Am J Transl Res 2025;17(12):9542-9554
www.ajtr.org /ISSN:1943-8141/AJTR0168024

Original Article

Radiomic analysis of multiple MRI sequences for diagnosing liver cancer vs. focal nodular hyperplasia

Yufang Cai*, Wenbin Liang*, Lihao Wei*, Ruigang Huang, Zhixian Wu

*Department of Medical Imaging, Zhangzhou Affiliated Hospital of Fujian Medical University, Zhangzhou 363000, Fujian, China. *Equal contributors and co-first authors.*

Received August 10, 2025; Accepted November 25, 2025; Epub December 15, 2025; Published December 30, 2025

Abstract: Purpose: To develop and validate a radiomics model that uses multiple magnetic resonance imaging (MRI) sequences to accurately distinguish hepatocellular carcinoma (HCC) from focal nodular hyperplasia (FNH), thereby improving diagnostic precision and decision-making. Methods: We conducted a retrospective analysis including 196 patients (97 in HCC and 99 in FNH) diagnosed at the Zhangzhou Affiliated Hospital of Fujian Medical University (August 2011-December 2021). Radiomics features were extracted from the MRI images. LASSO logistic regression models were constructed for feature selection and to differentiate HCC from FNH. The model was further validated using a temporally independent cohort of 91 patients (49 HCC, 42 FNH) from the same institution (January 2022-December 2023). The area under the curve (AUC), accuracy, sensitivity, and specificity were used to evaluate the model's performance. Results: We obtained 34 features for T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI), and contrast enhanced imaging (CEI). The radiomics model demonstrated high diagnostic performance, with AUCs of 0.992 and 0.958 in the training and internal validation, respectively. In the independent external validation set, the model maintained strong performance with an AUC of 0.903, sensitivity of 88.9%, and specificity of 87.2%. In the training and internal validation, the model also showed high accuracy (0.956 and 0.867, respectively) and sensitivity (0.957 and 0.900, respectively). The integrated T2WI + DWI + CEI (TDC)-clinical data model demonstrated higher diagnostic accuracy than the TDC-only model. Conclusion: The developed multimodal MRI radiomics model effectively differentiated HCC from FNH and offers a non-invasive diagnostic tool that surpasses traditional imaging techniques. Further research is warranted to confirm these findings and explore the model's applications in broader clinical settings.

Keywords: Hepatocellular carcinoma, focal nodular hyperplasia, MRI, radiomics, diagnostic model, liver cancer

Introduction

Hepatocellular carcinoma (HCC) represents a significant contributor to cancer-related mortality worldwide, with approximately 45% of HCC-related deaths occurring in China [1, 2]. This disease imposes a considerable economic burden on patients, healthcare systems, and society, while also substantially impairing patients' quality of life [3, 4]. Conversely, focal nodular hyperplasia (FNH) is a benign hepatic condition that presents diagnostic challenges. Although typically asymptomatic and not necessitating treatment, FNH is often difficult to distinguish from HCC using conventional imaging techniques, posing a diagnostic dilemma [5-7]. The imperative of accurately differentiating FNH from HCC cannot be overstated, as the management and therapeutic approaches for these conditions differ markedly [8]. An erroneous diagnosis may result in either unnecessary aggressive interventions for a benign condition or inadequate treatment for a potentially life-threatening malignancy.

Recent advancements in radiomics have significantly transformed the diagnostic process. Radiomics employs high-throughput techniques to extract and analyze extensive quantities of advanced quantitative imaging features from standard medical images [9, 10]. This approach has demonstrated potential in enhancing diagnostic accuracy, particularly in differentiating between malignant hepatic conditions, such as hepatocellular carcinoma (HCC), and benign conditions, such as focal nodular hyperplasia

https://doi.org/10.62347/WSYB2367

(FNH). For example, sophisticated radiomics models that integrate data from multiple magnetic resonance imaging (MRI) sequences have shown promise in improving diagnostic accuracy, which is essential for informing appropriate patient management and treatment strategies.

However, the optimal combination of MRI sequences for distinguishing FNH from HCC remains an area of active investigation. To address this, our study sought to develop and validate an advanced model that integrates T2-weighted imaging (T2WI) and diffusion-weighted imaging (DWI) with enhanced sequences. We hypothesize that this multi-parametric approach can provide a superior framework for the accurate non-invasive differentiation of FNH from HCC, leveraging the distinct and complementary information provided by each sequence.

Materials and methods

Participants

By reviewing the electronic medical records, we retrospectively analyzed the clinical and imaging data of 196 patients who were pathologically diagnosed with either FNH or HCC at Zhangzhou Affiliated Hospital of Fujian Medical University, Fujian Province, from August 2011 to December 2021. This study adhered to the Declaration of Helsinki and was approved by the Clinical Research Ethics Committee of Zhangzhou Hospital in Fujian Province, which waived the need for obtaining informed consent from the subjects. Prior to evaluating the clinical and imaging data, the patient information was anonymized. The following inclusion criteria were used: (1) surgical resection, pathological confirmation by immunohistochemical results of FNH or HCC; (2) contrast-enhanced MRI of the liver or upper abdomen within 15 days prior to the operation, with complete imaging data of T2WI, DWI, and contrast-enhanced sequences. The following exclusion criteria were used: (1) HCC patients who received chemotherapy or radiotherapy before the operation, (2) patients with unsatisfactory image quality or unusable images, and (3) patients with incomplete clinical data or unfeasible statistical analysis. Of the 196 patients, 99 had FNH and 97 had HCC. This cohort constituted the development set. To robustly assess model performance and mitigate overfitting, we employed a 10-fold cross-validation strategy on this entire development cohort for internal validation, rather than a static data split.

Moreover, the external validation was achieved through retrospective evaluation of our hospital's FNH and HCC case records (January 2022-December 2023), maintaining uniform criteria for clinical and follow-up data acquisition. The external validation set was temporally independent from the development set. Notably, the MRI scanner was the same during the study period, providing a consistent test of the model's generalizability.

The diagnosis for every patient included in this study, encompassing both the development cohort (2011-2021) and the external validation cohort (2022-2023), was definitively established by histopathological analysis of surgical resection or biopsy specimens.

Sample size consideration

As a retrospective study, we included all consecutive eligible patients to maximize the sample size. Although a formal a priori sample size calculation was not performed, the final model incorporated only 3 key predictors. With 97 events (HCC cases) and 3 variables, the Events Per Variable (EPV) ratio of 32.3 far exceeds the recommended threshold of 10-20, ensuring model stability.

Data

The outcome variable was the pathological diagnosis of FNH or HCC. The predictors were the radiomics features extracted from the MRI images by ITK-SNAP software (Version 3.8, http://www.itksnap.org).

Data sources and image acquisition

We obtained contrast-enhanced MRI data from 196 patients with hepatic nodular lesions who were diagnosed with either HCC or FNH at our hospital between August 2011 to December 2021. A Philips Achieva 1.5T superconducting MRI scanner with an abdominal phased-array coil was used to perform the MRI scans. The patients underwent breathing training before scanning, and the following MRI sequences and parameters were used: T2WI with a repetition time (TR) of 2000 ms, echo time (TE) of 80

ms, field of view (FOV) of 375 mm × 297 mm, matrix size of 268 × 182, slice thickness of 5 mm, and slice gap of 1 mm; DWI with a TR of 2187 ms, TE of 73 ms, FOV of 375 mm × 375 mm, matrix size of 152 × 110, and b value of 800 s/mm^2; and contrast-enhanced imaging (CEI) with TR of 4.1 ms, TE of 1.98 ms, FOV of 370 mm × 297 mm, matrix size of 284 × 175, slice thickness of 5 mm, and slice gap of 1 mm. A standardized contrast-enhanced protocol was followed using the extracellular agent gadopentetate dimeglumine. The MRI images were stored in Digital Imaging and Communications in Medicine format and downloaded from the Picture Archiving and Communication System of our hospital.

