



Evaluating artificial intelligence for a focal nodular hyperplasia diagnosis using magnetic resonance imaging: preliminary findings

- Mecit Kantarcı^{1,2}
 Volkan Kızılgöz¹
 Ramazan Terzi³
 Ahmet Enes Kılıç³
 Halime Kabalcı⁴
 Önder Durmaz¹
 Nil Tokgöz⁵
 Mustafa Harman⁶
 Ayşegül Sağır Kahraman⁷
 Ali Avanz⁸
 Sonay Aydın¹
 Gülsüm Özlem Elpek⁹
 Merve Yazol⁵
 Bülent Aydın⁸

¹Erzincan Binali Yıldırım University Faculty of Medicine, Department of Radiology, Erzincan, Türkiye

²Atatürk University Faculty of Medicine, Department of Radiology, Erzurum, Türkiye

³Digital Transformation Office, Presidency of the Republic of Türkiye, Ankara, Türkiye

⁴Erzincan Binali Yıldırım University Faculty of Medicine, Department of Anatomy, Erzincan, Türkiye

⁵Gazi University Faculty of Medicine, Department of Radiology, Ankara, Türkiye

⁶Ege University Faculty of Medicine, Department of Radiology, İzmir, Türkiye

⁷İnönü University Faculty of Medicine, Department of Radiology, Malatya, Türkiye

⁸Akdeniz University Faculty of Medicine, Department of General Surgery, Antalya, Türkiye

⁹Akdeniz University Faculty of Medicine, Department of Pathology, Antalya, Türkiye

Corresponding author: Volkan Kızılgöz

E-mail: volkankizilgoz@gmail.com

Received 25 October 2024; revision requested 04 December 2024; last revision received 09 February 2025; accepted 2 March 2025.



Epub: 26.03.2025

Publication date:

DOI: 10.4274/dir.2025.243095

PURPOSE

This study aimed to evaluate the effectiveness of artificial intelligence (AI) in diagnosing focal nodular hyperplasia (FNH) of the liver using magnetic resonance imaging (MRI) and compare its performance with that of radiologists.

METHODS

In the first phase of the study, the MRIs of 60 patients (30 patients with FNH and 30 patients with no lesions or lesions other than FNH) were processed using a segmentation program and introduced to an AI model. After the learning process, the MRIs of 42 different patients that the AI model had no experience with were introduced to the system. In addition, a radiology resident and a radiology specialist evaluated patients with the same MR sequences. The sensitivity and specificity values were obtained from all three reviews.

RESULTS

The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the AI model were found to be 0.769, 0.966, 0.909, and 0.903, respectively. The sensitivity and specificity values were higher than those of the radiology resident and lower than those of the radiology specialist. The results of the specialist versus the AI model revealed a good agreement level, with a kappa (κ) value of 0.777.

CONCLUSION

For the diagnosis of FNH, the sensitivity, specificity, PPV, and NPV of the AI device were higher than those of the radiology resident and lower than those of the radiology specialist. With additional studies focused on different specific lesions of the liver, AI models are expected to be able to diagnose each liver lesion with high accuracy in the future.

CLINICAL SIGNIFICANCE

AI is studied to provide assisted or automated interpretation of radiological images with an accurate and reproducible imaging diagnosis.

KEYWORDS

Artificial intelligence, deep learning, liver lesion, focal nodular hyperplasia, magnetic resonance imaging

Focal nodular hyperplasia (FNH) is the second most common benign tumor of the liver after hemangioma. The prevalence of FNH was found to be 0.4% to 3% in autopsy series.¹ FNH is believed to result from arterial malformations, and 60%–80% of cases are asymptomatic and are discovered incidentally.^{2,3} The imaging characteristics of FNH correspond well with histological properties and are observed as a solitary well-circumscribed lobulated mass in a cross-sectional imaging study (Figure 1).⁴ Magnetic resonance imaging (MRI) has a higher sensitivity than ultrasound and computed tomography (CT) imaging and a specificity of almost 100%.⁵ In MRI, a typical FNH is a solitary, well-defined, unencapsulated lesion with central scar formation.⁶ Approximately 35%–70% of FNH lesions do not have

these imaging features; they might have a pseudo capsule mimicking a true capsule, show washout-like hepatocellular carcinoma (HCC), or have no scar formation.^{7,8} The hepatobiliary phase (HBP) of MRI provides important data for the diagnosis of FNH, and 73%–90% of these lesions are observed with iso-intensity or hyperintensity in the HBP.⁹ Even though HCC and hepatic adenoma are usually hypointense in the HBP, these lesions may have upregulated hepatocyte-specific membrane transport proteins and, thus, may be observed as an iso- or hyperintense lesion in HBP images.⁴

Artificial intelligence (AI) is becoming a widespread method to interpret radiological images for research purposes, even in daily practice. It is expected to provide assisted or automated interpretation of radiological images with an accurate and reproducible imaging diagnosis. After obtaining images of the patients, AI may quickly interpret them and make critical diagnostic decisions for numerous patients. This may provide a quick and accurate diagnosis of many lesions located in different organs or systems in the future. Thus, all scientific studies targeted at developing AI for use as a diagnostic assistant can be considered a contribution to this topic. As a diagnostic tool, AI has been used in the detection and characterization of diffuse diseases or focal lesions of the liver and pancreas in recent studies. It has been applied to different imaging techniques, including ultrasound, CT, and MRI.¹⁰

The aim of the present study was to determine the effectiveness of AI in detecting the presence of FNH lesions of the liver and compare this diagnostic capacity of AI with that of radiologists. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated, considering the radiological and pathological results of the patients as the gold standard.

Main points

- The targeted long-term result is automated interpretation with an accurate diagnosis using artificial intelligence (AI) models for liver lesions; this study is part of the AI education program focusing on a specific liver lesion.
- A new scoring system is established to train the AI model to distinguish focal nodular hyperplasia (FNH) from other liver lesions.
- The AI model used in this research achieved sensitivity and specificity values higher than those of a radiology resident and lower than those of a radiology specialist for the diagnosis of FNH.



Figure 1. Schematic of classic focal nodular hyperplasia, a solitary well-circumscribed lobulated mass with central scar tissue. This illustration has been created by the Adobe Photoshop program (Adobe Inc., 2021. Adobe Photoshop, <https://www.adobe.com/products/photoshop.html>) based on the figures provided by the Complete Anatomy program (3D4 Medical, 2021; Complete Anatomy; retrieved from <https://3d4medical.com>)

Methods

Patients and the study workflow

This study was approved by the Ethics Committee of Erzincan Binali Yıldırım University (clinical trial number: 2023-13/6, date: 22.06.2023) and the requirement for informed consent was waived by the ethics committee due to the retrospective nature of the study. The study population constituted patients who had undergone MRI, and abdominal MRIs of 30 patients were used in the initial phase. In the first phase, the MRIs of patients (n = 30) who had been histologically diagnosed with FNH were introduced to the AI system using a segmentation program. In addition, the abdominal MRIs of 30 patients with no liver lesions were segmented using the same program. A scoring system was used to diagnose FNH. Then, 42 patients with various lesions, including FNH (n = 13), HCC (n = 5), low-grade dysplastic nodules (n = 1), hepatic adenoma (n = 3), biliary hamartoma (n = 1), primary hepatic neuroendocrine tumor (n = 1), colon cancer metastasis (n = 2), breast cancer metastasis (n = 2), stomach cancer metastasis (n = 1), pancreatic cancer metastasis (n = 1), hydatid cyst (n = 1), complex cyst (n = 1), biliary cystadenoma (n = 2), hemangioma (n = 4), simple cyst (n = 3), and a normal liver were reviewed by AI and two radiologists (a specialist with 18 years of experience and a radiology resident with 2.5 years of experience) independently in randomized order (Figure 2). Following the AI interpretations, sensitivity, specificity, the PPV, and the NPV were calculated. Then, the accuracy of the results from the AI model

and the two radiologists were compared. The radiological diagnosis (stable lesions with typical imaging features in follow-up examinations or typical imaging findings with primary tumor) and the histological results (obtained by biopsy procedures) were taken as the gold standard to reveal the sensitivity and specificity values.

