control outcomes (b). The MRIS held the highest area under curve (AUC) value (blue line) compared to other score types.

for the validation cohort. The median LTPFS time were not observed during the study period in either cohorts. However, the mean LTP time was 39.8 months (95% CI, 34.3–45.2 months) and 20.9 months (95% CI, 17.6–24.2 months) in the training and validation cohorts, respectively. Table 3 lists and compares the other clinical characteristics between the training and validation groups, demonstrating comparable baselines in these cohorts.

The Measurement Reliability of MRIAS 1MO

The measurement reliability for the four types of MRIAS 1MO scores (T1S, T2S, NES, and MRIS) was good, with ICC values ranging from 0.779 to 0.873. Notably, T2S exhibited the lowest ICC value among the scores. The ICC values for the inner area components of T1S and T2S were over 0.75. However, the ICC value for the marginal area component of T1S was less than 0.75. The ICC value of marginal area component of T2S was even lower. The ICC values of most MRIAS 1MO scores were over 0.75, indicating good reliability. The detailed ICC value of MRIAS 1MO scores are listed in Table 4.

The MRIAS 1MO Scores of Different Local Response Outcomes

The median values of four kinds MRIAS 1MO scores were significantly higher in lesions with LTP (T1S: 6 vs. 5, $p = 0.001$; T2S: 7 vs. 6, $p = 0.001$; NES: 13.5 vs. 11, $p < 0.001$; MRIS: 16 vs. 12, $p < 0.001$). The detailed scores comparison is listed in Table 5 and illustrated in Fig. 7a.

The ROC analysis (Fig. 7b) showed that T1S, T2S, NCES, and MRIS were able to discriminate local control outcomes. The AUC values were from 0.685 to 0.801. The AUC value of MRIS was the highest compared to other kinds of scores (Table 6). Therefore, MRIS was employed in following predictive model training.

The Predictive Model's Performance and Validation

Using univariate analysis, lesion diameter exceeding 30 mm and MRIS over 13.5 demonstrated associations with LTP (Fig. 8a). Multivariate analysis confirmed that both factors remained independent risk factors for LTP. MRIS over 13.5 had a hazard ratio of 4.28 (95% CI, 2.15–8.49), while lesion diameter larger than 30 mm had a hazard ratio of 2.06 (95% CI, 1.09–3.88) (Fig. 8b). The calibration curve of six, twelve, and twenty-four months showed satisfactory to good accuracy in the training cohort (Fig. 9a–c). In the validation cohort, there was satisfactory to good accuracy (Fig. 10a–c). The overall C index was 0.721 (95% CI, 0.656–0.785). The time-dependent C index curve showed consistent values above 0.700 over time in both cohorts (Fig. 9d and Fig. 10d). The value of C index at six, twelve, and twenty-four months were 0.722, 0.726, 0.718 in the training cohort. In the validation cohort, the C index values were 0.825, 0.794, and 0.764 at six, twelve, and twenty-four months, respectively.

Based on LTP probability by the model, patients were classified into either high or low-risk groups. The LTP-free survival (LTPFS) curves showed a significant difference between high and low-risk groups in LTPFS time. The mean LTPFS time of training cohort was longer in the low-risk group than the high-risk one (50.3 months vs. 36.1 months, $p = 0.016$) (Fig. 9e). In the validation cohort, median LTPFS time was 10.0 months (95% CI, 4.4–20.3 months) in the high-risk group. The median LTPFS time was not reached in the low-risk group. The mean LTPFS time was longer in the low-risk group (26.7 months vs. 17.1 months, $p = 0.003$) (Fig. 10e).

Table 3. The clinical characteristics of training and validation cohorts.

Variables	Training cohort	Validation cohort	Chi-square value	*p* value
	N	N		
Number of patients	73	48		
Number of metastases	110	60		
Age				
≥60 yrs	30	25	1.410	0.235
<60 yrs	43	23		
Gender				
Male	51	39	1.970	0.160
Female	22	9		
Primary tumor location				
Colon	43	36	3.311	0.069
Rectum	30	12		
Metastases time				
Synchronous	32	19	0.027	0.868
Metachronous	41	29		
Oligo-metastases				
Yes	69	47	0.843	0.359
No	4	1		
Liver metastases number				
≥5	4	3	0.032	0.859
<5	69	45		
Liver metastases location				
Right lobe	89	47	0.161	0.688
Left lobe	21	13		
Proximity of liver capsule				
Yes	50	25	0.226	0.635
No	60	35		
Proximity of intrahepatic vessels				
Yes	14	8	0.013	0.910
No	96	52		
Ablative margin				
≥5 mm	46	30	1.051	0.305
<5 mm	64	30		
Lesion diameter				
≥30 mm	21	16	1.309	0.253
<30 mm	89	44		
Local tumor progression				
Yes	46	24	0.053	0.818
No	64	36		