Image segmentation and radiomics feature extraction

We used ITK-SNAP software (version 3.8, http://www.itksnap.org) to segment the lesions from the MRI images. We selected three MRI sequences (T2WI, DWI, and CEI) and imported them into the software. For the contrast-enhanced imaging (CEI), radiomics features were extracted from the portal venous phase. This phase was selected for radiomics analysis because it provides the most consistent and reliable lesion-to-liver contrast for characterizing both FNH and HCC, and is less susceptible to flow-related artifacts compared to the arterial phase. To assess the inter-observer reproducibility of feature extraction, two radiologists who were blinded to the pathological results independently performed the manual segmentation. Both readers delineated the entire contour of each lesion slice by slice, carefully avoiding blood vessels and surrounding normal liver tissues. The segmented contours were merged to form a three-dimensional region of interest (ROI) for each lesion. A randomly selected subset of 50 cases (approximately 25% of the cohort) was used for this reproducibility analysis. After independent segmentation, the ROIs from both readers were used to extract radiomics features. The inter-observer agreement was quantified using the intraclass correlation coefficient (ICC) for each feature. Features with an ICC greater than 0.75 in the subset were considered to have good to excellent reproducibility and were retained for subsequent analysis. Any disagreements in the remaining cases were resolved through consensus discussion between the two readers. We used the PyRadiomics package (version 3.0, https://pyradiomics.readthedocs.io) to extract 1,688 radiomics features from each ROI. The extraction was performed using the default parameter settings of PyRadiomics unless otherwise specified. This included applying a fixed bin width of 25 for discretizing the image gray levels and calculating features from the original image without applying any filters. Feature classes extracted comprised first-order statistics, shape-based (3D) features, and texture features including Gray Level Co-occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), Gray Level Size Zone Matrix (GLSZM), Neighboring Gray Tone Difference Matrix (NGTDM), and Gray Level Dependence Matrix (GLDM).

Statistical analysis

Software and clinical data analysis: We used R software (version 3.8.0) for the analysis of clinical data. The Shapiro-Wilk test was used to assess the normality of the age distribution. The Kruskal-Wallis test was employed to compare age between the FNH and HCC groups. Categorical variables, such as gender, were compared using the chi-square test or Fisher's exact test, as appropriate. A p-value < 0.05 was considered statistically significant.

Radiomics feature processing and model construction: For radiomics data, we used Python (version 3.7.0) with the Scikit-learn (version 0.19.2) and Pyradiomics (version 3.0.1) packages for feature extraction, normalization, feature selection, and model construction.

Feature preprocessing and selection: Feature selection was performed separately for the features extracted from each MRI sequence (T2WI, DWI, and CEI) through a three-step pipeline: i) variance Thresholding: We first removed radiomics features with near-zero variance (variance threshold < 0.8) across the entire training cohort; ii) univariate feature selection: The remaining features were subjected to univariate logistic regression analysis. Features with a significant association with the pathological outcome (HCC vs. FNH, P < 0.05) were retained; iii) LASSO Regression: Finally, the Least Absolute Shrinkage and Selection Operator (LASSO) regression algorithm with 10-fold cross-validation was applied to further

Table 1. Comparison of the characteristics between the FNH and HCC groups

	HCC (97)	FNH (99)	U/χ^2	P
Age (years; median [IQR])	62.00 (54.00, 68.00)	33.00 (25.50, 42.00)	105.500	< 0.001
Sex			48.460	< 0.001
Male, n (%)	84 (86.6)	38 (38.4)		
Female, n (%)	13 (13.4)	61 (61.6)		

FNH, focal nodular hyperplasia; HCC, hepatocellular carcinoma; IQR, interquartile range.

reduce dimensionality and prevent overfitting. The optimal regularization parameter (λ) for LASSO was determined by selecting the value that yielded the most parsimonious model within one standard error of the minimum binomial deviance from the cross-validation. Regarding the 10-fold cross-validation for LASSO, the class distribution (FNH vs. HCC) in the training set was approximately balanced (99 vs. 97). Therefore, standard cross-validation without specific stratification was deemed appropriate, as it would likely yield folds with representative class proportions.

Model building and evaluation: The features with non-zero coefficients derived from the LASSO regression for each sequence were used to construct the final logistic regression (LR) predictive models. The diagnostic efficacy of each model was evaluated in both the training and validation sets using receiver operating characteristic (ROC) curve analysis. Performance was quantified by the area under the curve (AUC), accuracy, sensitivity, and specificity.

Results

Characteristics of the study population

The clinical characteristics of the FNH and HCC groups are shown in **Table 1**. The median age of the HCC group was 62.00 years (interquartile range, 54.00-68.00), which was significantly higher than the 33.00 years (interquartile range, 25.50-42.00) of the FNH group (U=105.500, P < 0.001). A significant difference was observed in sex distribution between the two groups (χ^2=48.460, P < 0.001), with a higher proportion of males in the HCC group (86.6%) compared to the FNH group (38.4%).

Radiomics feature selection and model construction

The inter-observer reproducibility was assessed on a randomly selected subset of 50 cases. Based on the intraclass correlation coefficient (ICC) analysis, 1,502 out of the total 1,688 features (89.0%) demonstrated excellent consistency (ICC > 0.75) and were therefore retained for subsequent feature selection and model construction. This confirms the high reliability of the feature extraction process. To reduce dimensionality and select important radiomics features, we performed three-step processing of the imaging features extracted from various sequences and their combinations. The following abbreviations are used for the different feature sets: (1) T2WI: Features from the T2-weighted imaging sequence alone. (2) TD: Combined features from both T2WI and DWI sequences. (3) TC: Combined features from both T2WI and CEI (contrast-enhanced imaging) sequences. (4) TDC: Combined features from all three sequences: T2WI, DWI, and CEI.

