Original Article

Machine Learning using MR İmaging Radiomics can Predict the Response of Large Hepatocellular Carcinoma to Transarterial Radioembolization

O Sarioglu, A Canturk, RC Yarol, H Gulmez[1], FC Sarioglu, E Derebek[2], A Gulcu

Departments of Radiology and [2]Nuclear Medicine, Dokuz Eylul University Faculty of Medicine, Izmir, [1]Department of Family Medicine, İzmir Democracy University, Izmir, Turkey

Received: 11-May-2025;
Revision: 01-Aug-2025;
Accepted: 05-Aug-2025;
Published: 27-Sep-2025

ABSTRACT

Background: Large tumor size is associated with poor outcomes in patients with hepatocellular carcinoma (HCC). Although some studies have evaluated the treatment response of HCC to transarterial radioembolization (TARE), none of them used radiomics features with machine learning (ML) models in large tumors. **Aim:** To assess the performance of ML models using radiomics to predict the treatment response of TARE in large HCC lesions. **Methods:** This study included 49 patients with a large (>5 cm) HCC who underwent TARE. Treatment response was determined according to modified response evaluation criteria in solid tumors (mRECIST) criteria from the 3-month follow-up MR examinations. Complete or partial response was categorized as the responder group, while stable or progressive disease was classified as the non-responder group. Segmentation was performed on axial T2-weighted (T2W) and contrast-enhanced (CE) T1-weighted images. Classification learning models were used to create prediction models for TARE response. **Results:** Forty-nine patients (9 female, 40 male; mean age 63.58 ± 8.77) were included. None of the clinical, laboratory, and radiologic characteristics except the neutrophil counts showed statistical significance. Radiomics models obtained from CE-T1 and T2W images demonstrated an accuracy rate of 79.6%, while the area under the curve (AUC) rates were 0.92 and 0.77, respectively. The clinical model showed an accuracy rate of 77.6% and an AUC of 0.65. No statistically significant difference was found among all the models (P = 0.092). **Conclusion:** ML-based models constructed with radiomics features obtained from MR images before the TARE procedure might predict response in large HCC lesions.

Keywords: *Hepatocellular carcinoma, machine learning, radiomics, transarterial radioembolization, treatment response*

Introduction

Hepatocellular carcinoma (HCC) is one of the most common causes of cancer-related deaths globally, and its incidence has been increasing gradually.[1] Transarterial radioembolization (TARE) with yttrium-90 has been recently recommended in the Barcelona Clinic Liver Cancer (BCLC) group guidelines.[2] Contrast-enhanced (CE) magnetic resonance (MR) imaging is the recommended modality for evaluating the response after TARE.[3] A definite evaluation of the response to TARE lets prevent further locoregional therapies in HCC patients, which may lead to liver failure.

Large tumor size has been reported to be a determinant of poor outcomes in patients with HCC.[4,5] Although the current BCLC staging system still categorizes HCC

Access this article online

Quick Response Code:

Website: www.njcponline.com

DOI: 10.4103/njcp.njcp_305_25

Address for correspondence: Dr. O Sarioglu, Department of Radiology, Dokuz Eylul University Faculty of Medicine, 35340 İnciraltı - Izmir, Turkey.
E-mail: orkun.sarioglu@deu.edu.tr

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

How to cite this article: Sarioglu O, Canturk A, Yarol RC, Gulmez H, Sarioglu FC, Derebek E, *et al*. Machine learning using MR imaging radiomics can predict the response of large hepatocellular carcinoma to transarterial radioembolization. Niger J Clin Pract 2025;28:1111-20.

larger than 5 cm as stage A, resection and transplant are not always suitable for those patients because of their relation with microvascular invasion and higher histological grade.[6,7] TARE has been demonstrated as an alternative, safe, and effective treatment in large tumors by securing a high dose of radiation.[8]

Radiomics, which can quantify pixel intensities, has been used increasingly in scientific studies and clinical practice.[9] Extensive quantitative information can be provided from a tumor detected by computed tomography (CT) or MR imaging using radiomics programs. Recently, a few studies have included radiomics features for evaluating the treatment response of TARE in HCC patients.[10,11] Moreover, machine learning (ML) models built by radiomics features extracted from MR imaging have been submitted to the literature.[12,13] Even so, there are no studies in the literature assessing the prediction of treatment response using radiomics features in large HCC masses. Thus, this study aimed to build an ML-based model using radiomics to predict the treatment response of TARE in large HCC lesions.

Materials and Methods

Study population

The institutional review board approved this retrospective study (27.03.2024, 2024/12-25) and waived the informed consent requirement. We retrieved consecutive patients diagnosed with a large (>5 cm) HCC who underwent TARE between January 2015 and December 2023 from our database. The diagnosis of HCC was established based on CT and/or MR imaging findings. The decision to perform TARE on these patients was determined by a multidisciplinary tumor board. Only those who had CE abdominal MRI before the TARE treatment were included in the study. Exclusion criteria consisted of insufficient image quality, any prior treatment for HCC, and lack of follow-up after the procedure. The flowchart of the study is summarized in Figure 1. Laboratory results and clinical conditions of the subjects were noted from the medical records.