Discussion

In the current study, we investigated the predictive value of post-RFA MR image signs in identifying LTP. Using a semi-quantitative approach, we assigned scores to post-ablation MRI appearances with the hypothesis that greater deviation from typical image signs correlated with an increased likelihood of non-complete ablation and LTP. Our findings aligned with this hypothesis, revealing higher scores in cases that experienced LTP (MRIS: 16 vs. 12). Based on this, we can describe the ablation zone on MRI with scores in a less subjective manner, offering prognostic value.

We developed the MRIAS 1MO, a semi-quantitative criteria focused on inner area signal intensity and marginal area morphology in post-RFA MRI findings. Ideally, the ablation zone was a sphere covering the entire target lesion, uniformly distributing energy to induce coagulative necrosis. The classical image appearance of inner area, homogeneous hyper- and hypo-signal intensity on T1WI and T2WI, represented the aforementioned tumor tissue destruction throughout the ablation zone, indicating complete ablation [18–20]. Hence, we assigned the lowest scores to these features. In previous studies, hyper-signal intensity on T1WI was considered as a prognosis factor in HCC [15]. However, the signal intensity distribution was rarely

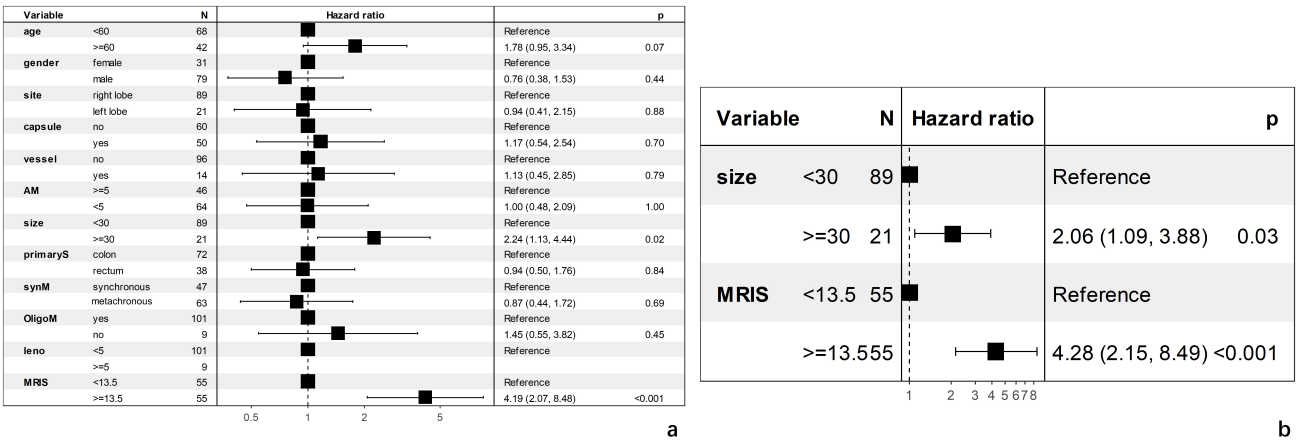

Fig. 8. The local tumor progression (LTP) prediction model building procedure. Among MRIS, lesion morphological features, and clinical characteristics, lesion size and MRIS were significantly related to LTP (a). The LTP prediction model was built based on MRIS and lesion size (b). MRIS, MRIAS 1MO overall scores.