First, we used variance thresholding to exclude features with a variance threshold < 0.8. Second, we performed univariate LR to identify features significantly associated with the diagnosis (HCC vs. FNH). Finally, LASSO regression further reduced the feature set and prevented overfitting. The LASSO regression algorithm selected 23, 27, 24, and 34 radiomic features from the T2WI, TD, TC, and TDC feature sets, respectively (as visualized in **Figures 1**-**4**). 2WI provides the foundational information, and DWI adds unique, complementary value. The TC model (T2WI + CEI), comprising 24 features, retained 23 features from T2WI and 1 additional feature from CEI. This suggests that the hemodynamic information from CEI, while highly specific, may partially overlap with the textural information captured by T2WI. Critically, the superior TDC model (T2WI + DWI + CEI), with 34 features, was not merely the sum of its parts. It integrated a unique set of features, indicating a synergistic effect where the combination of all three sequences captures diagnostic information that is not accessible when any sequence is used in isolation. The significant performance gain of the TDC model under-

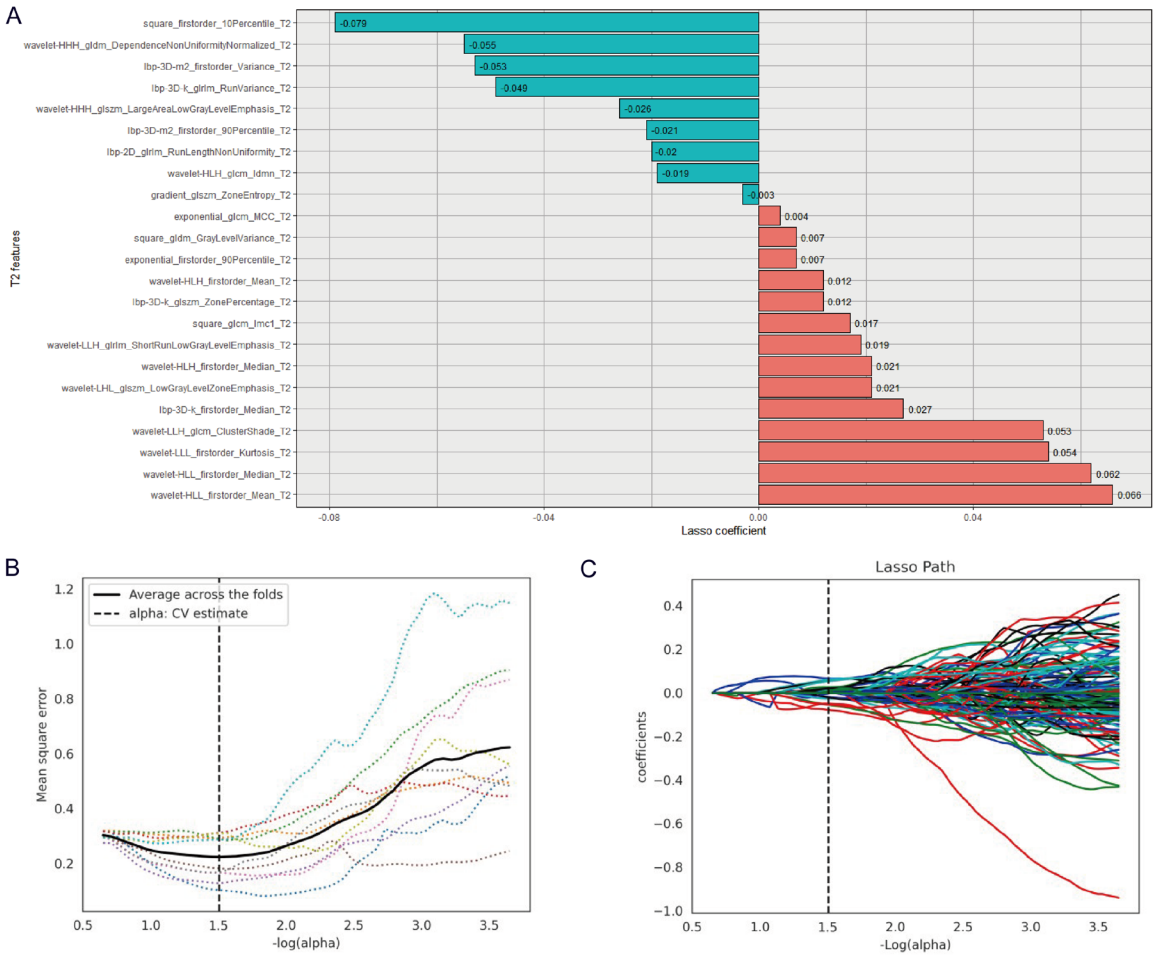

Figure 1. T2WI single-sequence group radiomics model construction. A: Image features and correlation coefficients of the minimum absolute contraction and selection operator (LASSO); B: Best tuning parameter (λ) selected by 10-fold cross-validation in the LASSO regression model; C: λ used to obtain the radiomics features of the non-zero series (each colored line represents the change of its coefficient).

scores that features reflecting tissue structure (T2WI), cellularity (DWI), and vascularity (CEI) each provide indispensable and complementary diagnostic value. These selected features were then used to build the corresponding predictive models.

Diagnostic performance of the three models

Using imaging characteristics, we constructed predictive models and evaluated their diagnostic efficacy for FNH via ROC curves. As detailed in **Table 2** and **Figure 5**, the T2WI-based model produced AUC values of 0.942 (95% CI: 0.909-0.968) in the training cohort and 0.814 (95% CI: 0.707-0.895) in the validation cohort. For the TD-based model, the AUCs were 0.967 (95% CI: 0.941-0.988) and 0.891 (95% CI: 0.802-0.957) in the training and validation sets, respectively. The TC-based model produced AUC values of 0.963 (95% CI: 0.940-0.984) for training and 0.901 (95% CI: 0.820-0.970) for validation. Of particular note, the TDC-based model attained the highest training AUC (0.989, 95% CI: 0.978-0.997) and the strongest validation concordance (AUC: 0.944, 95% CI: 0.886-0.987), with 90.0% sensitivity and 86.7% specificity, underscoring its exceptional diagnostic utility.

Building and validation of the combined diagnostic model

To integrate clinical risk factors with the radiomics signature, we developed a combined model. The Rad-score from the optimal TDC radiomics model, along with the clinical variables of age and sex (which were selected due