Diagnosing focal nodular hyperplasia

A standardized method was used to simplify the interpretation, and for reproducibility and repeatability regarding the FNH diagnosis, only the axial plane images, including T1-weighted, T2-weighted, and T1-weighted enhanced (arterial, portal, and venous phase) images, and HBP images were evaluated. Typically, FNH is hypointense or isointense on T1-weighted images and hyperintense or isointense on T2-weighted images, showing intense contrast medium enhancement in the arterial phase and retaining contrast in the portal and venous phases.¹¹ The central FNH scar is best seen on MRI. The scar is monitored as hypointense on a pre-contrast T1-weighted sequence, substantially hyperintense on T2-weighted images, and becomes hyperintense on HBP images because of the accumulation of the contrast medium in the fibrous tissue. Most FNH lesions are iso- or hyperintense on HBP images (Figure 3).⁷

Magnetic resonance imaging protocol and selected sequences

All the MRIs were acquired using a 1.5T MRI scanner (Magnetom Era, Siemens, Erlangen, Germany) with a standard abdom-

inal coil. The axial sequences, including the T1-weighted and T2-weighted images, as well as the contrast-enhanced phases, were evaluated. The MRIs were segmented using dedicated software, ensuring the precise identification of focal lesions. All the contrast-enhanced T1-weighted images were obtained using gadoxetate disodium (Primovist®) through intravenous injection at a dosage of 0.1 mmol/kg (maximum dose, 20 mL) and a rate of 2 mL/s, followed by saline flush (50 mL at the rate of 2 mL/s). Postcontrast images were analyzed, including the late arterial phase (15–20s postinjection), portal venous phase (60–70s postinjection), delayed phase (3–5 min postinjection), and HBP (20 min postinjection). The axial plane T1- and T2-weighted, arterial, portal, venous, and HBP enhanced T1-weighted MRIs were introduced to the AI model and interpreted by the two radiologists through the liver lesion diagnostic process in relation to FNH. The following technical parameters were applied to both the enhanced and non-enhanced series: T2 weighted: time of repetition (TR): 1,200 ms, time of echo (TE): 95 ms, number of excitations (NEX): 1, slice thickness: 6 mm; T1 weighted: TR: 6.94 ms, TE: 2.39 ms, NEX: 1, slice thickness: 3.3 mm.

Segmentation

The segmentation process was performed by an anatomist and a radiologist (with 3 and 27 years of experience, respectively) using the same monitor as that used for segmentation of the abdominal MRIs and at the same time. The radiologist decided on the presence and locations of the liver lesions for each patient. The anatomist had learned about image maps, anatomical details, and liver lesions from an experienced radiologist. The anatomist consulted with the radiologist at every step. Each patient MR examination was also checked at the end of the segmentation session by the experienced radiologist for every segmented anatomical part or liver lesion. Segmentation of the axial images was performed with 3D Slicer software (v5.3.0, <http://www.slicer.org>) manually. The liver borders, FNH lesions (if any), lesions other than FNH, the main branches of the portal veins, and the hepatic veins were segmented on six sequences in the axial plane, as described before in the diagnosing FNH section. Only the focal lesions were tagged, and fibrosis or other diffuse parenchymal signal alterations were not segmented. The main portal veins and main hepatic veins were segmented in each patient. All FNH lesions in the liver were segmented if the patient had

more than one lesion. The FNH lesions were segmented based on the lesion borders, and scar formation was also segmented in typical FNH lesions. The FNH lesion, scar formation of the FNH lesion, liver, main portal vein, main hepatic veins, and lesions other than FNH were tagged with different colors before being introduced to the AI model. Based

on the axial slices, three-dimensional (3D) reconstruction images were also obtained by the segmentation program (Figure 4). After the AI training session, a radiology resident (with 2.5 years of experience), a radiology specialist (with 18 years of experience), and the AI model evaluated the random dataset that included the FNH and other lesions.

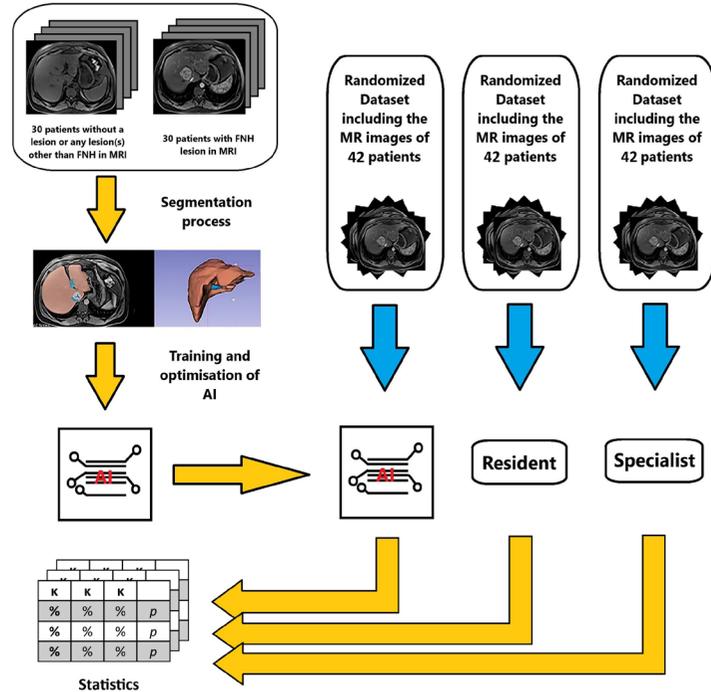


Figure 2. Study workflow. After the segmentation process and training the AI model, a randomized dataset was evaluated by AI, a radiology resident, and a radiology specialist independently. AI, artificial intelligence; FNH, focal nodular hyperplasia; MRI, magnetic resonance imaging.