Image acquisition

All examinations were performed on a 1.5-T scanner (Philips Achiva Release 1.8, Eindhoven, The Netherlands) with a body coil. The specific parameters of axial T2-weighted (T2W) imaging were as follows: time of repetition (TR): 3.500 ms, time of echo (TE): 100 ms, layer thickness: 5 mm, layer spacing: 1 mm, matrix: 256 × 512, and field of view (FOV): 450 mm. The parameters of dynamic CE imaging were as follows: TR: 5.4 ms, TE: 0.1 ms, thickness: 3.6 mm, layer spacing: 0 mm, matrix: 260 × 224 mm, FOV: 450 mm, and parallel acceleration factor: 2. Intravenous gadobutrol (10 mL of 1 mmol/mL; Gadovist; Bayer, Berlin, Germany) was administered with a dose of 0.1 mL/kg in CE images. The flow rate was 2 mL/s. Image quality was evaluated to determine whether the images were suitable for radiomics measurements by two fifth-year radiology residents (A.C. and R.Y.) and a 7-year experienced interventional radiologist (O.S.) in consensus.

TARE procedure

TARE procedures were performed by an interventional radiologist (A.G.) with more than 15 years of experience. Hepatic angiograms and cone-beam CT were initially performed to plan the procedure and evaluate the lung shunt, distribution, and any possible extrahepatic accumulation by using technetium-99 m macro-aggregated albumin. One week later, TARE was done using Thera-Sphere (Boston Scientific) or SIR-Spheres (Sirtex) microspheres, depending on the operator's preference. Yttrium-90 microspheres were injected through the hepatic artery into the tumor, securing the normal liver parenchyma as much as possible. The dose was calculated using the Medical Internal Radiation Dosimetry dosimetry in Thera-Sphere and partition dosimetry in SIR-Sphere. The dose was planned to be >120 Gy in all procedures.

Study endpoints and assessments

The objective response of the index tumor represented the primary outcome of the study. Treatment response was assessed by the 7-year experienced interventional radiologist (O.S.) according to modified response evaluation criteria in solid tumors (mRECIST)[14] criteria from the 3-month follow-up MR examinations. Complete or partial response was categorized as the responder group, while stable or progressive disease was classified as the non-responder group.

Image analysis and segmentation

A freely available software (LIFEx, Version 7.4.0, www.lifexsoft.org) was used to derive radiomics features.[15] After all patients were anonymized, MR images performed within 1 month before TARE were downloaded from workstations in DICOM format and uploaded to the LIFEx software. One radiologist with 5 years of experience (A.C.), who was blinded to the TARE treatment results, performed image analysis and segmentation. Semi-automatic segmentation was constructed using a volumetric method on axial turbo spin echo T2W and arterial phase CE-T1 images [Figure 2]. The pseudocapsule was also included in the measurements during tumor segmentation. The voxel values of each lesion in three axes (x, y, z) were recorded, and the median values of these recorded data

were obtained. These median values were then utilized as optimized parameters in the texture analysis.[16] The intensity range was normalized by means of Z-scoring [mean ± 3 SD]. Image intensities were discretized into 128 fixed bins.[17] Finally, 175 radiomics features were obtained from each lesion. Exhaustive information is given in the Online Supplement Table.

Model training and evaluation

Radiomics features derived from preoperative CE-T1 and T2W MR images were transferred to the MATLAB R2020a (Math-Works, Natick, Massachusetts) program. Radiomic features with statistical significance were selected for the model creation. Receiver operating characteristic (ROC) curve analyses were performed for each statistically significant parameter to evaluate the ability to predict response. Classification learning models (CLMs) were used to create prediction models for TARE response. Bagged trees, fine Gaussian support vector machine (SVM), linear SVM, and quadratic SVM models were used, as they had the highest accuracy, positive predictive value (PPV), and negative predictive value (NPV) rates. A 5-fold cross-validation technique was utilized to identify the prediction rates of treatment response for TARE before the procedure. Accuracy, sensitivity, specificity, PPV, and NPV rates were noted for each model. ROC analysis and area under the curve (AUC) were used to evaluate the classification performance. We calculated the methodological radiomics score, a quality scoring tool for radiomics research endorsed by EuSoMII.[18] Our study's score was 80.3, with an excellent quality category.

Statistical analysis

Statistical analysis was done using the Statistical Package for the Social Sciences (SPSS) version 23.0 (SPSS Inc., Chicago, IL, USA). Continuous variables were defined as means and standard deviation, while categorical variables were defined as frequency counts and percentages. The Kolmogorov-Smirnov and Shapiro-Wilk tests were used to test normality. The Chi-square or Fisher's exact test was performed for categorical variables to compare the responder and non-responder groups. Mann–Whitney U test or independent *t*-test was used to compare continuous variables. The Friedman and Wilcoxon signed-rank tests were used to compare the results among the models. A *P* value of less than 0.05 indicated statistical significance.

Results

Clinical characteristics

A total of 49 patients (9 female, 40 male; mean age: 63.58 ± 8.77) met the inclusion criteria and were included in the study. The mean mass size was 7.56 ± 2.13 (range: 5–11 cm). Thera-Sphere was