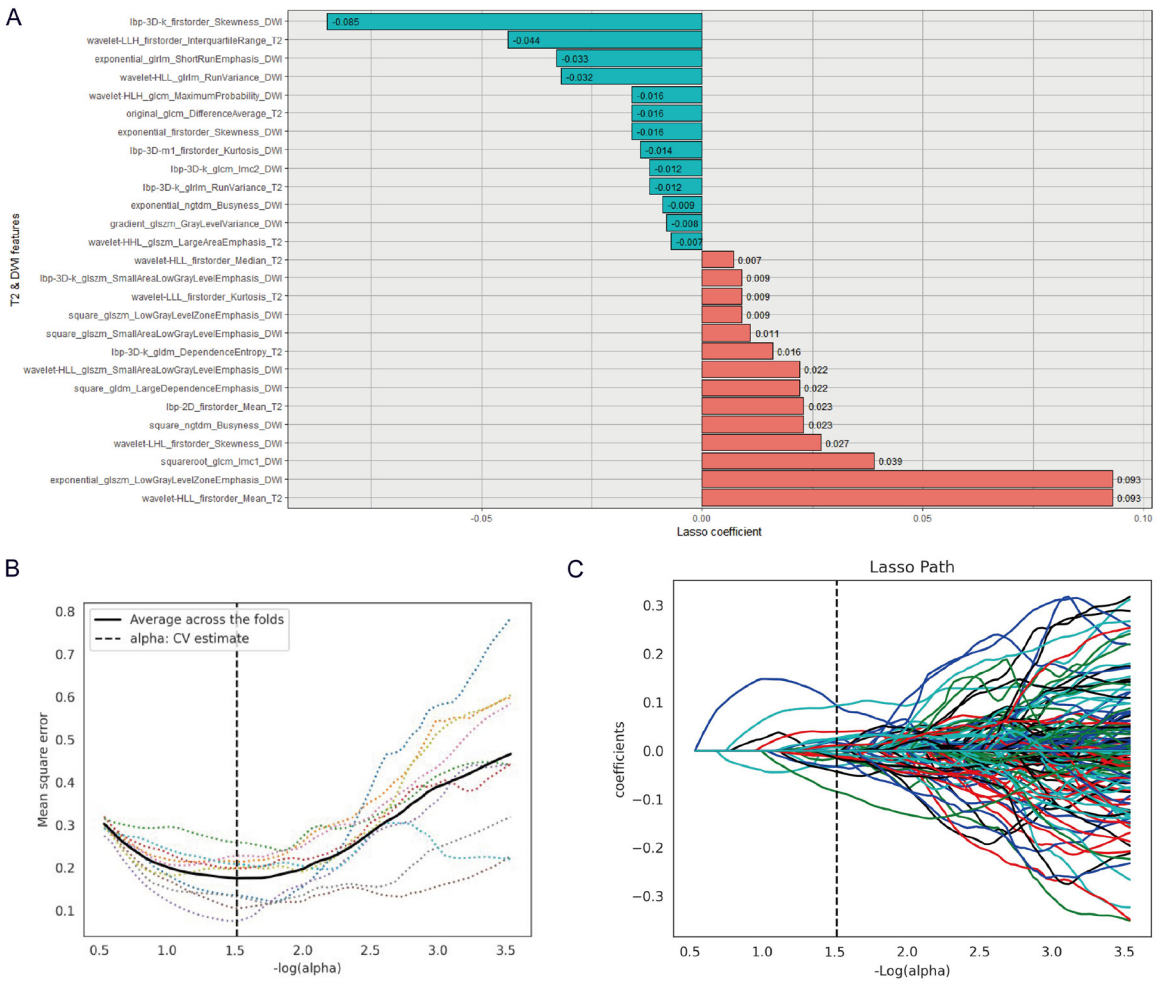

Figure 2. TD (T2WI + DWI) combined sequence dimensionality reduction and radiomics model construction. A: Image features and correlation coefficients of the minimum absolute contraction and selection operator (Lasso); B: Best tuning parameter (λ) selected by 10-fold cross-validation in the LASSO regression model; C: λ used to obtain the radiomics features of the non-zero series (each colored line represents the change in its coefficient).

to their significant association with the diagnosis in univariate analysis, as shown in **Table 1**, and their universal availability in clinical practice), were incorporated as continuous and categorical predictors, respectively, into a multivariate logistic regression analysis to build the combined clinical-radiomics model. The performance of this combined model was then evaluated and compared against the radiomics-only model in both the training and validation cohorts.

Figure 6 presents the combined diagnostic model along with a nomogram for predicting FNH probability. The model's discriminatory power was assessed via ROC analysis. In the training cohort (**Figure 7**; **Table 3**), the model achieved an AUC of 0.992 (95% CI: 0.979-1.000), accompanied by a sensitivity of 97.1% and specificity of 98.5%. The 10-fold cross-validation cohort yielded an AUC of 0.958 (95% CI: 0.904-0.993), with 86.7% sensitivity, 86.7% specificity, a positive predictive value (PPV) of 86.7%, a negative predictive value (NPV) of 86.0%, and an F1-score of 0.867. Importantly, this combined model (TDC + clinical data) outperformed the TDC-only model (**Figure 5D**), confirming its enhanced diagnostic accuracy.

External validation

The model's FNH/HCC classification ability was tested via an independent validation set derived from hospital records and follow-up data (January 2022-December 2023), with 49 patients with HCC (27 male, 22 female) and 42 patients with FNH (14 male, 28 female) (**Table**

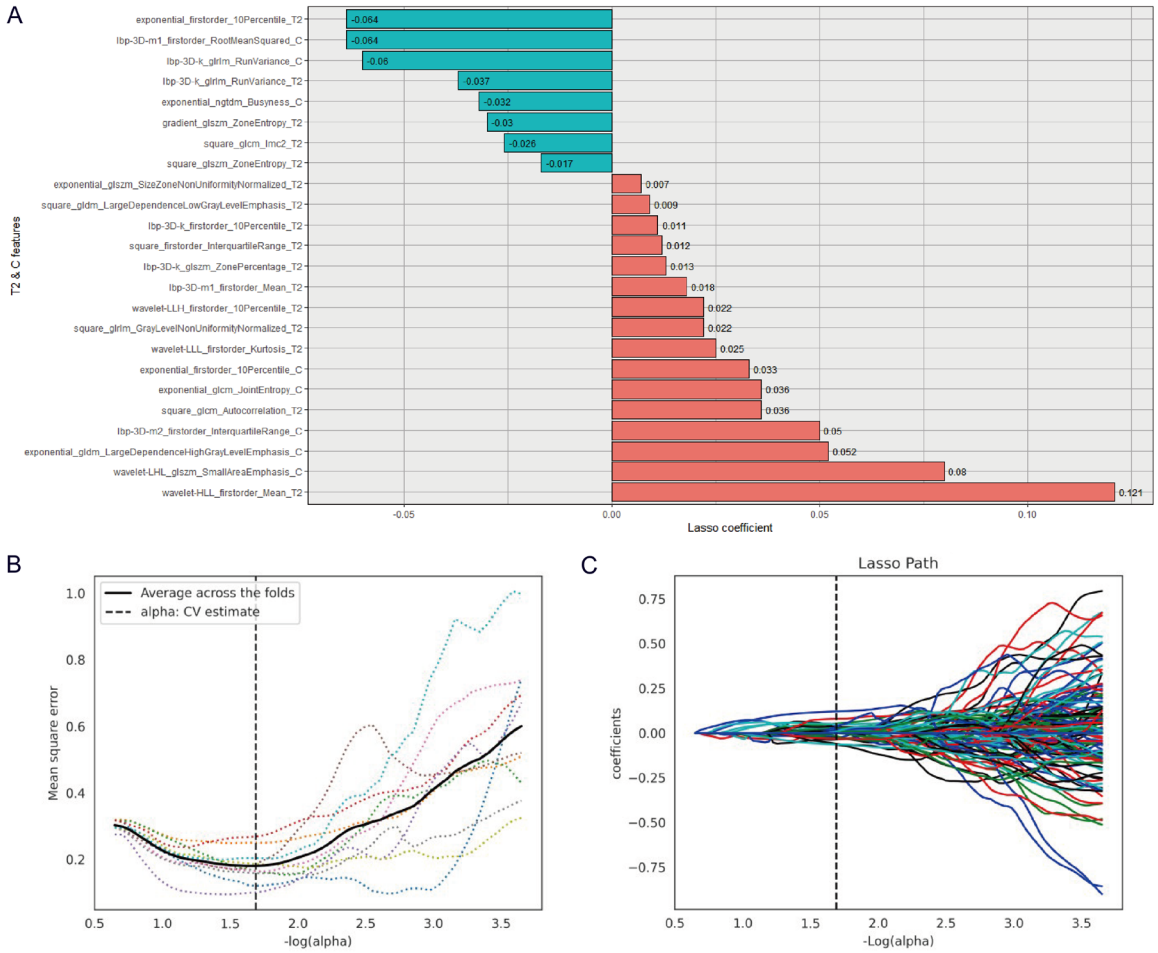

Figure 3. TC (T2WI + CEI) combined sequence dimensionality reduction and radiomics model construction. A: Image features and correlation coefficients of the minimum absolute contraction and selection operator (Lasso); B: Best tuning parameter (λ) selected by 10-fold cross-validation in the LASSO regression model; C: λ used to obtain the radiomics features of the non-zero series (each colored line represents the change of its coefficient).