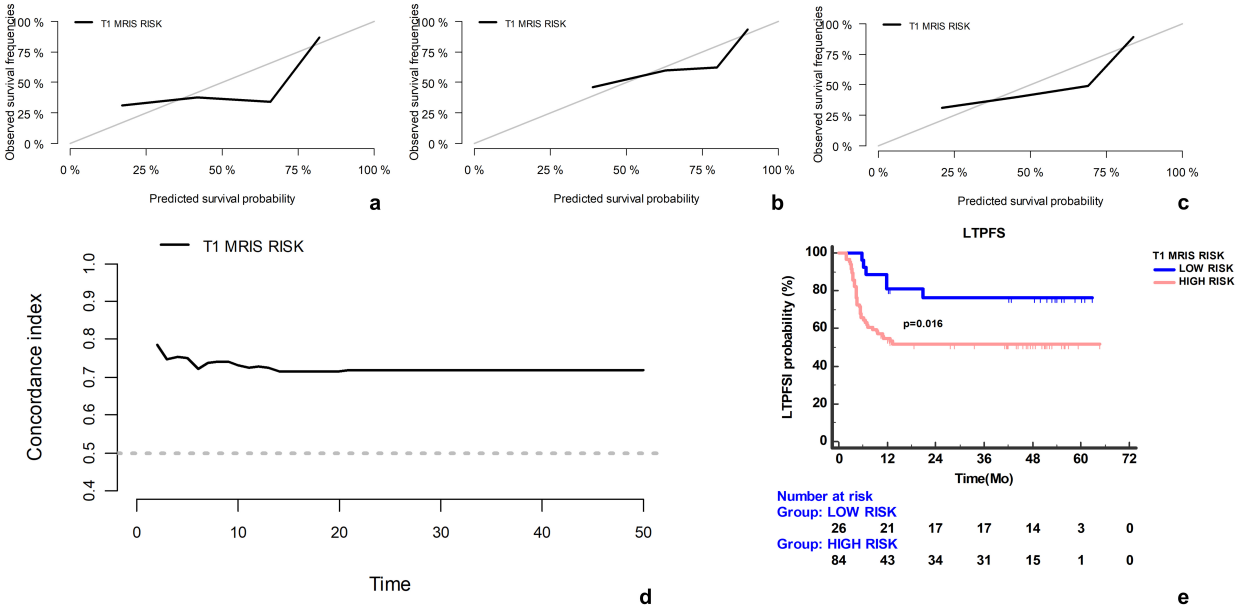

Fig. 9. The local tumor progression (LTP) prediction model evaluation in the training cohort. The calibration curve of six, twelve, and twenty-four months showed satisfactory to good accuracy (a–c). The C index values at six, twelve, and twenty-four months were 0.722, 0.726, and 0.718 according to time-dependent C index curve (d). Based on the LTP probability by the model, patients were classified into either high or low-risk groups. The LTP free survival (LTPFS) curves showed a significant difference between high and low-risk groups in LTPFS time (e).

considered. In the current study, we deemed that the even distribution of hyper- and hypo-signal intensity may be representative of homogeneous thermal energy transduced to the whole target lesion.

The other major component of MRIAS 1MO was scoring of the marginal area. The ablative margin is an important and well-known factor related to LTP. Unlike the surgical margin, which relies on direct examination of viable tumor tissue via microscope, assessing the ablative margin is solely based on evaluating ablation zone coverage through imaging. A five-millimeter coverage is often considered a safe margin. However, there is no universally recognized method for measuring the ablative margin. Previous studies have introduced three primary measurement approaches. The first involves halving the difference between lesion and ablation zone diameters to determine the ablative margin [21]. This method operates based on the hypothesis that the lesion and ablation zone centers are coincident, which may lead to either over or underestimation due to its deviation from the actual situation. The second approach was reported in a retrospective analysis focused on LTP [16]. The authors measured distances from lesion and