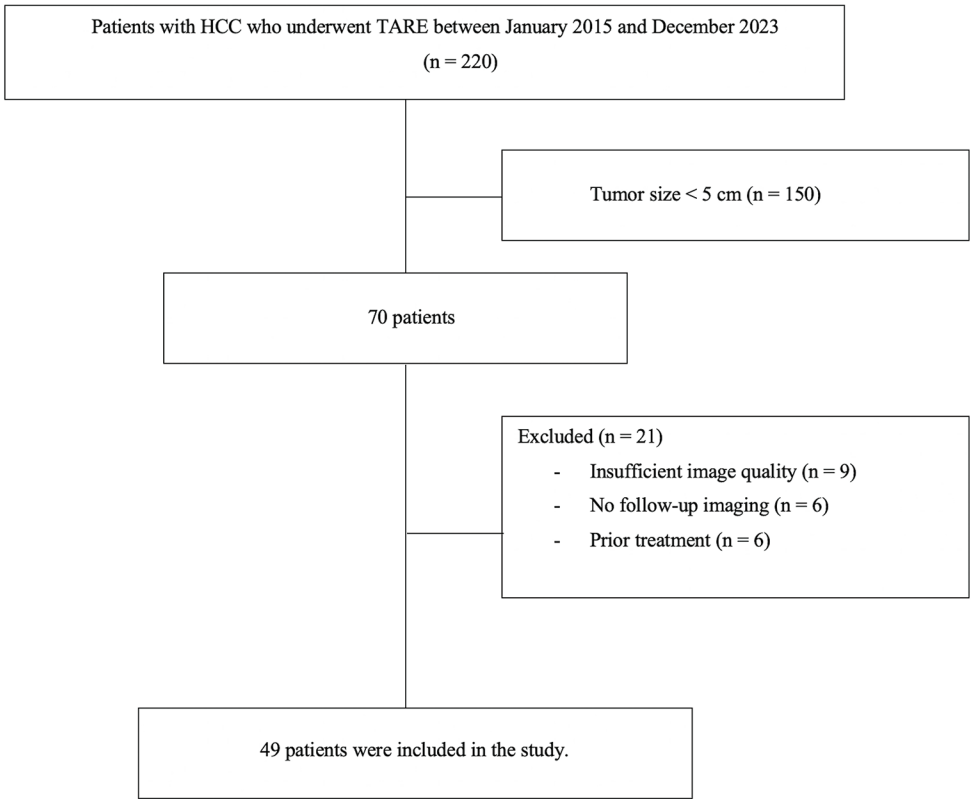


Figure 1: Flowchart of the study

Table 1: Clinical, imaging, and laboratory features of the cohort

	All of the patients (*n*=49)	Responder group (*n*=38)	Non-responder group (*n*=11)	*P*
Age (years)	63.58±8.77	63.34±9.01	64.45±8.23	0.778
Female/Male	9/40	8/31	1/9	0.864
Largest tumor size (cm)	7.56±2.13	7.59±2.24	7.40±1.67	0.998
Number of tumors				
1	26	22	4	0.210
2	7	6	1	
3	6	4	2	
4	3	3	0	
5+	7	3	4	
Cirrhosis etiology				
HBV	22 (44.9%)	15 (68.2%)	7 (31.8%)	0.117
HCV	8 (16.3%)	5 (62.5%)	3 (37.5%)	
ETOH	11 (22.5%)	11 (100%)	0	
NASH	5 (10.2%)	5 (100%)	0	
Other	3 (6.1%)	2 (66.7%)	1 (33.3%)	
PVTT				
Yes	12	9	3	0.913
No	37	29	8	
Child-Pugh score	6.94±1.04	6.95±1.1	6.91±0.83	0.818
Sodium	139.16±3.50	138.95±3.84	139.91±1.86	0.841
Creatinine	0.87±0.22	0.85±0.23	0.93±0.17	0.140
Total bilirubin	1.22±0.81	1.28±0.89	1.01±0.39	0.631
Albumin	3.71±0.47	3.70±0.47	3.77±0.52	0.623
INR	1.13±0.30	1.14±0.32	1.10±0.24	0.772
AST	55.71±22.50	57.16±24.54	50.73±12.90	0.369
ALT	47.43±43.12	50.26±47.35	37.64±22.18	0.368
Neutrophils (10^9/L)	3.57±1.52	3.31±1.33	4.47±1.85	**0.038**
Lymphocytes (10^9/L)	1.27±0.61	1.27±0.64	1.29±0.53	0.876
Platelets (10^9/L)	166.98±94.63	154.11±90.98	211.45±97.72	0.080
NLR	3.40±2.33	3.18±2.27	4.16±2.50	0.131
PLR	143.91±94.87	142.87±100.54	147.51±76.04	0.549
AFP (ng/mL)	1067.57±1438.67	1168.29±1590.05	819.64±638.23	0.876
Sphere type				
Thera-Sphere	37			0.785
SIR-Sphere	12			

HBV: hepatitis B virus, HCV: hepatitis C virus, ETOH: chronic alcohol use, NASH: nonalcoholic steatohepatitis, PVTT: portal vein tumor thrombus, NLR: neutrophil to lymphocyte ratio, PLR: platelet to lymphocyte ratio, AFP: α-fetoprotein. Bold values indicate statistically significant

administered in 37 patients, and SIR-Spheres were used in 12. At the 3-month follow-up, MR images revealed that 22 patients had a complete response, 16 had a partial response, six had stable disease, and five had progressive disease. Therefore, 38 patients (77.5%) were classified as responders, while 11 patients (22.5%) were classified as non-responders.

The clinical and laboratory features of the study population, along with comparisons between the responder and non-responder groups, are shown in Table 1. None of the clinical characteristics, including Child-Pugh score, cirrhosis etiology, age, and sex, showed statistical significance between the responder and non-responder groups. Radiologic features such as tumor size, number of tumors, and portal vein tumor thrombus did not show any statistical difference. Among the laboratory findings, only neutrophil counts were found to be significantly different between the groups. Neutrophil counts were found to be significantly higher in the non-responder group (P = 0.038). Other inflammatory markers, including the neutrophil to lymphocyte ratio (NLR) or platelet to lymphocyte ratio

Table 2: Comparison of the radiomics feature values between the two groups on T2-weighted images