4). As shown in **Figure 8**, Nomogram-assisted risk evaluation yielded an AUC of 0.903 (95% CI: 0.834-0.972), with a sensitivity of 88.9%, a specificity of 87.2%, a PPV of 89.8%, a NPV of 88.1%, and an F1-score of 0.893, underscoring the nomogram's excellent differentiation of FNH from HCC.

Representative cases

To illustrate the clinical context and the decision-making process facilitated by our model, we present the following case. Although we are unable to share the original MRI due to patient privacy regulations, we provide a detailed radiological description that aligns with the input features of our model.

Case 1: a 38-year-old woman with histologically proven FNH

T2-weighted Imaging: The lesion appeared slightly hyperintense compared to the surrounding liver parenchyma.

Diffusion-weighted Imaging (DWI): The lesion showed restricted diffusion, presenting as hyperintense.

Contrast-enhanced Imaging (CEI): On the arterial phase, the lesion demonstrated strong and homogeneous enhancement. During the portal venous and delayed phases, the enhancement became isointense to the liver, with a characteristic persistently enhancing central scar becoming visible in the delayed phase.

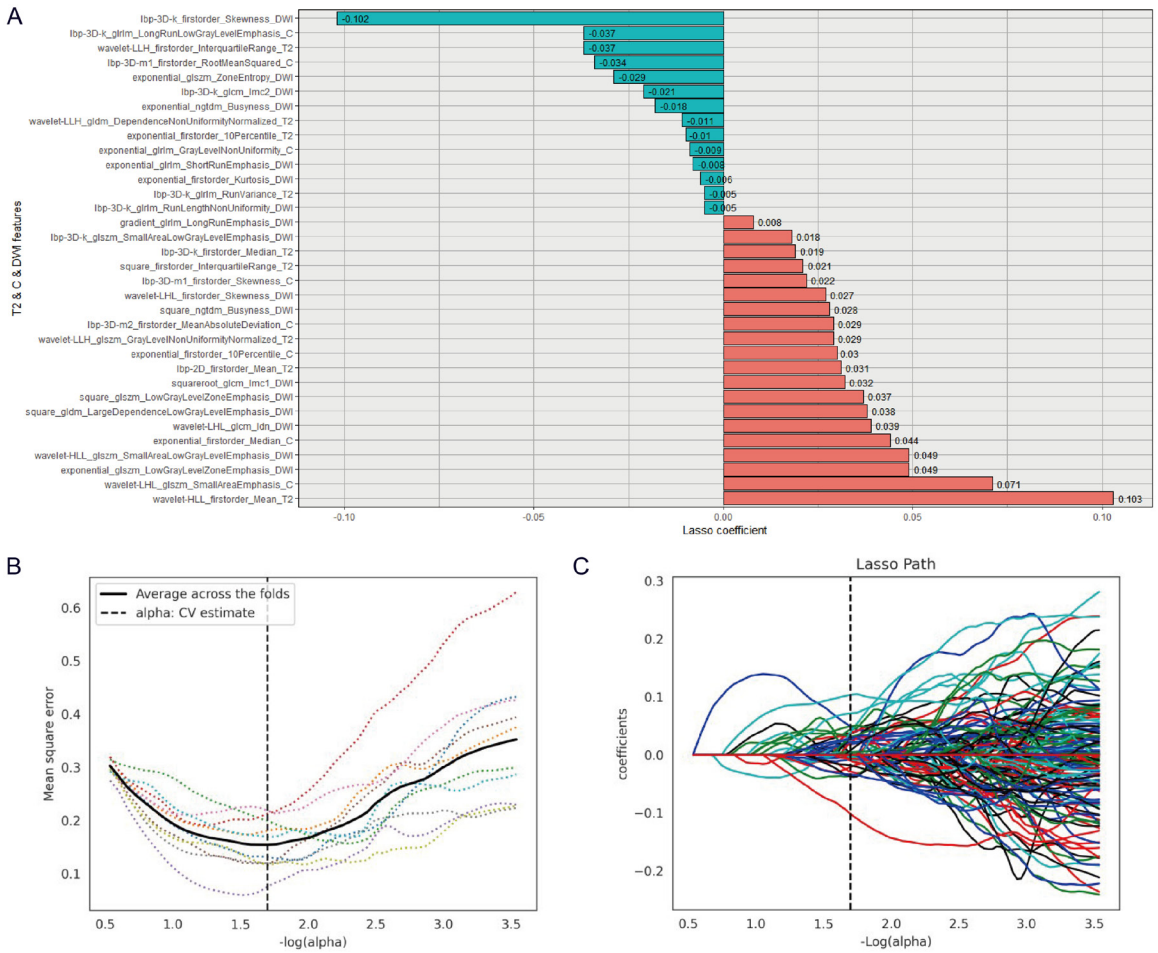

Figure 4. TDC (T2WI + DWI + CEI) combined sequence dimensionality reduction and radiomics model construction. A: Image features and correlation coefficients of the minimum absolute contraction and selection operator (Lasso); B: Best tuning parameter (λ) selected by 10-fold cross-validation in the LASSO regression model; C: λ used to obtain the radiomics features of the non-zero series (each colored line represents the change in its coefficient).

The radiomics analysis of this lesion's ROI on the aforementioned sequences yielded a high Rad-score, and the combined clinical-radiomics model correctly predicted a high probability of FNH, which was consistent with the pathological diagnosis. This case exemplifies a typical presentation where the model can reinforce a confident diagnosis.

Discussion

In this study, we developed and validated a multivariate radiomics model based on T2WI, DWI, and contrast-enhanced MRI for distinguishing between HCC and FNH in patients with hepatic nodular lesions. The accurate differentiation between these two entities is crucial for determining appropriate treatment strategies and avoiding unnecessary interventions.

It is crucial to accurately distinguish between HCC and FNH for determining the appropriate course of treatment and preventing unwarranted therapeutic procedures. Conventional MRI has been widely used for the differential diagnosis of FNH and HCC [11], but these examinations have limitations. An example is that for diagnosing HCC, the sensitivity and specificity of conventional MRI are relatively low, especially for small HCCs. Conventional MRI also might not be able to distinguish between HCC and other benign liver lesions, such as FNH [3]. Even with these limitations, conventional MRI is an important diagnostic tool for liver lesions.

Capitalizing on advancements in radiomics, several approaches have been suggested to distinguish between benign and malignant liver tumors. A previous study developed a CT-based

Table 2. Differential diagnostic performance of the LR model based on different sequences of MRI for FNH agent HCC

Model	Data Group	AUC	Accuracy	Sensitivity	Specificity
T2WI	Training cohort	0.942 (0.909-0.968)	0.846	0.841	0.851
	Validation cohort	0.814 (0.707-0.895)	0.750	0.867	0.633
TD	Training cohort	0.967 (0.941-0.988)	0.897	0.899	0.896
	Validation cohort	0.891 (0.802-0.957)	0.833	0.867	0.800
TC	Training cohort	0.963 (0.940-0.984)	0.890	0.899	0.881
	Validation cohort	0.901 (0.820-0.970)	0.800	0.900	0.700
TDC	Training cohort	0.989 (0.978-0.997)	0.956	0.957	0.955
	Validation cohort	0.944 (0.886-0.987)	0.867	0.900	0.833

LR, logistic regression; MRI, Magnetic Resonance Imaging; FNH, focal nodular hyperplasia; HCC, hepatocellular carcinoma; AUC, Area under the curve; T2WI, T2-weighted imaging; T2WI and DWI (Diffusion Weighted Imaging), referred to as TD, T2WI and CEI (Contrast Enhanced Imaging), referred to as TC, T2WI, DWI and C, referred to as TDC.