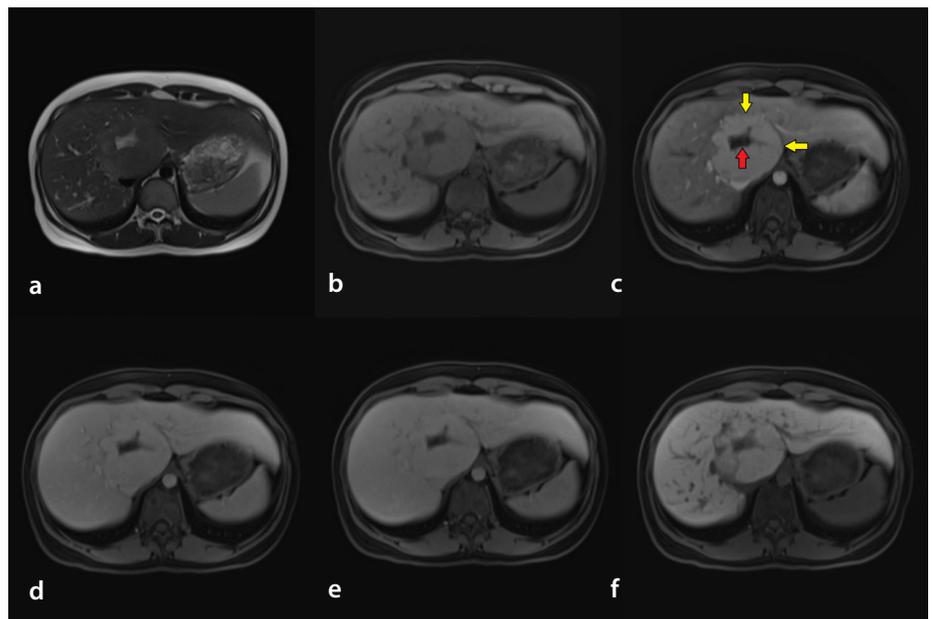


Figure 3. Classic focal nodular hyperplasia with radial scar tissue. Axial plane magnetic resonance images with T2-weighted (a), T1-weighted pre-contrast (b), arterial phase contrast-enhanced T1-weighted (c), portal phase contrast-enhanced T1-weighted (d), venous phase contrast-enhanced T1-weighted (e), and hepatobiliary phase (f) images (yellow arrows indicate the lesion in an enhanced T1-weighted axial plane image, and the red arrow shows the typical central scar of the lesion).

Artificial intelligence protocol and data preprocessing

The workflow developed for FNH detection with AI from MRIs is presented in Figure 5. The workflow consists of two stages: segmentation and the FNH detection process. The process from dataset preparation to FNH detection with AI is explained in detail in this section.

The MRI data provided were converted from nrrd format to .nii.gz format, and a data standard was created. For the 3D modeling, the data were converted to Medical Segmentation Decathlon format.¹² To produce a more generalizable result, the five-fold cross-validation method was applied instead of random split for the algorithms.

Deep learning architectures

Organs such as the liver, veins, and gallbladder can be detected in MRIs thanks to deep learning architecture such as object recognition, semantic segmentation, and instance segmentation. In this study, a decision-making process was used to focus on the intensity of FNH so that it could be detected by AI. Thus, the use of segmentation algorithms was deemed more appropriate. Moreover, it was decided to use 3D segmentation algorithms instead of two-dimensional (2D) segmentation algorithms to access the temporal information between MRI slices. In this study, the nnU-Net deep learning algorithm, a deep learning-based semantic segmentation model developed with both 2D and 3D U-Net configurations, was used.¹³ We chose to use this algorithm for this study because it can automatically configure appropriate preprocessing, network architecture, training parameters, and post-processing processes according to the data in the medical imaging.

Artificial intelligence-training and testing

The 3D nnU-Net model training was performed in three categories—the liver, vein, and FNH—using model configurations prepared based on the data of 60 patients, 30 with FNH and 30 without. Model training was conducted with the five-fold cross-validation method. Thirty nnU-Net models were trained in five-fold form over six phases: T1 weighted, T2 weighted, arterial, portal, and HBP. The hyperparameters used for model training are shown in Table 1. Optimal model selection was made according to the highest average validation Dice score. The most successful 3D nnU-Net model selected was tested on 30 test patients.

Artificial intelligence-evaluation metrics

The metrics used to evaluate the segmentation model performance provide a key tool for measuring the sensitivity, accuracy, and overall effectiveness of the developed model. In this study, the Dice score (Sørensen–Dice coefficient) metric was used. The Dice score is a metric that measures how well the region predicted by the model overlaps with the actual labeled region. This metric, used to evaluate the similarity between two clusters, is calculated with the following formula:¹⁴

$$\text{Dice Score} = 2x \frac{|\text{Prediction} \cap \text{Ground truth}|}{|\text{Prediction}| + |\text{Ground truth}|}$$

In this formula, prediction represents the segmentation region predicted by the model, and ground truth represents the ground truth region. The Dice score has a value between 0 and 1, with a value closer to 1 indicating greater overlap. A high Dice score indicates that the model performs segmentation correctly, whereas a low value indicates that the model's predictions are incompatible with the actual data.

Artificial intelligence-registration

Six phases were used to decide whether a patient had an FNH liver lesion. When six deep learning models are developed for six phases and used separately, a situation occurs if the lesion can be found in one phase

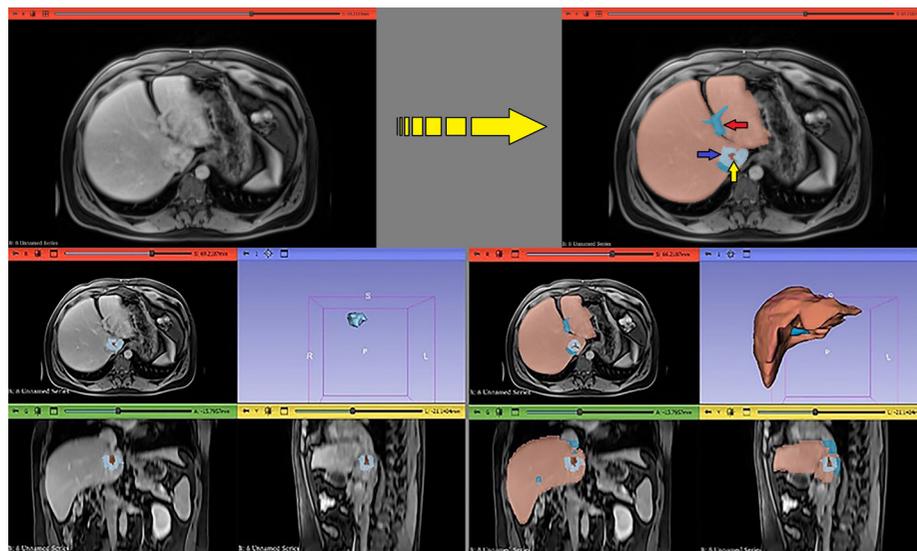


Figure 4. Magnetic resonance images after segmentation of the anatomical structures and focal nodular hyperplasia (FNH). After segmentation of all of the axial slices, either FNH or anatomical formation of the liver has been coded and tagged as a space-occupying structure by the segmentation program. In the right upper corner, the left main portal vein (red arrow), FNH (blue arrow), and central FNH scar formation (yellow arrow) are tagged with different colors after segmentation.