Radiomics features	Responder group (*n*=38)	Non-responder group (*n*=11)	*P*
First-order			
Intensity Histogram Maximum Gray Level	127.05±3.38	122.72±9.84	0.032
GLCM			
Angular Second Moment	0.0005±0.0002	0.0006±0.0002	0.047
Inverse Difference	0.173±0.02	0.193±0.02	0.021
Inverse Difference Moment	0.09±0.01	0.11±0.02	0.021
GLRLM			
Short Runs Emphasis	0.978±0.005	0.974±0.004	0.026
Run Percentage	0.97±0.008	0.96±0.010	0.042
GLSZM			
Small Zone Emphasis	0.80±0.03	0.77±0.05	0.039
Normalized Zone Size Non-Uniformity	0.59±0.06	0.55±0.08	0.037
Zone Percentage	0.69±0.06	0.64±0.08	0.035

GLCM=gray level co-occurrence matrix, GLRLM=gray level run-length matrix, GLSZM=gray level size zone matrix

Table 3: Comparison of the radiomics feature values between the two groups on contrast-enhanced T1-weighted images

Radiomics features	Responder group (*n*=38)	Non-responder group (*n*=11)	*P*
First-order			
Intensity Histogram Entropy Log10	1.89±0.04	1.83±0.1	0.026
Intensity Histogram Entropy Log2	6.30±0.1	6.09±0.4	0.026
Intensity Histogram Uniformity	0.01±0.002	0.01±0.008	0.039
GLCM			
Joint Maximum	0.003±0.003	0.006±0.005	0.039
Entropy Log2	11.74±0.39	11.12±1.01	0.013
Entropy Log10	3.53±0.11	3.34±0.30	0.013
Difference Average	9.82±1.78	8.54±2.79	0.037
Sum Entropy	11.74±0.38	11.08±1.01	0.010
Angular Second Moment	0.0005±0.0022	0.0010±0.001	0.029
Normalized Inverse Difference	0.93±0.01	0.93±0.02	0.035
GLSZM			
Large Zone Emphasis	8.11±9.75	6.49±9.80	0.035
Zone Size Variance	5.29±8.91	5.83±8.60	0.031

GLCM=gray level co-occurrence matrix, GLSZM=gray level size zone matrix

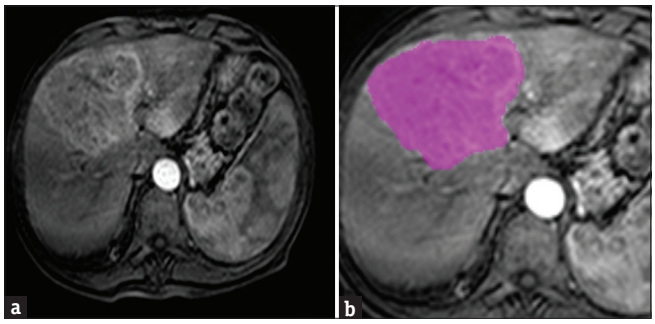


Figure 2: (a) A large heterogeneously enhanced mass is seen on the arterial phase of a contrast-enhanced T1-weighted image. (b) Segmentation of the tumor

(PLR), did not show any difference (P = 0.131 and P = 0.549, respectively).

Radiomics features

On preoperative T2W images, nine radiomics features, namely one first-order feature (Intensity Histogram Maximum Grey Level), three gray-level co-occurrence matrix (GLCM) features (GLCM_Angular Second Moment, GLCM_Inverse Difference, and GLCM_Inverse Difference Moment), two gray-level run-length matrix (GLRLM) features (GLRLM_Short Run Emphasis and GLRLM_Run Percentage), and three gray-level size zone matrix (GLSZM) features (GLSZM_Small Zone Emphasis, GLSZM_Normalized Zone Size Non Uniformity, and GLSZM_Zone Percentage), demonstrated statistically significant differences between the groups. On CE-T1 images, 12 radiomics features,

namely three first-order features (Intensity Histogram Entropy Log10, Intensity Histogram Entropy Log2, and Intensity Histogram Uniformity), seven GLCM features (GLCM_Joint Maximum, GLCM_Joint Entropy Log10, GLCM_Joint Entropy Log2, GLCM_Difference Average, GLCM_Sum Entropy, GLCM_Angular Second Moment, and GLCM_Normalized Inverse Difference), and two GLSZM features (GLSZM_Large Zone Emphasis and GLSZM_Zone Size Variance) were significantly different. All radiomics variables showing significant differences are elaborated in Tables 2 and 3. The ROC curve analysis results and optimal cutoff values for predicting response according to values on T2W and CE-T1 images are shown in Table 4.

Performance of the models to predict treatment response

The model utilizing radiomics features obtained from CE T1 (CE-T1) images demonstrated an accuracy rate of 79.6% for predicting response, with an area under the curve (AUC) of 0.92. The sensitivity and specificity rates for this model were 79% and 100%, respectively. In contrast, the model based on radiomics features extracted from T2W images achieved an accuracy rate of 79.6% and an AUC of 0.77, with a sensitivity of 80% and a specificity of 67%. There was no difference in the model performances between CE-T1 and T2W radiomics features (P = 0.157). When only clinical and laboratory parameters were used, the model showed an accuracy rate of 77.6% and an AUC of 0.65. The sensitivity and specificity values of the clinical and