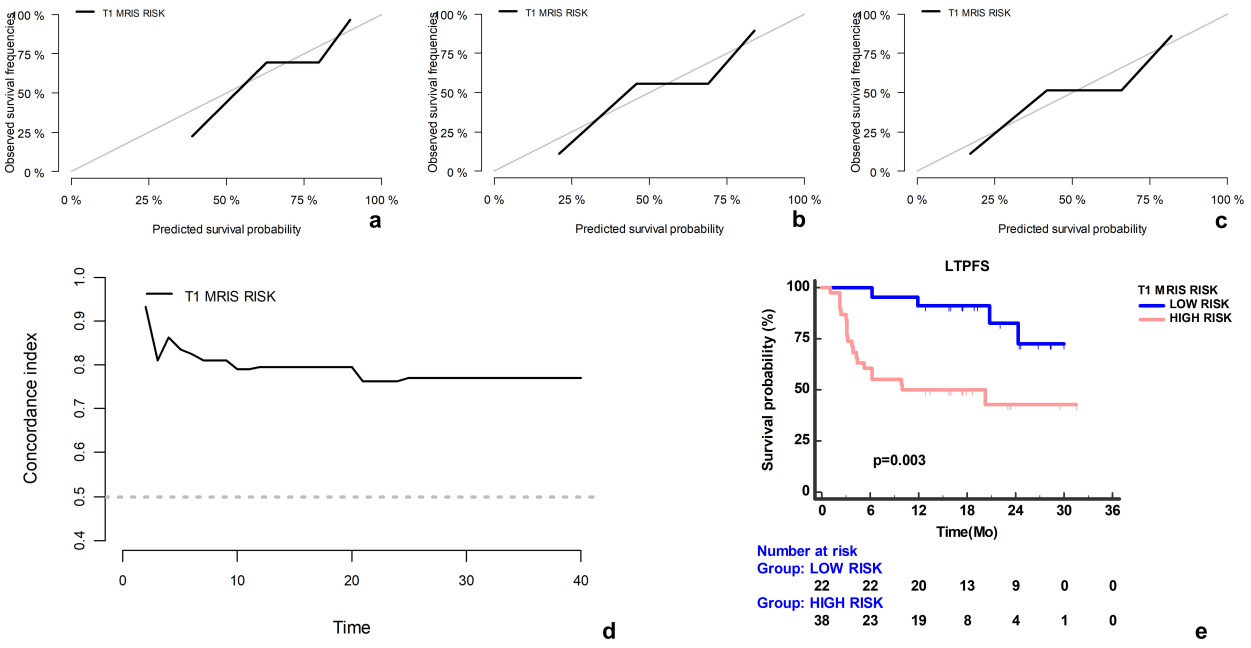

Fig. 10. The local tumor progression (LTP) prediction model evaluation in the validation cohort. The calibration curve of six, twelve, and twenty-four months showed satisfactory to good accuracy (a–c). The C index values at six, twelve, and twenty-four months were 0.825, 0.794, and 0.764 according to time-dependent C index curve (d). The LTP free survival (LTPFS) curve showed significantly shorter LTPFS time in high-risk group (e). The median LTPFS time was 10.0 months (95% CI, 4.4–20.3 months) in the high-risk group. The median LTPFS time was not reached in the low-risk group.

Table 4. Measurement Reliability of MRIAS 1MO scores.

Scores name	ICC	95% CI
T1S	0.825	0.745–0.878
Inner area of T1S	0.771	0.666–0.843
Marginal area of T1S	0.737	0.617–0.820
T2S	0.779	0.678–0.849
Inner area of T2S	0.747	0.631–0.826
Marginal area of T2S	0.513	0.290–0.666
NCES	0.826	0.746–0.881
MRIS	0.873	0.815–0.913

MRIAS 1MO, MR Image Appearance Scoring at One Month after RFA; ICC, intraclass correlation coefficient; CI, confident interval.

ablation zone edges to intrahepatic landmarks such as cysts, bile ducts, vessels, and the liver edge. Subtracting these two distances was considered as the ablative margin. We used this method to calculate the ablative margin, using 5 mm as the cutoff point. However, ablative margin less than 5 mm was not identified as a prognostic factor in the current study. The discrepancy may arise from measurement biases caused by non-contrast CT in our study. In addition, a study consisting of over 2500 liver tumors found no link between ablative margin and LTP risk [22]. Especially in CRCLM patients, the ablative margin was not related to local control [23]. The last method for measuring the ablative margin involved pre and post-ablation image registration [24]. This method considered the irregular shapes of lesion and ablation zones. However, image registration required solid and powerful motion correlation algorithms for the liver, which is a non-rigid and mobile organ.

Therefore, we attempted to assess the ablative margin by evaluating morphological features of the marginal area, primarily the thickness and continuity. Typically, the marginal area, formed by liver edema and inflammation, should exhibit a thin, continuous, and enhanced ring. The discontinuous and irregular enhancement were found to be risk factors of incomplete ablation [15].