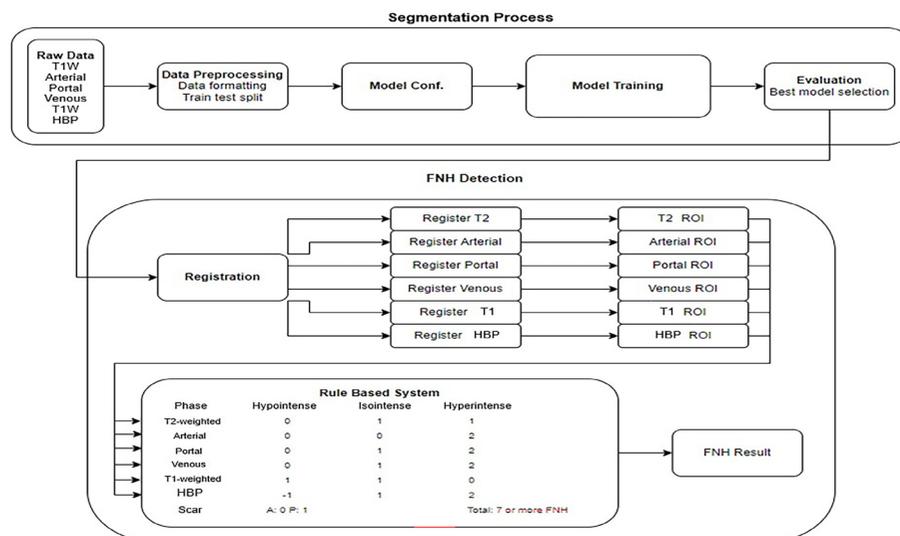


Figure 5. Artificial intelligence workflow. T1W, T1-weighted; HBP, hepatobiliary phase; ROI, region-of-interest; FNH, focal nodular hyperplasia.

and not in another. In this case, deficiencies in the evaluation exist in terms of AI. Therefore, a 3D registration process was used in this study. The 3D registration process is used to align the position and orientation of images in the 3D space. This process is generally performed to obtain geometric harmony between a reference (base) and a moving image. In this study, a reference phase and the other five phases were registered separately. Since the most successful deep learning model was developed on the arterial phase, the reference phase was determined as the arterial phase. The registration process shown in Figure 6a has been produced automatically in 3D Slicer (Figure 6). The elastix registration method was used in this process.¹⁵

Region-of-interest extraction

When performing phase checks for FNH, specialist physicians make decisions by focusing on the surroundings of the FNH region. However, deep learning models segment all the relevant locations for the liver, vein, and FNH. To solve this, the region-of-interest (ROI) extraction post-processing method was used. For ROI extraction, as shown in Figure 6b, the region segmented by the deep learning model as FNH was increased by 30%, and only the liver and vein segmentations around the FNH label were obtained. Since the arterial phase is the reference phase, the regions predicted by the deep learning model in the arterial phase were mapped onto the other five registration phases, and ROI extraction was completed for the six phases.

Rule-based system

The average pixel intensity was measured using the signal intensity of the liver, vein, and FNH segmentations within the ROI regions, six phases apart, and extracted. To make an intensity decision, the liver average pixel intensity of each phase was compared with the FNH average pixel intensity. To determine the lesion as hypo-, iso-, or hyperintense relative to the liver tissue, the surrounding liver parenchyma (the adjacent 30% of the area of the lesion) was considered (Figure 7). A comparison table for the intensity decision and the scoring system for each phase is shown in Table 1, and the decision regarding the presence of FNH is made according to the MR intensity obtained. To enable AI to determine the presence of FNH, a strict pattern must be followed. A lack of information in the literature and the absence of any widely used or accepted rule to enable AI to decide accurately, compelled the researchers

to find a new pathway. Therefore, a new scoring system was established based on the MR signal features of the lesion. The images of patients in the training session (the images of 30 patients with at least one FNH lesion) were used for the preliminary testing to optimize the scoring system. According to this scoring system, for the unenhanced series, 1 point was allocated to iso- or hyperintensity in T2-weighted images and 1 point to hypo-

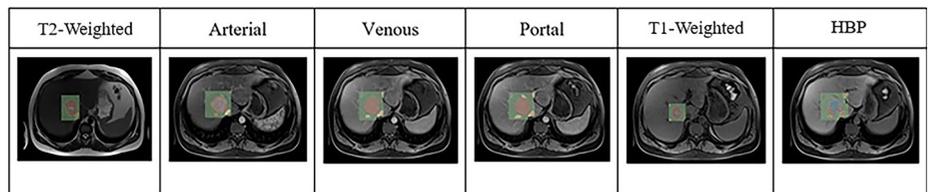
or iso-intensity in T1-weighted images. For the dynamic contrast-enhanced series, the signal intensity was identified relative to the surrounding liver parenchyma. According to this rule, hyperintensity in the arterial phase and iso- or hyperintensity in the portal, venous, and HBPs were all allocated 1 point. A lesion with scar tissue was considered as 2 points. In total, 7 or more points were considered to be FNH according to the morpholog-

Table 1. Model hyperparameters, intensity decision, and the rule-based system of this study

Model hyperparameters			
Hyperparameters	Values		
Epoch number	1,000		
Batch size	2		
Learning rate	Poly learning rate scheduler (initial learning rate: 0.01)		
Optimizer	Stochastic gradient descent		
Momentum	0.99		
Weight decay	3e-05		
Loss function	Robust cross entropy loss Memory efficient soft Dice loss		
Intensity decision (SI unit)			
Value (liver SI-FNH SI)	Decision of intensity		
Value < -10	Hypointense		
-10 < value < 10	Isointense		
Value > 10	Hyperintense		
Rule-based system			
T2 weighted	0	1	1
T1 weighted	1	1	0
Arterial phase	0	0	2
Portal phase	0	1	2
Venous phase	0	1	2
Hepatobiliary phase	-1	1	2
Scar	Absence of the scar scored as 0 points, and presence of the scar scored as 1 point		
Result: 7 or more points were considered focal nodular hyperplasia			
SI, signal intensity; FNH, focal nodular hyperplasia.			



a-Registration example of the MRI phases



b-ROI extraction of the MRI phases

Figure 6. Registration process of each magnetic resonance sequence and the region-of-interest extraction. HBP, hepatobiliary phase; ROI, region-of-interest; MRI, magnetic resonance image.

ical appearance in the MRIs. If one or more lesions in the liver were consistent with FNH on the MRIs, the patient was accepted as FNH positive by each reviewer.

Statistical analysis

All statistical analyses were calculated using IBM SPSS statistics v22.0 for Windows. Sensitivity, specificity, the PPV, the NPV, and accuracy were calculated using the chi-square test, with the radiological and histological results considered as the gold standard. The area under curve (AUC) values were calculated and presented as 95% confidence intervals (CIs). Cohen's kappa analysis was performed to reveal the agreement levels between reviewers, and Koo et al.'s¹⁶ classification method was used to represent the agreement levels. According to these agreement levels, values less than 0.50 indicated poor agreement, values between 0.50 and 0.75 showed moderate agreement, values between 0.75 and 0.90 revealed good agreement, and values greater than 0.90 demonstrated excellent agreement.¹⁶ A *P* value < 0.05 was considered to represent a statistically significant difference.

Results

The training of the 30 nnU-Net models was conducted in the form of five-fold cross validation for six phases: T2 weighted, pre-contrast (T1 weighted), arterial, portal, venous, and HBP. The nnU-Net deep learning algorithm automatically adjusts model configurations according to the data. Figure 2 shows the Dice score and mean validation Dice score for the liver, vein, and FNH classes over five-fold means \pm standard separately for each phase.