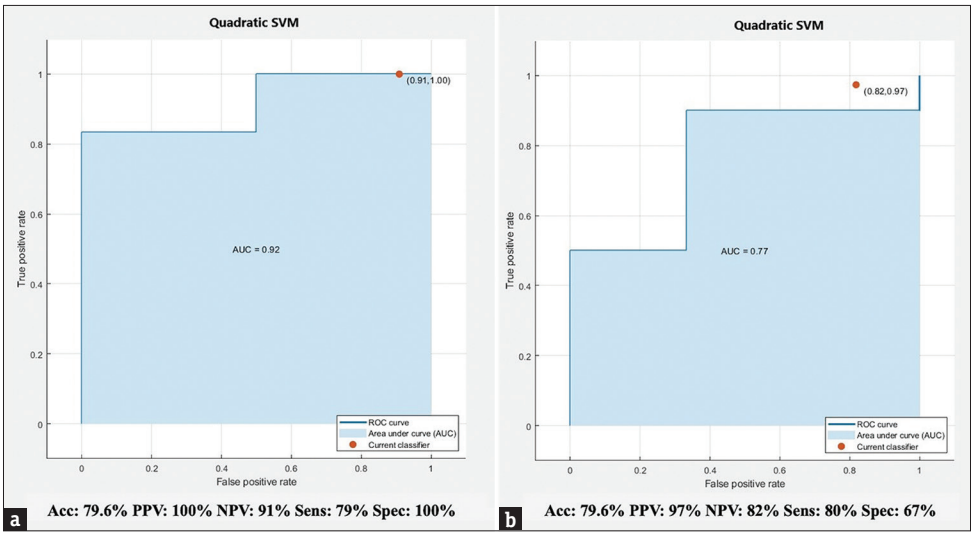


Figure 3: Receiver operating characteristic (ROC) curves of the radiomics models: (a) contrast-enhanced T1-weighted radiomics, (b) T2-weighted radiomics. Acc: accuracy, PPV: positive predictive value, NPV: negative predictive value, Sens: sensitivity, Spec: specificity

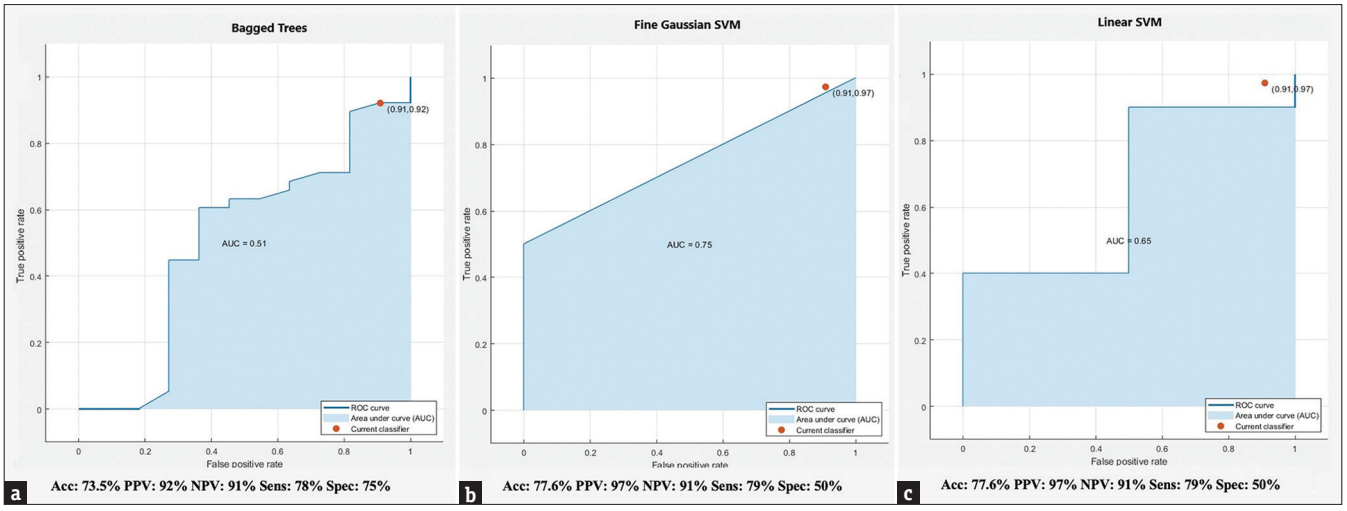


Figure 4: ROC curves of the clinicoradiomics models: (a) contrast-enhanced T1-weighted radiomics + clinical features, (b) T2-weighted radiomics + clinical features, (c) clinical features alone

Table 4: Diagnostic performances of the radiomics features with the cutoff values for predicting response on T2-weighted and contrast-enhanced T1-weighted imaging

	Area under the curve	95% confidence interval	Cutoff value	Sensitivity (%)	Specificity (%)	*P*
Radiomics features on T2-weighted imaging						
Intensity Histogram Maximum Gray Level	0.638	0.431–0.844	>126.5	89	40	0.045
GLCM_Angular Second Moment	0.699	0.526–0.871	<0.000487	70	65	0.047
GLCM_Inverse Difference	0.727	0.539–0.916	<0.193	81	65	0.023
GLCM_Inverse Difference Moment	0.730	0.529–0.930	<0.114	82	73	0.021
GLRLM_Short Runs Emphasis	0.722	0.527–0.918	>0.9725	90	64	0.026
GLRLM_Run Percentage	0.703	0.505–0.902	>0.9625	86	65	0.042
GLSZM_Small Zone Emphasis	0.706	0.501–0.910	>0.768	79	73	0.039
GLSZM_Normalized Zone Size Non-Uniformity	0.708	0.506–0.911	>0.540	79	73	0.037
GLSZM_Zone Percentage	0.711	0.512–0.909	>0.628	85	65	0.035
Radiomics features on CE T1-weighted imaging						
Intensity Histogram Entropy Log10	0.722	0.544–0.901	>1.897	74	73	0.026
Intensity Histogram Entropy Log2	0.722	0.544–0.901	>6.304	74	73	0.026
Intensity Histogram Uniformity	0.706	0.541–0.871	<0.0154	80	55	0.039
GLCM_Joint Maximum	0.706	0.522–0.899	<0.0045	79	64	0.039
GLCM_Entropy Log2	0.749	0.572–0.925	>11.677	71	73	0.013
GLCM_Entropy Log10	0.749	0.572–0.925	>3.5152	71	73	0.013
GLCM_Difference Average	0.708	0.506–0.910	>9.27	69	73	0.037
GLCM_Sum Entropy	0.758	0.579–0.938	>11.32	92	55	0.010
GLCM_Angular Second Moment	0.718	0.538–0.898	<0.000445	63	64	0.029
GLCM_Normalized Inverse Difference	0.489	0.087–0.492	>0.9344	35	65	0.035
GLSZM_Large Zone Emphasis	0.587	0.278–0.501	>7.91	92	30	0.035
GLSZM_Zone Size Variance	0.715	0.513–0.927	<5.13	71	73	0.031