The timing of imaging is crucial for evaluating the ablation zone, especially considering the dynamic changes in its appearance post-RFA, particularly in the early phases [25]. Early assessments are useful for detecting incomplete ablations and guiding re-ablation procedures instead of prognosis prediction. Although the immediate image findings complied with typical post-ablation signs, the early evaluation like those done 24 hours post-RFA were reported with limited insights into LTP prediction [26]. Furthermore, early abnormalities in peripheral liver perfusion can complicate the evaluation of the ablative margin. In this study, the routine follow-up time was one month post-RFA. Moreover, most cases of LTP occurred after six months follow-up post-RFA according to previous results [10]. Imaging at one month would not delay successive treatments once LTP was found. Therefore, we used one month post-RFA as imaging time.

Table 5. MRIAS 1MO scores comparison of local control outcomes.

Scores name	No LTP Median (IQR)	LTP Median (IQR)	*p* value	ROC results				
				Cutoff point	AUC (95% CI)	Youden index	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)
T1S	5 (2.0)	6 (2.3)	0.001	≥5.5	0.690 (0.595–0.775)	0.299	67.4 (52.0–80.5)	62.5 (49.5–74.3)
T2S	6 (2.0)	7 (2.3)	0.001	≥7.5	0.685 (0.590–0.771)	0.269	45.7 (30.9–61.0)	81.3 (69.5–89.9)
NCES	11 (3.0)	13.5 (3.3)	<0.001	≥12.5	0.741 (0.649–0.820)	0.318	63.0 (47.5–76.8)	68.8 (55.9–79.8)
MRIS	12 (3.0)	16 (3.3)	<0.001	≥13.5	0.801 (0.714–0.871)	0.448	76.1 (61.2–87.4)	68.8 (55.9–79.8)

MRIAS 1MO, MR Image Appearance Scoring at One Month after RFA; LTP, local tumor progression; IQR, interquartile range; AUC, area under curve; ROC, receiver operating characteristic; CI, confidence interval.

Table 6. AUC value comparison of MRIAS 1MO scores.

	p value
MRIS VS NCES	**0.006**
MRIS VS T1S	**0.010**
MRIS VS T2S	**0.003**
NCES VS T1S	0.178
NCES VS T2S	0.119
T1S VS T2S	0.947

AUC, area under the curve; MRIAS 1MO, MR Image Appearance Scoring at One Month after RFA. The AUC value of MRIS was significantly higher than other three types of scores (with bold *p* values).

We investigated the prognostic value of other clinicopathological factors on LTP. Lesion diameter larger than 3 cm emerged as a well-established predictor of LTP, consistent with findings from other studies. Although a shattered RFA probe might extend the ablation zone to 5 cm, the necrotic tissue could hinder adequate thermal energy transducing to the marginal area, potentially resulting in non-fatal tumor ablation and unseen residual tumor cells on imaging. Factors like lesion proximity to the liver capsule, intrahepatic vessels, and lesion location weren't found to be associated with LTP in our study, potentially indicating advancements in RFA technology and physicians' skills. In addition, we failed to find clinical factors related to the primary tumor, such as oligo metastases, synchronous metastases, and primary tumor location to be predictors of LTP. This inconsistent result might rise from variance in the characteristics of participants.

The prognostic predictive models based on radiomics were reported with amazing performance [27–29]. Nevertheless, the power of these models in clinical practice remains uncertain due to the lack of external validation. The poor reliability and reproducibility of radiomics and large variance of image quality impeded effective external validation. Furthermore, the need for specific imaging protocols and analysis software limits the widespread adoption of radiomics models across medical centers with varying expertise. Contrarily, our study analyzed routine MR images from widely used protocols with conventional image interpretation methods. This approach was independent on imaging equipment, imaging protocols, and specific analysis software. To mitigate subjective influences on image interpretation, we developed a semi-quantitative criterion, MRIAS 1MO, and evaluated its reliability at the study's outset. Reliability analysis demonstrated the independence of this criterion from image readers.

The current study had several limitations. First, the ICC values of T1S and T2S of the marginal area were lower than 0.75, indicating challenges in precise and clear marginal area evaluation. A simplified and explicit expression would be necessary in future studies. Secondly, the current MRIAS 1MO only considered a single imaging time point. The prognostic value of dynamic changes of MR images at multiple time points evoked our interest and would be investigated in successive studies. Finally, despite using a validation set from a different time frame, this study's internal validation design has limitations. External validation from multiple centers is essential to further evaluate the reliability of the MRIAS 1MO criteria.

Conclusions

The semi-quantitative MRIAS 1MO criteria developed in the current study yielded good reliability and prognostic potential for LTP prediction in CRCLM patients treated with RFA.