Among the 30 nnU-Net models trained, the results were given singularly for each MR sequence. The arterial phase images had the highest performance in terms of the average validation Dice score (0.7998), and this sequence was chosen as the best model in the category of both FNH class success and high average validation Dice score (Table 2).

In this study, 5 of the 13 FNH lesions were typical FNH lesions with scar formation. The list of patients with the histological and interpretation results of each reviewer are presented in Table 3. Two patients had more than one FNH lesion in the liver, as seen in the dataset presented in Table 3. The dimensions of the FNH lesions measured on the axial plane are presented in Table 4. The mean was 2.78 ± 1.84 for the FNH dimensions.

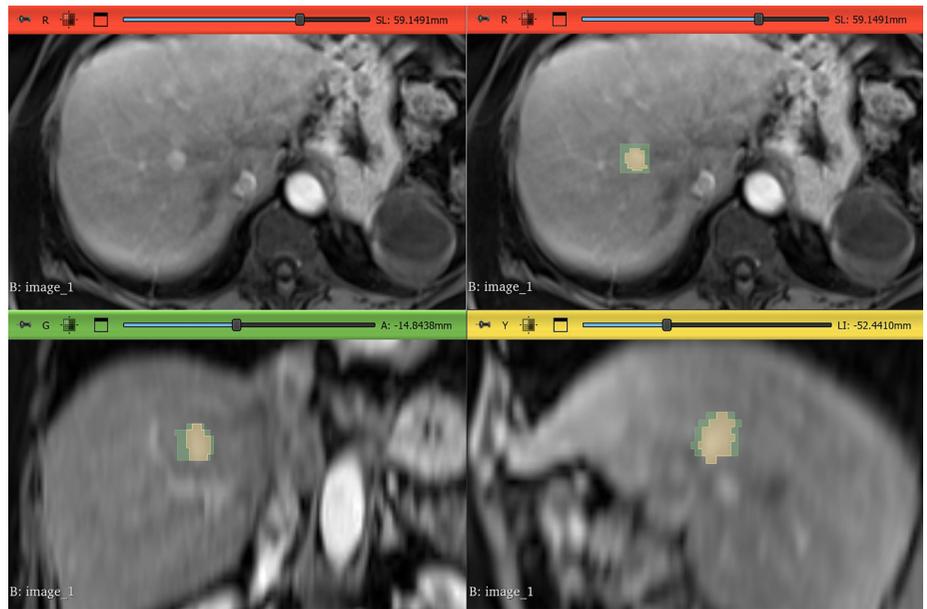


Figure 7. Surrounding liver parenchyma (the adjacent 30% of the area of the lesion) was considered to determine the lesion as hypo-, iso-, or hyperintense relative to the liver tissue.

Table 2. Fold mean results

Sequence	Liver	Vein	FNH	EMA Dice	Mean val Dice
T2 weighted	0.949 \pm 0.002	0.586 \pm 0.05	0.254 \pm 0.253	0.571 \pm 0.08	0.512 \pm 0.03
T1 weighted	0.954 \pm 0.01	0.734 \pm 0.02	0.136 \pm 0.160	0.589 \pm 0.04	0.560 \pm 0.01
Arterial phase	0.955 \pm 0.01	0.680 \pm 0.02	0.733 \pm 0.246	0.780 \pm 0.09	0.712 \pm 0.05
Venous phase	0.961 \pm 0.01	0.746 \pm 0.04	0.645 \pm 0.221	0.767 \pm 0.08	0.671 \pm 0.07
Portal phase	0.946 \pm 0.04	0.752 \pm 0.03	0.608 \pm 0.246	0.759 \pm 0.09	0.665 \pm 0.07
HBP	0.948 \pm 0.02	0.629 \pm 0.10	0.436 \pm 0.262	0.651 \pm 0.09	0.633 \pm 0.06

FNH, focal nodular hyperplasia; EMA, exponential moving average; Val, validated; HBP, hepatobiliary phase.

Table 3. Reviewer results for each patient (“+” indicates the presence of FNH, and “-” represents the absence of FNH in the liver)

	Reference diagnosis	RD	H	Resident	AI	Specialist
1	Normal liver	+	-	-	-	-
2	Hemangioma	+	-	-	-	-
3	Hydatid cyst	+	-	-	-	-
4	PHNT	-	+	-	-	-
5	Simple cyst	+	-	-	-	-
6	Hemangioma	+	-	+	-	-
7	Breast cancer metastasis	-	+	-	-	-
8	FNH	+	+	+	+	+
9	FNH	+	-	+	-	-
10	Stomach cancer metastasis	-	+	-	-	-
11	Hemangioma	+	-	+	-	-
12	Hepatocellular carcinoma	-	+	-	+	-
13	Biliary hamartoma	-	+	-	-	-
14	Complex cyst	-	+	-	-	-
15	Low-grade dysplastic nodule	-	+	-	-	-
16	FNH	+	+	+	+	+
17	Breast cancer metastasis	+	-	-	-	-
18	Colon cancer metastasis	-	+	-	-	-

The liver interpretations on the MRIs according to each reviewer were compared with the histopathological results regarding the presence of FNH. The sensitivity, specificity, PPV, and NPV obtained from the radiology resident, the radiology specialist, and the AI model are presented in Table 5. According to these results, the diagnostic parameters of the AI model were better than those of the resident and lower than those of the specialist.

The results of the radiology resident and the AI model showed poor agreement ($\kappa = 0.374$), and the results of the radiology resident and the radiology specialist indicated good agreement ($\kappa = 0.602$). The results of the radiology specialist and the AI model revealed good agreement ($\kappa = 0.777$) (Table 6). The AUC values with 95% CI were 0.794 (0.630–0.959) for the radiology resident, 0.833 (0.682–0.983) for the AI model, and 0.944 (0.851–1.000) for the radiology specialist (Table 7). The accuracy values were 0.833, 0.905, and 0.952 for the radiology resident, AI model, and radiology specialist, respectively.

Discussion

The AI model used in this study had 76.9% sensitivity, 96.6% specificity, a 90.9% PPV, and a 90.3% NPV for the diagnosis of FNH of the liver. The AI results were better than those of the radiology resident and lower than those of the radiology specialist. Additionally, the AI results indicated a good level of agreement with the specialist.

FNH is a conservatively managed lesion for most patients, and surgery is not required in the management of this condition. Only patients with pedunculated, exophytic, or expanding lesions are considered for surgery.¹⁷ Hepatic adenoma, however, is treated by surgery because of its well-known complications, including spontaneous hemorrhage and malignant transformation.¹⁸ HCC is another lesion that occurs in the differential diagnosis of FNH, and HCC may also occur in a non-cirrhotic liver.¹⁹ The spectrum of patients that AI will evaluate should comprise all these lesions as well as cirrhotic livers that may have diagnostic challenges. Another important discussion point is distinguishing hepatic adenomas from FNH lesions. This may not be easy to accomplish using MRIs. Most adenomas (reported to be between 75% and 90%) are hypointense in the HBP, whereas FNH is iso- or hyperintense compared with the surrounding liver parenchyma, and these different lesion properties make the diagnosis easier in daily practice.