GLCM=gray level co-occurrence matrix, GLRLM=gray level run-length matrix, GLSZM=gray level size zone matrix, CE=contrast-enhanced

laboratory model were 79% and 50%, respectively. Another model using both clinical and CE-T1 radiomics features showed an accuracy rate of 73.5%. Lastly, the model built with clinical and T2 radiomics features demonstrated an accuracy of 77.6%. The ROC analyses for all models, including accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), are illustrated in Figures 3 and 4. No statistically significant difference was found among all the models ($P = 0.092$).

Discussion

The present study assessed the performance of ML-based models to predict TARE response in large HCCs by using preprocedural radiomics and clinical data. We demonstrated that ML-based models with radiomics may contribute to the prediction of TARE response. Specifically, models that utilized radiomics data from pre-procedural CE T1-weighted (CE-T1) and T2W images achieved an accuracy of 79.6%. Conversely, when only clinical and laboratory features were used, all predictive metrics—including accuracy, sensitivity, specificity, and AUC values—showed a decline.

There are only a few studies that evaluated TARE response in HCC patients using an ML radiomics approach. One of them was conducted by Marinelli *et al*.,[12] which assessed the performance of an ML-based model for the prediction of early response by using MR imaging radiomics obtained before and early after treatment. Another study built ML models with radiomics from preprocedural MR imaging and clinical variables.[13] The biggest difference between our study and these two studies was that we included only HCC masses larger than 5 cm. Data regarding response rates of TARE in unresectable, large HCC lesions is limited in the literature.[19,20] To our knowledge, this is the first study to evaluate the performance of radiomics-based ML models specifically for large HCCs. We believe that our encouraging results using ML models to predict the response to TARE in large HCCs could contribute valuable insights for future research in this area.

The capacity of radiomics features to predict treatment response was previously evaluated by Aujay *et al*.[21] In their study, GLRLM and shape features were found to differ between the responder and non-responder groups. Similarly, we noticed that two GLRLM features (short runs emphasis and run percentage) were significantly different between the groups. Two main differences from this study were that we used ML and included large lesions. In a comparable study,[13] four first-order,

one shape, and three GLCM features were selected for the ML models. We also found one first-order and three GLCM features significantly different.

Radiomics obtained from different imaging modalities have been evaluated to predict response to TARE so far.[22] Pre-treatment[23] and post-treatment[24] PET/CT images and arterial and portal phase MR images[20] were used to extract radiomics features. In the present study, we used arterial phase CE-T1 and T2W MR images in texture analysis. Moreover, ML-based models were created in contrast to the many studies conducted before. We did not find any significant difference in the model performances between T2W and CE-T1 images. However, AUC (0.92), PPV (100%), and specificity (100%) rates were much higher when using CE-T1 images. Although our radiomics-based models achieved high AUC values, the absence of statistically significant differences among the models ($P = 0.092$) is likely attributable to the limited sample size, which may reduce statistical power and widen confidence intervals. In such small cohorts, even models with strong classification performance may not reach statistical significance. Nonetheless, the CE-T1-based model exhibited 100% specificity and high AUC, which suggests a potentially valuable clinical tool—especially in settings where traditional clinical predictors fall short. These findings indicate that radiomics-based ML models may aid in early treatment stratification in patients with large HCC lesions. However, further validation using larger and independent cohorts is essential to confirm these preliminary results and assess their generalizability. Performance rates decreased when using clinical features alone, or clinicoradiomics models with T2W and CE-T1 images. The close numeric range of clinical data might have caused these results. Further studies are needed to validate our findings.

Apart from radiomics features, we evaluated the clinical parameters to predict response to TARE. None of the clinical and laboratory variables, except neutrophils, differed between the groups in our study. Only neutrophil count was significantly higher in the non-responder group. This result was not surprising as neutrophil count is reported to be a significant marker of inflammation and immune response.[25] Age, Child-Pugh score, albumin, and bilirubin levels have been demonstrated as prognostic factors in different studies.[26,27] An ML model to predict TARE response using clinical and laboratory features was built and demonstrated an accuracy rate of 77.6%; however, the specificity (50%) and AUC (0.65) values were low. This likely reflects the inability of routine clinical and laboratory parameters to capture the biological heterogeneity of large HCCs or to reflect the microenvironmental and vascular changes induced by TARE. Integrating biomarkers and perfusion metrics into future models may offer a more comprehensive biological profile, potentially improving predictive accuracy and clinical decision-making.