Availability of Data and Materials

The data sets generated during and/or analyzed during the current study are not publicly available, but are available from the corresponding authors on reasonable request.

Author Contributions

GYZ and WJP designed the study. GYZ and PL designed the methodology of study and analyzed all data. XHH, GZW, LCX, YHW and WTL performed the ablation procedure as well as documented all data related to RFA. GYZ and PL wrote the original draft. WJP and WTL made thoroughly review and editing of the manuscript. All authors contributed to editorial changes in the manuscript. All

authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

The Institutional Review Board of Fudan University Shanghai Cancer Center approved this study. The approval number was 1604159-3-1712A. All participants in this study signed informed consents.

Acknowledgment

Not applicable.

Funding

This study has been supported by Shanghai Pudong New Area Health Committee Science Project (Item No. PW2023A-47).

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Siegel RL, Wagle NS, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2023. CA: A Cancer Journal for Clinicians. 2023; 73: 233–254.

[2] Chow FCL, Chok KSH. Colorectal liver metastases: An update on multidisciplinary approach. World Journal of Hepatology. 2019; 11: 150–172.

[3] Ismaili N. Treatment of colorectal liver metastases. World Journal of Surgical Oncology. 2011; 9: 154.

[4] Heinemann V, von Weikersthal LF, Decker T, Kiani A, Vehling-Kaiser U, Al-Batran SE, *et al*. FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab as first-line treatment for patients with metastatic colorectal cancer (FIRE-3): a randomised, open-label, phase 3 trial. The Lancet. Oncology. 2014; 15: 1065–1075.

[5] Simkens LHJ, van Tinteren H, May A, ten Tije AJ, Creemers GJM, Loosveld OJL, *et al*. Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): a phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. Lancet. 2015; 385: 1843–1852.

[6] Czauderna C, Luley K, von Bubnoff N, Marquardt JU. Tailored Systemic Therapy for Colorectal Cancer Liver Metastases. International Journal of Molecular Sciences. 2021; 22: 11780.

[7] Nieuwenhuizen S, Dijkstra M, Puijk RS, Geboers B, Ruarus AH, Schouten EA, *et al*. Microwave Ablation, Radiofrequency Ablation, Irreversible Electroporation, and Stereotactic Ablative Body Radiotherapy for Intermediate Size (3-5 cm) Unresectable Colorectal Liver Metastases: a Systematic Review and Meta-analysis. Current Oncology Reports. 2022; 24: 793–808.

[8] Gillams A, Goldberg N, Ahmed M, Bale R, Breen D, Callstrom M, *et al*. Thermal ablation of colorectal liver metastases: a position paper by an international panel of ablation experts, The Interventional Oncology Sans Frontières meeting 2013. European Radiology. 2015; 25: 3438–3454.

[9] Takahashi H, Berber E. Role of thermal ablation in the management of colorectal liver metastasis. Hepatobiliary Surgery and Nutrition. 2020; 9: 49–58.

[10] Wang Y, Zhang GY, Xu LC, He XH, Huang HZ, Li GD, *et al*. Clinical Outcomes and Predictors in Patients With Unresectable Colorectal Cancer Liver Metastases Following Salvage Percutaneous Radiofrequency Ablation: A Single Center Preliminary Experience. Technology in Cancer Research & Treatment. 2020; 19: 1533033820963662.

[11] Solbiati L, Ahmed M, Cova L, Ierace T, Brioschi M, Goldberg SN. Small liver colorectal metastases treated with percutaneous radiofrequency ablation: local response rate and long-term survival with up to 10-year follow-up. Radiology. 2012; 265: 958–968.

[12] Chung WS, Park MS, Shin SJ, Baek SE, Kim YE, Choi JY, *et al*. Response evaluation in patients with colorectal liver metastases: RECIST version 1.1 versus modified CT criteria. AJR. American Journal of Roentgenology. 2012; 199: 809–815.

[13] Ayaz UY, Ayaz S. SUVmax-to-HU ratio in diagnosis of hepatic metastases of colon cancer on FDG PET/CT. A new semiquantitative parameter. Annali Italiani di Chirurgia. 2023; 94: 27–35.