Table 3. Continued

	Reference diagnosis	RD	H	Resident	AI	Specialist
19	Colon cancer metastasis	+		–	–	–
20	Hepatic adenoma		+	+	+	+
21	Hemangioma	+		–	+	–
22	Simple cyst	+		–	–	–
23	Pancreas cancer metastasis	+		–	–	–
24	Hepatocellular carcinoma		+	–	–	–
25	FNH	+		–	+	+
26	FNH	+		+	–	+
27	Simple cyst	+		–	–	–
28	FNH	+		+	+	+
29	Biliary cystadenoma		+	–	–	–
30	Biliary cystadenoma		+	–	–	–
31	FNH	+		+	+	+
32	FNH	+		–	+	+
33	FNH		+	–	+	+
34	FNH	+		+	–	+
35	FNH		+	–	+	+
36	FNH	+		+	+	+
37	FNH	+		+	+	+
38	Hepatocellular carcinoma		+	–	–	–
39	Hepatic adenoma	+		–	–	–
40	Hepatocellular carcinoma		+	–	–	–
41	Hepatocellular carcinoma		+	–	–	–
42	Hepatic adenoma		+	–	–	–

AI, artificial intelligence; RD, radiological diagnosis (stable in follow-up examinations or typical imaging findings with primary tumor), H, histologically confirmed lesions; PHNT, primary hepatic neuroendocrine tumor; FNH, focal nodular hyperplasia.

Table 4. Dimensions of the FNH lesions in axial plane

	Patient number*	Lesion 1 (TR × AP)	Lesion 2 (TR × AP)	Lesion 3 (TR × AP)
1	Patient number 8	3.62 × 2.97 cm		
2	Patient number 9	1.56 × 1.39 cm		
3	Patient number 16	1.47 × 1.42 cm		
4	Patient number 25	3.39 × 2.79 cm		
5	Patient number 26	0.80 × 0.89 cm		
6	Patient number 28	1.98 × 1.59 cm	2.96 × 2.54 cm	6.65 × 7.50 cm
7	Patient number 31	7.17 × 5.49 cm		
8	Patient number 32	5.41 × 5.15 cm		
9	Patient number 33	3.51 × 3.30 cm		
10	Patient number 34	1.28 × 1.49 cm	1.94 × 2.01 cm	
11	Patient number 35	2.24 × 1.66 cm		
12	Patient number 36	0.91 × 0.86 cm		
13	Patient number 37	1.75 × 2.33 cm		

*Represents the patient numbers in the study, also shown in Table 3; TR, maximum transvers diameter of the lesion; AP, maximum anteroposterior diameter of the lesion. TR, time of repetition; AP, anterioposterior.

The uptake and excretion of hepatocyte-specific contrast agents into the biliary system is facilitated by hepatocyte-specific membrane transport proteins, which are not present in other cells. HCC and hepatic adenoma are

usually hypointense in the HBP; however, these lesions may have upregulated hepatocyte-specific membrane transport proteins, which make them appear as iso- or hyperintense lesions in HBP images. Approximately

25% of inflammatory hepatic adenomas and 40–80% of beta-catenin-activated hepatic adenomas are reported to appear as iso- or hyperintense on HBP images, and this overlap makes diagnosis challenging. Moreover, beta-catenin-activated hepatic adenomas have the highest risk for malignant transformation (40%).⁴

This study is a step forward in using AI to diagnose one of the most common hepatic nodular lesions, FNH. Not only typical but also atypical nodular hyperplasia lesions, which have been histologically confirmed, were evaluated through AI as a reviewer. The AI model provided a relatively high sensitivity value along with 96.6% specificity in diagnosing FNH in the liver with one or more nodular lesions, including non-FNH lesions.

In the literature, researchers have included many parameters in their studies. These include the contrast curve, gray-level histogram, and gray-level co-occurrence matrix texture properties, as well as risk factors, such as the presence of steatosis, known primary tumors, or cirrhosis, and MR sequences such as dynamic contrast-enhanced T1-weighted with T2-weighted images, for the classification of focal liver lesions.²⁰ In the present study, a simplified approach, using only certain MR sequences that were unaware of other risk factors or medical conditions, was used to understand the diagnostic success of AI. The nnU-Net deep learning algorithm was chosen for this study, which is considered to be highly impactful for object identification with successful segmentation capabilities. This algorithm was designed to optimize 2D or 3D image segmentation tasks and is usable for any given input geometrical type. This deep learning modality optimally segments organs using CT images based on the use of differences of densities.²¹ In this research, signal intensity was used as the indicator of FNH lesions using the same algorithm and both 2D and 3D U-Net configurations.

In this study, arterial phase images had the highest performance for the average validation Dice score. Dice scores were important for determining the anomaly and starting to implement further calculations to reveal the lesion characterization regarding the presence of FNH. The ground truth-based border drawn by the radiologists was analyzed along with a prediction based on the border of the model. The Dice score represents the overall segmentation performance and indicates the success of the segmentation through the prediction ability of

the model. The Dice score ranges from 0 (no overlap compared with the segmented borders of the radiologist) to 1 (perfect overlap compared with the segmented borders of

the radiologist). This method was used in the literature for similar purposes, such as the segmentation of HCC in the liver.²²

Table 5. Sensitivity, specificity, and positive and negative predictive values of the reviewers with 95% confidence intervals

Results for the resident				
Radiology resident		Radiology resident		Total
		Positive	Negative	
Pathology result	Positive	9	4	13
	Negative	3	26	29
Total		12	30	42
Sensitivity: 0.692 (0.388 ± 0.896)				<i>P</i> < 0.001
Specificity: 0.897 (0.715 ± 0.972)				
Positive predictive value: 0.750 (0.428 ± 0.933)				
Negative predictive value: 0.867 (0.683 ± 0.956)				
Results for artificial intelligence				
AI		AI		Total
		Positive	Negative	
Pathology result	Positive	10	3	13
	Negative	1	28	29
Total		11	31	42
Sensitivity: 0.769 (0.459 ± 0.938)				<i>P</i> < 0.001
Specificity: 0.966 (0.803 ± 0.998)				
Positive predictive value: 0.909 (0.571 ± 0.995)				
Negative predictive value: 0.903 (0.730 ± 0.974)				
Results for the radiology specialist				
Radiology specialist		Radiology specialist		Total
		Positive	Negative	
Pathology result	Positive	12	1	13
	Negative	1	28	29
Total		13	29	42
Sensitivity: 0.923 (0.620 ± 0.995)				<i>P</i> < 0.001
Specificity: 0.966 (0.803 ± 0.998)				
Positive predictive value: 0.923 (0.620 ± 0.995)				
Negative predictive value: 0.966 (0.803 ± 0.998)				

AI, artificial intelligence.

Table 6. Kappa values for the comparison of each reviewer

	Value	Asymptotic standard error ^a	Approximate T ^b	Approximate significance
Resident vs. AI	0.374	0.155	2.428	0.015
Specialist vs. AI	0.777	0.106	5.037	<0.001
Resident vs. specialist	0.602	0.135	3.905	<0.001

^a, not assuming a null hypothesis; ^b, using the asymptotic standard error assuming a null hypothesis. AI, artificial intelligence.