Our study had some limitations. The retrospective design and the small number of subjects were the major limitations. We accept that the small number may decrease the robustness and generalizability of the study. Moreover, this limitation may cause some overfitting risks in our models, especially in the clinicoradiomics models, due to redundant variables. To mitigate these risks, we applied a 5-fold cross-validation strategy and limited model complexity by selecting simpler, well-established classifiers appropriate for small datasets. Additionally, we used feature selection and normalization techniques to reduce dimensionality and enhance model stability. Nonetheless, we recognize that validation on a larger, independent cohort is necessary to confirm our findings and improve generalizability. Second, controversy might happen during the evaluation of treatment response as unexpected imaging findings may occur after TARE.[28] Therefore, we assessed the 3-month follow-up MR images as early imaging can lead to inaccurate observation due to necrosis, hemorrhage, and edema. Moreover, a 5-fold cross-validation technique was used in the present study. Holdout validation set was not performed because the number of cases was small. Thus, separating the data into training and test would have caused information loss. The external validation method is a very important process in ML models. This method allows us to evaluate how well an ML model performs on new, unseen data samples. It offers simplicity, flexibility, and speed in assessing different algorithm techniques. In the future, external validation of these models with multi-institutional data is needed. Lastly, traditional imaging features such as enhancement pattern, tumor margin, or vascularity were not quantitatively assessed or included as separate variables in the present study. Our analysis focused solely on radiomics features extracted from preprocedural CE-T1 and T2W images. We believe that future work comparing radiomics-based models with conventional imaging descriptors could offer valuable insights and further validate the added benefit of radiomics in this setting.

In conclusion, our results demonstrated that ML-based models formed with radiomics features obtained from MR images before the TARE procedure could predict response in large HCC lesions. Further data with a large cohort is required to endorse our results.

Statement of authorship

Dr. Orkun Sarioglu, corresponding author

Address: Department of Radiology, Dokuz Eylul University Faculty of Medicine, Izmir, Turkey,

Tel: +90 232 412 22 22 Fax: +90 232 412 97 98
E-mail: orkun.sarioglu@deu.edu.tr

Dr. Ali Canturk

Address: Department of Radiology, Dokuz Eylul University Faculty of Medicine, Izmir, Turkey, Tel: +90 232 412 22 22 Fax: +90 232 412 97 98
E-mail: alicanturk.md@gmail.com

Dr. Raif Can Yarol

Address: Department of Radiology, Dokuz Eylul University Faculty of Medicine, Izmir, Turkey, Tel: +90 232 412 22 22 Fax: +90 232 412 97 98
E-mail: raifyarol@gmail.com

Dr. Hakan Gulmez

Address: Department of Family Medicine, İzmir Democracy University, Izmir, Turkey, Tel: +90 232 260 10 01 Fax: +90 232 260 10 04
E-mail: hakan.gulmez@gmail.com

Dr. Fatma Ceren Sarioglu

Address: Department of Radiology, Dokuz Eylul University Faculty of Medicine, Izmir, Turkey, Tel: +90 232 412 22 22 Fax: +90 232 412 97 98
E-mail: drcerenunal@gmail.com

Dr. Erkan Derebek

Address: Department of Nuclear Medicine, Dokuz Eylul University Faculty of Medicine, Izmir, Turkey, Tel: +90 232 412 22 22 Fax: +90 232 412 97 98
E-mail: erkan.derebek@deu.edu.tr

Dr. Aytac Gulcu

Address: Department of Radiology, Dokuz Eylul University Faculty of Medicine, Izmir, Turkey, Tel: +90 232 412 22 22 Fax: +90 232 412 97 98
E-mail: aytacgulcudeu@gmail.com

Acknowledgements

This study was carried out in Dokuz Eylul University Faculty of Medicine, Izmir, Turkey.

Authors' contributions

Concept - Orkun Sarioglu, Aytac Gulcu; Design - Orkun Sarioglu, Ali Canturk; Supervision – Aytac Gulcu; Data Collection and/or Processing – Ali Canturk, Raif Can Yarol; Analysis and/or Interpretation - Orkun Sarioglu, Hakan Gulmez; Literature Search - Orkun Sarioglu, Ali Canturk, Raif Can Yarol; Writing - Orkun Sarioglu; Critical Reviews – Erkan Derebek, Aytac Gulcu.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The study was approved by Institutional Review Board of Dokuz Eylul University.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Salem R, Gabr A, Riaz A, Mora R, Ali R, Abecassis M, *et al*. Institutional decision to adopt Y90 as primary treatment for hepatocellular carcinoma informed by a 1,000-patient 15-year experience. Hepatology 2018;68:1429-40.
2. Reig M, Forner A, Rimola J, Ferrer-Fàbrega J, Burrel M, Garcia-Criado Á, *et al*. BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. J Hepatol 2022;76:681-93.
3. Kloeckner R, Otto G, Biesterfeld S, Oberholzer K, Dueber C, Pitton MB. MDCT versus MRI assessment of tumor response after transarterial chemoembolization for the treatment of hepatocellular carcinoma. Cardiovasc Intervent Radiol 2010;33:532-40.
4. Pawlik TM, Delman KA, Vauthey JN, Nagorney DM, Ng IO, Ikai I, *et al*. Tumor size predicts vascular invasion and histologic grade: Implications for selection of surgical treatment for hepatocellular carcinoma. Liver Transpl 2005;11:1086-92.
5. Choi GH, Han DH, Kim DH, Choi SB, Kang CM, Kim KS, *et al*. Outcome after curative resection for a huge (>or=10 cm) hepatocellular carcinoma and prognostic significance of gross tumor classification. Am J Surg 2009;198:693-701.
6. Wu MY, Qiao Q, Wang K, Ji GW, Cai B, Li XC. development and validation of pre- and post-operative models to predict recurrence after resection of solitary hepatocellular carcinoma: A multi-ınstitutional study. Cancer Manag Res 2020;15:3503-12.
7. Gao SX, Liao R, Wang HQ, Liu D, Luo F. A nomogram predicting microvascular ınvasion risk in BCLC 0/A hepatocellular carcinoma after curative resection. Biomed Res Int 2019;25:9264137.
8. Kim HC, Kim YJ, Lee JH, Suh KS, Chung JW. Feasibility of boosted radioembolization for hepatocellular carcinoma larger than 5 cm. J Vasc Interv Radiol 2019;30:1-8.
9. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, *et al*. Radiomics: The bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749-62.
10. Li M, Fu S, Zhu Y, Liu Z, Chen S, Lu L, *et al*. Computed tomography texture analysis to facilitate therapeutic decision making in hepatocellular carcinoma. Oncotarget 2016;7:13248-59.
11. Blanc-Durand P, Van Der Gucht A, Jreige M, Nicod-Lalonde M, Silva-Monteiro M, Prior JO, *et al*. Signature of survival: A 18F-FDG PET based whole-liver radiomic analysis predicts survival after 90Y-TARE for hepatocellular carcinoma. Oncotarget 2017;9:4549-58.
12. Marinelli B, Chen M, Stocker D, Charles D, Radell J, Lee JY,