[14] Kim SM, Shin SS, Lee BC, Kim JW, Heo SH, Lim HS, *et al*. Imaging evaluation of ablative margin and index tumor immediately after radiofrequency ablation for hepatocellular carcinoma: comparison between multidetector-row CT and MR imaging. Abdominal Radiology. 2017; 42: 2527–2537.

[15] Sheng RF, Zeng MS, Ren ZG, Ye SL, Zhang L, Chen CZ. Intrahepatic distant recurrence following complete radiofrequency ablation of small hepatocellular carcinoma: risk factors and early MRI evaluation. Hepatobiliary & Pancreatic Diseases International. 2015; 14: 603–612.

[16] Wang X, Sofocleous CT, Erinjeri JP, Petre EN, Gonen M, Do KG, *et al*. Margin size is an independent predictor of local tumor progression after ablation of colon cancer liver metastases. Cardiovascular and Interventional Radiology. 2013; 36: 166–175.

[17] Ahmed M, Solbiati L, Brace CL, Breen DJ, Callstrom MR, Charboneau JW, *et al*. Image-guided tumor ablation: standardization of terminology and reporting criteria–a 10-year update. Radiology. 2014; 273: 241–260.

[18] Sainani NI, Gervais DA, Mueller PR, Arellano RS. Imaging after percutaneous radiofrequency ablation of hepatic tumors: Part 2, Abnormal findings. AJR. American Journal of Roentgenology. 2013; 200: 194–204.

[19] Sainani NI, Gervais DA, Mueller PR, Arellano RS. Imaging after percutaneous radiofrequency ablation of hepatic tumors: Part 1, Normal findings. AJR. American Journal of Roentgenology. 2013; 200: 184–193.

[20] Bréhier G, Besnier L, Delagnes A, Oberti F, Lebigot J, Aubé C, *et al*. Imaging after percutaneous thermal and non-thermal ablation of hepatic tumour: normal appearances, progression and complications. The British Journal of Radiology. 2021; 94: 20201327.

[21] Takahashi H, Akyuz M, Aksoy E, Karabulut K, Berber E. Local recurrence after laparoscopic radiofrequency ablation of malignant liver tumors: Results of a contemporary series. Journal of Surgical Oncology. 2017; 115: 830–834.

[22] Yu J, Liang P, Yu XL, Cheng ZG, Han ZY, Mu MJ, *et al*. Local tumour progression after ultrasound-guided microwave ablation of liver malignancies: risk factors analysis of 2529 tumours. European Radiology. 2015; 25: 1119–1126.

[23] Liu CH, Arellano RS, Uppot RN, Samir AE, Gervais DA, Mueller PR. Radiofrequency ablation of hepatic tumours: effect of post-ablation margin on local tumour progression. European Radiology. 2010; 20: 877–885.

[24] Tani S, Tatli S, Hata N, Garcia-Rojas X, Olubiyi OI, Silverman SG, *et al*. Three-dimensional quantitative assessment of ablation margins based on registration of pre- and post-procedural MRI

and distance map. International Journal of Computer Assisted Radiology and Surgery. 2016; 11: 1133–1142.

[25] Tsuda M, Rikimaru H, Majima K, Yamada T, Saito H, Ishibashi T, *et al*. Time-related changes of radiofrequency ablation lesion in the normal rabbit liver: findings of magnetic resonance imaging and histopathology. Investigative Radiology. 2003; 38: 525–531.

[26] Ringe KI, Wacker F, Raatschen HJ. Is there a need for MRI within 24 hours after CT-guided percutaneous thermoablation of the liver? Acta Radiologica. 2015; 56: 10–17.

[27] Staal FCR, Taghavi M, van der Reijd DJ, Gomez FM, Imani F, Klompenhouwer EG, *et al*. Predicting local tumour progression after ablation for colorectal liver metastases: CT-based radiomics of the ablation zone. European Journal of Radiology. 2021; 141: 109773.

[28] Taghavi M, Staal F, Gomez Munoz F, Imani F, Meek DB, Simões R, *et al*. CT-Based Radiomics Analysis Before Thermal Ablation to Predict Local Tumor Progression for Colorectal Liver Metastases. Cardiovascular and Interventional Radiology. 2021; 44: 913–920.

[29] Yuan C, Wang Z, Gu D, Tian J, Zhao P, Wei J, *et al*. Prediction early recurrence of hepatocellular carcinoma eligible for curative ablation using a Radiomics nomogram. Cancer Imaging. 2019; 19: 21.