Table 7. Area under the curve for each reviewer's results

Test result variable (s)	Area	Standard error	<i>P</i> values	Asymptotic 95% confidence interval	
				Lower bound	Upper bound
Resident	0.794	0.084	0.003	0.630	0.959
AI	0.833	0.077	0.001	0.682	0.983
Specialist	0.944	0.048	0.000	0.851	1.000

AI, artificial intelligence.

According to the results presented in Table 2, each imaging phase demonstrates distinct characteristics in segmenting different parts of the liver. For instance, the HBP exhibited the highest overall liver segmentation performance, with a Dice score of 0.948 ± 0.02 , whereas the portal phase achieved the best vein segmentation performance, with a Dice score of 0.752 ± 0.03 . Similarly, the most effective segmentation for FNH was observed in the arterial phase, yielding a Dice score of 0.733 ± 0.246 . In summary, based on the mean Dice score across phases, the arterial phase proved most effective in segmenting the three liver components, achieving a Dice score of 0.712 ± 0.05 . This might be a result of the increased signal difference between the lesion and the surrounding parenchyma. The ability of the model to distinguish the lesion borders was considered superior for the arterial phase images. Consequently, the arterial phase was chosen as the foundational model. Specifically, the first-fold model of the arterial phase, which achieved a mean Dice score of 0.7998, was selected as the base model for FNH detection and segmentation. Subsequently, the scoring system was considered for analyzing the lesion, particularly for the diagnosis of FNH.

There are attempts in the literature to use automatized AI models to diagnose focal liver lesions. In Goehler et al.'s²³ study, the researchers tried to detect liver metastases and evaluate changes in tumor size on consecutive MR examinations. A convolutional neural network (CNN) and Kuhn–Munkres algorithm were used for 64 patients with neuroendocrine tumors with two consecutive liver MR examinations using gadoxetic acid. The results of this study indicated that this evaluation system was 91% concordant with the radiologists' decision, and the sensitivity and specificity were 0.85 and 0.92, respectively. In addition, the model was capable of assessing the interval change in tumor burden between two MRI examinations.²³ A computer-assisted diagnosis system, the liver artificial neural network (ANN), was analyzed by Zhang et al.²⁴ regarding its feasibility for identifying focal liver lesions. Using an ANN technique, this system classified the liver lesions into five categories. Their investigation used 320 MRIs (from 80 patients); however, the system was human assisted, and a radiologist had to delineate an ROI for the lesion. The five hepatic categories for the lesions in their study were cavernous hemangioma, HCC, hepatic cyst, dysplasia in cirrhosis, and metastasis. This liver ANN system was developed to assist the radiolo-

gists, giving a second opinion with a training accuracy of 100% and a testing accuracy of 93%.²⁴ For the diagnosis of focal liver lesions, Hamm et al.²⁵ performed a study using multi-phasic MRIs, and 92% sensitivity, 98% specificity, and 92% accuracy were achieved with their CNN. In the same study, the model displayed a sensitivity of 90% for the diagnosis of HCC, whereas the radiologist achieved 70%.²⁵ Jansen et al.²⁰ utilized a system of automatic classification to classify focal liver lesions using MRIs and the risk factors for a more accurate diagnosis. They achieved an overall accuracy for focal liver lesions of 0.77. The sensitivity and specificity values for hepatic hemangioma were 84% and 82%, respectively, for hepatic cyst, 93% and 93%, for hepatic adenoma, 80% and 78%, for HCC, 73% and 56%, and for metastasis, 62% and 77%.²⁰ Zhen et al.²⁶ analyzed the efficiency of a deep learning-based tool based on the fact that dynamic contrast-enhanced MRI provides the most precise diagnosis of hepatic tumors. In their analysis, enhanced and unenhanced MRIs, along with relevant patient clinical information, were used. The results indicated that the deep learning-based system differentiated malignant from benign focal liver lesions well using only unenhanced images (AUC: 0.946; 95% CI: 0.914–0.979 vs. AUC: 0.951; 95% CI: 0.919–0.982, $P = 0.664$). Moreover, the performance of the deep learning-based system was improved when combining unenhanced images with clinical data to classify malignancies as metastatic tumors (AUC = 0.998; 95% CI: 0.989–1.000), HCC (AUC: 0.998; 95% CI: 0.989–1.000), HCC (AUC: 0.985; 95% CI: 0.960–1.000), and other primary malignancies (AUC: 0.963; 95% CI: 0.896–1.000). Compared with the pathological examination, the agreement was 91.9%, and the sensitivity and specificity values for almost every liver lesion category achieved the same accuracy as those of experienced radiologists.²⁶ A study by Stollmayer et al.²⁷ used deep learning with 2D and 3D networks to diagnose FNH, HCC, and liver metastases on hepatocyte-specific contrast-enhanced MRIs. In total, 216 MRIs from 69 patients were analyzed. Overall, the 2D model performed better, with AUCs of 0.990, 0.966, and 0.960, respectively, for the investigated liver lesions.²⁷ Wang et al.'s²⁸ CNN-based model differentiated various focal liver lesions as either benign or malignant. Then, detailed classification was performed depending on tumor types. A total of 557 images were separated into a training and a testing set, and the AUCs for the classifications were 0.969 and 0.919, respectively. Seven focal liver lesions—liver cyst, cavernous hemangioma,

hepatic abscess, FNH, HCC, intrahepatic cholangiocarcinoma, and hepatic metastasis—were investigated in their research using seven MR sequences (T2 weighted, diffusion weighted, apparent diffusion coefficient, T1 weighted, late arterial phase, portal venous phase, and delayed phase), and the accuracy for performing the seven-way classification was 79.6%.²⁸ The present study focused on a specific lesion, and it is difficult to compare the results with those of other studies in which some of the lesions were grouped and some were focused on distinguishing benign and malignant lesions.

To the best of our knowledge, this is the first study focused solely on the presence of FNH of the liver, and the results indicate promising results for the future. The sensitivity and specificity of the AI model were 76.9% and 96.6%, respectively, which were lower than those of the radiology specialist. The AUC value was 0.833 (95% CI: 0.682–0.983) and the accuracy was 0.905 for the AI model for indicating the presence of FNH using six MR sequences. These AUC values seem to be lower than some values from previous studies, and a higher accuracy value was obtained than from some other investigations. However, this study cannot be compared exactly with the other studies mentioned above. Datasets from previous studies including various lesions cannot be compared with the dataset from this study since this research was solely focused on FNH lesions. The AI results were better than those of the radiology resident; however, they were lower than those of the radiology specialist. Nonetheless, it was remarkable that a good agreement level was indicated between the radiology specialist and the AI model according to the results of this study. This might be caused by the lack of diversity among the FNH lesions in the MRIs experienced by the AI model in the training session. The radiology specialist's years of experience cannot be compared with the AI's training image dataset, which only included 30 patients with FNH lesions. This gap between AI and the radiology specialist might be compensated for by introducing a larger number and greater variety of FNH lesions to the AI model.