et al. Early prediction of response of hepatocellular carcinoma to yttrium-90 radiation segmentectomy using a machine learning MR ımaging radiomic approach. J Vasc Interv Radiol 2023;34:1794-801.
13. İnce O, Önder H, Gençtürk M, Cebeci H, Golzarian J, Young S. Prediction of response of hepatocellular carcinoma to radioembolization: Machine learning using preprocedural clinical factors and MR ımaging radiomics. J Vasc Interv Radiol 2023;34:235-43.
14. Marrero JA, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM, *et al*. Diagnosis, staging, and management of hepatocellular carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases. Hepatology 2018;68:723-50.
15. Nioche C, Orlhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, *et al*. LIFEx: A freeware for radiomic feature calculation in multimodality ımaging to accelerate advances in the characterization of tumor heterogeneity. Cancer Res 2018;78:4786-9.
16. Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. Magn Reson Imaging 2004;22:81-91.
17. Declaux G, Denis de Senneville B, Trillaud H, Bioulac-Sage P, Balabaud C, Blanc JF, *et al*. Assessment of a multivariable model using MRI-radiomics, age and sex for the classification of hepatocellular adenoma subtypes. Res Diagn Interv Imaging 2024;5;10:100046.
18. Kocak B, Akinci D'Antonoli T, Mercaldo N, Alberich-Bayarri A, Baessler B, Ambrosini I, *et al*. METhodological RadiomICs Score (METRICS): A quality scoring tool for radiomics research endorsed by EuSoMII. Insights Imaging 2024;15:8.
19. Tzedakis S, Sebai A, Jeddou H, Garin E, Rolland Y, Bourien H, *et al*. Resection postradioembolization in patients with single large hepatocellular carcinoma. Ann Surg 2023;278:756-62.
20. Meerun MA, Allimant C, Rivière B, Herrero A, Panaro F, Assenat E, *et al*. Large, multifocal or portal vein-invading hepatocellular carcinoma (HCC) downstaged by Y90 using personalized dosimetry: Safety, pathological results and outcomes after surgery. Hepatobiliary Surg Nutr 2023;12:351-65.
21. Aujay G, Etchegaray C, Blanc JF, Lapuyade B, Papadopoulos P, Pey MA, *et al*. Comparison of MRI-based response criteria and radiomics for the prediction of early response to transarterial radioembolization in patients with hepatocellular carcinoma. Diagn Interv Imaging 2022;103:360-6.
22. Mirza-Aghazadeh-Attari M, Srinivas T, Kamireddy A, Kim A, Weiss CR. Radiomics features extracted from pre- and postprocedural ımaging in early prediction of treatment response in patients undergoing transarterial radioembolization of hepatic lesions: A systematic review, meta-analysis, and quality appraisal study. J Am Coll Radiol 2024;21:740-51.
23. Karahan Şen NP, Alataş Ö, Gülcü A, Özdoğan Ö, Derebek E, Çapa Kaya G. The role of volumetric and textural analysis of pretreatment 18F-fluorodeoxyglucose PET/computerized tomography images in predicting complete response to transarterial radioembolization in hepatocellular cancer. Nucl Med Commun 2022;43:807-14.
24. Wei L, Cui C, Xu J, Kaza R, El Naqa I, Dewaraja YK. Tumor response prediction in 90Y radioembolization with PET-based radiomics features and absorbed dose metrics. EJNMMI Phys 2020;7:74.
25. Buonacera A, Stancanelli B, Colaci M, Malatino L. Neutrophil to lymphocyte ratio: An emerging marker of the relationships between the ımmune system and diseases. Int J Mol Sci 2022;23:3636.
26. Reincke M, Schultheiss M, Doppler M, Verloh N, Uller W, Sturm L, *et al*. Hepatic decompensation after transarterial radioembolization: A retrospective analysis of risk factors and outcome in patients with hepatocellular carcinoma. Hepatol Commun 2022;6:3223-33.
27. Rognoni C, Ciani O, Sommariva S, Facciorusso A, Tarricone R, Bhoori S, *et al*. Trans-arterial radioembolization in intermediate-advanced hepatocellular carcinoma: Systematic review and meta-analyses. Oncotarget 2016;7:72343-55.
28. Spina JC, Hume I, Pelaez A, Peralta O, Quadrelli M, Garcia Monaco R. Expected and unexpected ımaging findings after 90Y transarterial radioembolization for liver tumors. Radiographics 2019;39:578-95.