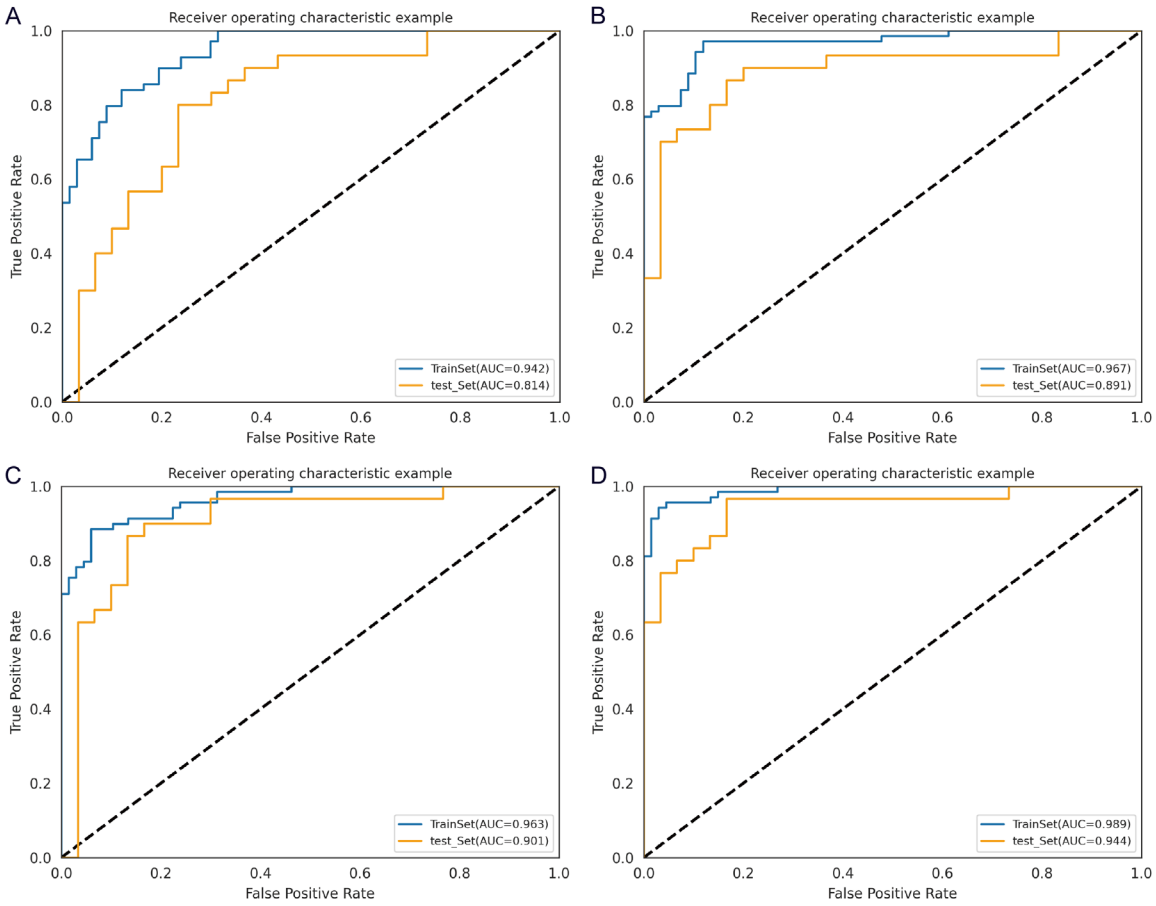

Figure 5. Receiver operating characteristic curves of the training group and test group in the classifier models. A: T2WI single-sequence group model; B: TD (T2WI + DWI) combined sequence; C: TC (T2WI + CEI) combined sequence; D: TDC (T2WI + DWI + CEI) combined sequence. AUC: Area under the curve. The specific data of the ROC curve are presented in **Table 2**.

radiomics model that showed good preoperative discrimination between HCC and FNH, with AUCs of 0.964 and 0.865 in the training set and the validation set, respectively [5]. The MRI-based model achieved good differentiation of HCC from FNH in both the training and validation datasets (0.956 and 0.941, respectively) [12]. Also, there was a highly accurate

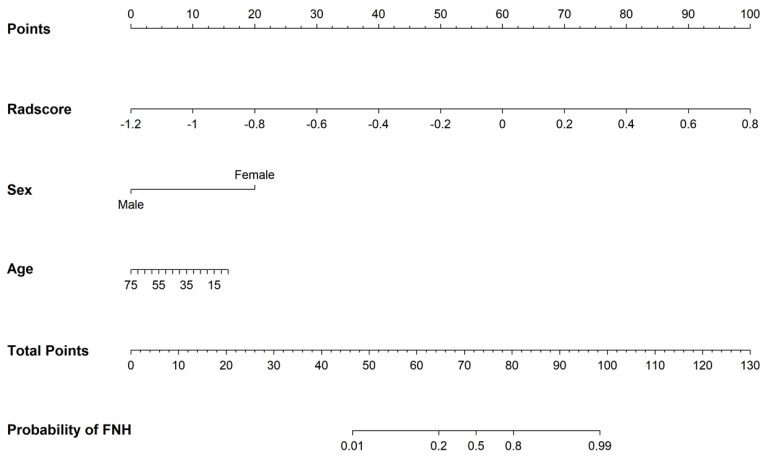

Figure 6. Nomogram of the combined diagnostic model combining TDC radiomics with clinical parameters. FNH, focal nodular hyperplasia.

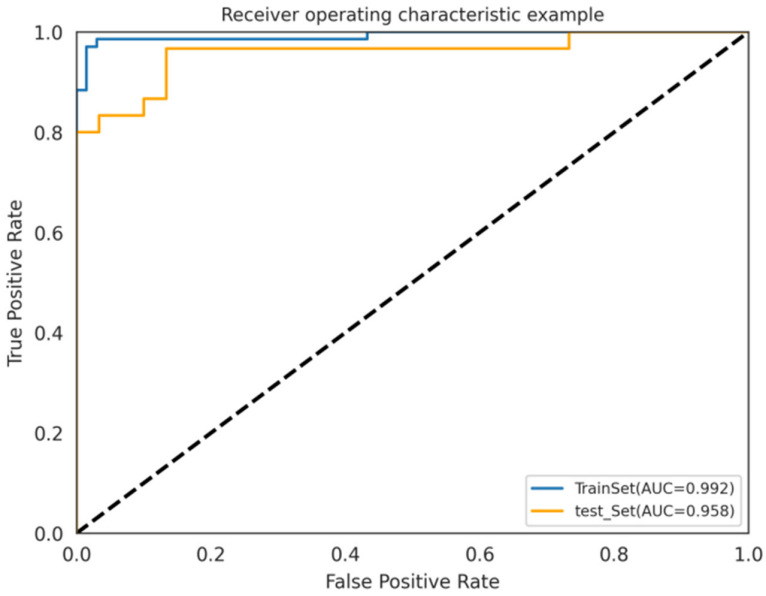

Figure 7. ROC curve analysis evaluating the diagnostic performance of the combined model for FNH identification. ROC, receiver operating characteristic; FNH, focal nodular hyperplasia.

diagnostic model using Gd-EOB-DTPA-enhanced MRI to preoperatively distinguish HCC from FNH in challenging cases where both showed iso- or hyperintensity in the hepatobiliary phase [13]. It is worth noting that while Gd-EOB-DTPA-enhanced MRI provides valuable functional information, our model utilizes more widely available contrast agents, potentially offering a more accessible alternative without compromising diagnostic performance. Another study developed a CT-based radiomics model to identify hepatic lesions (including HCC and FNH), and the classifiers had good diagnostic performance, with AUC values > 0.900 in the training and validation groups [14]. A radiomics model based on T2W MRI, designed to discriminate between malignant and benign solid liver lesions, exhibited AUC values ranging from 0.74-0.86 across various validation datasets [15].

Our study advances this field through two key methodological innovations. First, we established a comprehensive multi-sequence fusion approach, systematically integrating features from T2WI, DWI, and CEI. This TDC model capitalizes on the complementary strengths of each sequence. While the TC model (T2 + CEI) contained only one more feature than the T2-only model, this does not necessarily indicate redundant information in CEI. Rather, it reflects our stringent feature selection process which prioritized the most discriminative features. The retained CEI feature likely captures unique vascular and perfusion characteristics crucial for differentiation, complementing the architectural information from T2WI. T2WI emphasizes the contrast between tissues on the basis of their T2 relaxation times, which is useful for visualizing differences in water content and detecting abnormalities, such as edema or inflammation [16, 17]. DWI measures the random motion of water molecules within tissues, which can indicate cellular structure changes [18]. It is particularly helpful for identifying areas of restricted diffusion, which are often observed in acute stroke or certain tumor types. CEI refers to the use of contrast agents to enhance the visibility of blood vessels and tissue perfusion in scans [19, 20]. The superior performance of our TDC model, achieving AUCs of 0.992 and 0.958 in training and 10-fold

Table 3. Diagnostic performance of the nomogram for FNH prediction

	AUC	ACC	Sensitivity	Specificity	95% CI
Training set	0.992	0.978	0.971	0.985	0.979-1.000
Validation set	0.958	0.867	0.867	0.867	0.904-0.993

FNH, focal nodular hyperplasia; AUC, area under the curve; CI, confidence interval.

Table 4. Clinical characteristics of patients in the external validation cohort

	HCC (n=49)	FNH (n=42)	$t/\chi^2/U$	p
Age (years)	55.12±11.00	38.00±11.83	7.150	0.000
Sex			4.329	0.038
Male	27 (55.10)	14 (33.33)		
Female	22 (44.90)	28 (66.67)		
Rad-score (IQR)	0.2 (-0.1, 0.5)	0.4 (0.225, 0.7)	747.500	0.024

FNH, focal nodular hyperplasia; HCC, hepatocellular carcinoma; IQR, interquartile range.

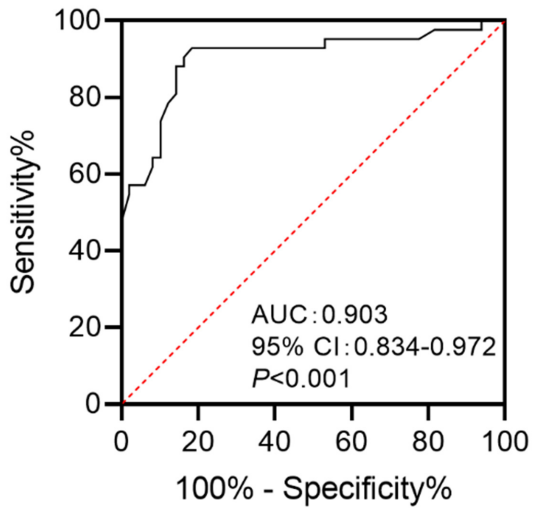

Figure 8. External validation (ROC analysis). ROC, receiver operating characteristic; AUC, area under the curve.

cross-validation cohorts, respectively, demonstrates the synergistic value of this multi-parametric approach.

To address feature biological significance, we analyzed key predictors in our model. The radiomics features selected in our final model demonstrate plausible biological correlates with the underlying pathophysiology of HCC and FNH. For example, the texture feature primarily from T2WI likely captures architectural heterogeneity, with higher values in HCC corresponding to its disorganized structure with variable cellularity, necrosis, and fibrosis versus FNH's organized architecture with uniform hepatocytes and characteristic central scar. Similarly, features from DWI may reflect cellular density differences between hypercellular HCC and normal hepatocyte arrangement in FNH with relatively free water diffusion. CEI-derived features potentially capture the distinct vascular patterns, and HCC typically demonstrates neoangiogenesis with chaotic arterial enhancement and washout, while FNH exhibits homogeneous arterial enhancement with sustained enhancement in later phases due to its organized vascular architecture. These interpretations, while requiring further validation, provide plausible biological explanations for our model's discriminative capability based on established pathophysiology [21, 22]. Then, we developed a clinically integrated framework combining the TDC radiomics signature with essential clinical parameters. The combined model showed improved diagnostic accuracy, and its robust performance in external validation (AUC: 0.903) underscores clinical potential. By synthesizing multi-parametric MRI data with clinical variables, our approach can deliver a more comprehensive diagnostic profile than single-modality radiomics, advancing the pursuit of accurate, non-invasive diagnosis in hepatology.

To facilitate the clinical application of our combined diagnostic model, we have constructed a practical nomogram (**Figure 6**). This nomogram serves as a bridge between the complex algorithm and clinical practice, allowing radiologists to manually calculate a personalized probability of FNH for each patient by summing points

assigned to their specific radiomics signature, age, and sex. While this study represents the development and validation phase, our immediate future work includes the development of a user-friendly web-based calculator or mobile application. This tool would automate the scoring process, integrating seamlessly into the radiologist's workflow by inputting the Rad-score and clinical parameters to instantly output the FNH probability. Furthermore, we plan to conduct clinical utility studies to evaluate radiologists' diagnostic confidence and accuracy with and without the assistance of this model, which is a critical step towards its widespread acceptance and adoption.

In this study, we elected to employ the logistic regression (LR) classifier due to its high interpretability, computational efficiency, and lower risk of overfitting, especially given that our final model was parsimonious with only 3 features. This aligns with our goal of developing a clinically transparent and deployable tool. While we acknowledge that testing a broader range of classifiers (e.g., Random Forest, Support Vector Machines) and dimensionality reduction methods (e.g., Principal Component Analysis) represents a valuable future direction, the strong performance of our carefully tuned LR model (AUC: 0.903 on external validation) demonstrates its effectiveness as a robust solution for the task at hand. Future work will include a comprehensive benchmark of various algorithms to further optimize performance.

This study has several limitations. Its retrospective, single-center design with a small sample size may limit generalizability and introduce selection bias. Furthermore, the cohort contained an insufficient number of small lesions (≤ 2 cm) to perform a meaningful subgroup analysis. While our model relied on portal venous phase features, its performance with hepatobiliary-specific contrast agents (e.g., gadoxetate disodium) remains unverified and requires future investigation. The research scope was also restricted to differentiating HCC from FNH, and the clinical relevance of the imaging features remains unclear. Methodologically, the limited range of algorithms used and the absence of decision curve and calibration analyses affect the evaluation of model robustness and clinical utility. Despite these limitations, the model demonstrated balanced sensitivity and specificity (86.7% each) in external validation, indicating its potential to address key diagnostic challenges.

Conclusion

The study findings support the use of the multimodal MRI model based on T2WI, DWI, and the C-sequence can be used to distinguish between HCC and FNH. The shorthand diagnostic model constructed by use of the LR classifier model gave the best performance, and it can assist in the differential diagnosis of HCC and FNH, thus improving the accuracy of the differential diagnosis and realizing individualized precision medicine. Our study results have important implications for accurate diagnosis and proper treatment of HCC and FNH. Further research is needed to validate these findings and to explore the potential of this approach in other clinical settings.

Acknowledgements

We would like to acknowledge the Health Youth Research Project of Fujian Province (2020QNA076), for providing the monetary support necessary to conduct this research. The Health Youth Research Project of Fujian Province has played a pivotal role in making this project possible.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Zhixian Wu, Department of Medical Imaging, Zhangzhou Affiliated Hospital of Fujian Medical University, Zhangzhou 363000, Fujian, China. Tel: +86-0596-2082372; E-mail: wzxzzsyy@163.com

References

[1] Kohn CG, Singh P, Korytowsky B, Caranfa JT, Miller JD, Sill BE, Marshall AC and Parikh ND. Humanistic and economic burden of hepatocellular carcinoma: systematic literature review. Am J Manag Care 2019; 25: SP61-SP73.

[2] Singh SP, Madke T and Chand P. Global epidemiology of hepatocellular carcinoma. J Clin Exp Hepatol 2025; 15: 102446.

[3] Aly A, Ronnebaum S, Patel D, Doleh Y and Benavente F. Epidemiologic, humanistic and

economic burden of hepatocellular carcinoma in the USA: a systematic literature review. Hepat Oncol 2020; 7: HEP27.
[4] Hwang SY, Danpanichkul P, Agopian V, Mehta N, Parikh ND, Abou-Alfa GK, Singal AG and Yang JD. Hepatocellular carcinoma: updates on epidemiology, surveillance, diagnosis and treatment. Clin Mol Hepatol 2025; 31: S228-S254.
[5] Nie P, Yang G, Guo J, Chen J, Li X, Ji Q, Wu J, Cui J and Xu W. A Ct-based radiomics nomogram for differentiation of focal nodular hyperplasia from hepatocellular carcinoma in the non-cirrhotic liver. Cancer Imaging 2020; 20: 20.
[6] Azizaddini S and Mani N. Liver Imaging. Statpearls. Treasure Island (FL) ineligible companies. Disclosure: Nisha Mani declares no relevant financial relationships with ineligible companies: 2025.
[7] Zhu M, Li H, Wang C, Yang B, Wang X, Hou F, Yang S, Wang Y, Guo X and Qi X. Focal nodular hyperplasia mimicking hepatocellular adenoma and carcinoma in two cases. Drug Discov Ther 2021; 15: 112-117.
[8] Murakami T and Tsurusaki M. Hypervascular benign and malignant liver tumors that require differentiation from hepatocellular carcinoma: key points of imaging diagnosis. Liver Cancer 2014; 3: 85-96.
[9] Maino C, Vernuccio F, Cannella R, Franco PN, Giannini V, Dezio M, Pisani AR, Blandino AA, Faletti R, De Bernardi E, Ippolito D, Gatti M and Inchingolo R. Radiomics and liver: where we are and where we are headed? Eur J Radiol 2024; 171: 111297.
[10] Mao HY, Hu JC, Zhang T, Fan YF, Wang XM, Hu CH and Yu YX. Hepatocellular Carcinoma (Hcc) and Focal Nodular Hyperplasia (Fnh) showing Iso- or hyperintensity in the hepatobiliary phase: differentiation using Gd-Eob-Dtpa enhanced mri radiomics and deep learning features. BMC Med Imaging 2025; 25: 397.
[11] Rao PN. Nodule in liver: investigations, differential diagnosis and follow-up. J Clin Exp Hepatol 2014; 4 Suppl 3: S57-62.
[12] Ding Z, Lin K, Fu J, Huang Q, Fang G, Tang Y, You W, Lin Z, Lin Z, Pan X and Zeng Y. An Mr-based radiomics model for differentiation between hepatocellular carcinoma and focal nodular hyperplasia in non-cirrhotic liver. World J Surg Oncol 2021; 19: 181.
[13] Mao HY, Shen BQ, Zhang JY, Zhang T, Cai W, Fan YF, Wang XM, Yu YX and Hu CH. Gd-Eob-Dtpa enhanced mri nomogram model to differentiate hepatocellular carcinoma and focal nodular hyperplasia both showing Iso- or hyperintensity in the hepatobiliary phase. BMC Med Imaging 2024; 24: 211.
[14] Zhao X, Liang P, Yong L, Jia Y and Gao J. Radiomics study for differentiating focal hepatic lesions based on unenhanced Ct images. Front Oncol 2022; 12: 650797.
[15] Starmans MPA, Miclea RL, Vilgrain V, Ronot M, Purcell Y, Verbeek J, Niessen WJ, Ijzermans JNM, de Man RA, Doukas M, Klein S and Thomeer MG. Automated assessment of T2-weighted Mri to differentiate malignant and benign primary solid liver lesions in noncirrhotic livers using radiomics. Acad Radiol 2024; 31: 870-879.
[16] Nougaret S, Horta M, Sala E, Lakhman Y, Thomassin-Naggara I, Kido A, Masselli G, Bharwani N, Sadowski E, Ertmer A, Otero-Garcia M, Kubik-Huch RA, Cunha TM, Rockall A and Forstner R. Endometrial cancer Mri staging: updated guidelines of the European society of urogenital radiology. Eur Radiol 2019; 29: 792-805.
[17] Veraart J, Novikov DS and Fieremans E. Te dependent diffusion imaging (Teddi) distinguishes between compartmental T(2) relaxation times. Neuroimage 2018; 182: 360-369.
[18] Ni P, Lin Y, Zhong Q, Chen Z, Sandrasegaran K and Lin C. Technical advancements and protocol optimization of diffusion-weighted imaging (Dwi) in liver. Abdom Radiol (NY) 2016; 41: 189-202.
[19] Reimer P, Schneider G and Schima W. Hepatobiliary contrast agents for contrast-enhanced Mri of the liver: properties, clinical development and applications. Eur Radiol 2004; 14: 559-578.
[20] Dietrich CF, Albrecht T, Becher H, Harvey CJ, Jenssen C, Lim AK, Möller K and Greis C. History of contrast enhanced ultrasound (Ceus). Med Ultrason 2024; 26: 405-416.
[21] Moga TV, Lupusoru R, Danila M, Ghiuchici AM, Popescu A, Miutescu B, Ratiu I, Burciu C, Bizerea-Moga T, Voron A, Sporea I and Sirli R. Challenges in diagnosing focal liver lesions using contrast-enhanced ultrasound. Diagnostics (Basel) 2024; 15: 46.
[22] Dai H, Xiao Y, Fu C, Grimm R, von Busch H, Stieltjes B, Choi MH, Xu Z, Chabin G, Yang C and Zeng M. Deep learning-based approach for identifying and measuring focal liver lesions on contrast-enhanced Mri. J Magn Reson Imaging 2025; 61: 111-120.