Important factors such as feasibility, ethical concerns, precision, safety, and overall acceptability influence the application speed of auto-diagnosis systems in medicine. Collaboration between healthcare professionals and AI-based diagnostic systems remains a mandatory objective for succeeding in this difficult task, and AI can still not replace skilled diagnosticians.²⁹

There are limitations to this study, which must be considered when interpreting its outcomes. First, the presence of FNH was determined with only six MR sequences on axial planes to standardize, simplify, and easily compare the interpretation results. It also helped the standardization of the segmentation process, which should have been performed meticulously as part of a long-lasting process. However, a standard interpretation of the liver MRIs needed all the sequences obtained during the imaging procedure. If the patient had one or more lesions consistent with FNH, they were accepted as FNH positive by each reviewer, and chi-square tests were performed using these results. The AI model used in the study indicated the results regarding the presence or absence of FNH as an outcome. To compare the results and calculate interobserver reliability accurately, the study was planned in this way. This methodological approach might be criticized in terms of its appropriateness for indicating the sensitivity and specificity values. A lesion-based model rather than patient-based evaluation results would provide more accurate outcomes. The detection of the lesion was based on signal properties and dynamic enhancement patterns, but the borders of the lesion were underestimated. A morphological approach using the border attributes would be a more realistic approach, similar to the routine radiological liver interpretations on MRIs. Having more patients with hepatic adenomas and HCC to evaluate the ability of AI in distinguishing FNH from other lesions would be better. Some of the patients in this study were used in the AI training process, and some were not suitable for the investigation because of motion artifacts or image distortions. Moreover, we could only share the results of patients confidently diagnosed either radiologically or histologically. Although a variety of lesions with different histological and imaging features were evaluated in this study, additional studies with larger sample sizes are needed to confirm the results of this investigation. Due to the extremely detailed and very long-lasting process of segmentation, the proximal branches of the hepatic and portal veins were mapped to introduce them to the AI model. It was expected that the more distant segments would be perceived by the AI model, as it was part of the program. To minimize the AI model's possible segmentation and interpretation errors, the more distant segments of the vessels might also be drawn manually.

In conclusion, the AI model provided remarkable sensitivity, specificity, PPV, and

NPV results regarding the detection of FNH in this study. The potential of AI should not be underestimated since this current investigation indicated that AI achieved better results than a radiology resident. Through multidisciplinary studies based on the increasing interest of physicians and engineers, AI might become a crucial element in diagnostics and play a major role in the detection and characterization of liver lesions. Targeted studies focused on specific lesions may be combined in the same diagnostic tool, using the experience of all focal lesions of the liver to widen the spectrum of lesions recognized by AI.

Footnotes

Conflict of interest disclosure

Sonay Aydın, MD, is Section Editor in Diagnostic and Interventional Radiology. He had no involvement in the peer-review of this article and had no access to information regarding its peer-review. Other authors have nothing to disclose.

References

- Pompili M, Ardito F, Brunetti E, et al. Benign liver lesions 2022: Guideline for clinical practice of Associazione Italiana Studio del Fegato (AISF), Società Italiana di Radiologia Medica e Interventistica (SIRM), Società Italiana di Chirurgia (SIC), Società Italiana di Ultrasonologia in Medicina e Biologia (SIUMB), Associazione Italiana di Chirurgia Epatobilio-Pancreatica (AICEP), Società Italiana Trapianti d'Organo (SITO), Società Italiana di Anatomia Patologica e Citologia Diagnostica (SIAPEC-IAP) - Part II - Solid lesions. *Dig Liver Dis*. 2022;54(12):1614-1622. [\[Crossref\]](#)
- Basturk O, Farris AB, Adsay NV. Chapter 15 - Immunohistology of the pancreas, biliary tract, and liver, editor(s): David J. Dabbs. *Diagnostic Immunohistochemistry (Third Edition)*. 2011; p.541-592 [\[Crossref\]](#)
- Ding Z, Lin K, Fu J, et al. An MR-based radiomics model for differentiation between hepatocellular carcinoma and focal nodular hyperplasia in non-cirrhotic liver. *World J Surg Oncol*. 2021;19(1):181. [\[Crossref\]](#)
- LeGout JD, Bolan CW, Bowman AW, et al. Focal nodular hyperplasia and focal nodular hyperplasia-like lesions. *Radiographics*. 2022;42(4):1043-1061. [\[Crossref\]](#)
- Kamel IR, Liapi E, Fishman EK. Focal nodular hyperplasia: lesion evaluation using 16-MDCT and 3D CT angiography. *AJR Am J Roentgenol*. 2006;186(6):1587-1596. [\[Crossref\]](#)
- Giambelluca D, Taibbi A, Midiri M, Bartolotta TV. The "spoke wheel" sign in hepatic focal nodular hyperplasia. *Abdom Radiol (NY)*. 2019;44(3):1183-1184. [\[Crossref\]](#)

- European Association for the Study of the Liver (EASL). EASL clinical practice guidelines on the management of benign liver tumours. *J Hepatol*. 2016;65(2):386-398. [\[Crossref\]](#)
- Murakami T, Tsurusaki M. Hypervascular benign and malignant liver tumors that require differentiation from hepatocellular carcinoma: key points of imaging diagnosis. *Liver Cancer*. 2014;3(2):85-96. [\[Crossref\]](#)
- Suh CH, Kim KW, Kim GY, Shin YM, Kim PN, Park SH. The diagnostic value of Gd-EOB-DTPA-MRI for the diagnosis of focal nodular hyperplasia: a systematic review and meta-analysis. *Eur Radiol*. 2015;25(4):950-960. [\[Crossref\]](#)
- Berbis MA, Paulano Godino F, Royuela Del Val J, Alcalá Mata L, Luna A. Clinical impact of artificial intelligence-based solutions on imaging of the pancreas and liver. *World J Gastroenterol*. 2023;29(9):1427-1445. [\[Crossref\]](#)
- Hussain SM, Terkivatan T, Zondervan PE, Lanjouw E, de Rave S, Ijzermans JN, de Man RA. Focal nodular hyperplasia: findings at state-of-the-art MR imaging, US, CT, and pathologic analysis. *Radiographics*. 2004;24(1):3-17;discussion 18-9. [\[Crossref\]](#)
- Antonelli M, Reinke A, Bakas S, et al. The medical segmentation decathlon. *Nat Commun*. 2022;13(1):4128. [\[Crossref\]](#)
- Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods*. 2021;18(2):203-211. [\[Crossref\]](#)
- Bertels J, Eelbode T, Berman M, et al. Optimizing the Dice Score and Jaccard index for medical image segmentation: theory and practice. In: *Medical Image Computing and Computer Assisted Intervention - MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13-17, 2019, Proceedings, Part II*. 2019. p. 92-100. [\[Crossref\]](#)
- <https://github.com/lasooan/SlicerElastix#slicerelastix> [\[Crossref\]](#)
- Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-163. Erratum in: *J Chiropr Med*. 2017;16(4):346. [\[Crossref\]](#)
- Hanna EJ, Ismail N, Arsalane A, Quenum C, Moumen A, Sacrieru D, Kabbej M, Khadra J. A case of liver rupture in a patient with focal nodular hyperplasia at 33 weeks of gestation: a multidisciplinary management. *Gynecology and Obstetrics Clinical Medicine*. 2022; 2: 46-48. [\[Crossref\]](#)
- Tsilimigras DI, Rahnamai-Azar AA, Ntanasis-Stathopoulos I, et al. Current approaches in the management of hepatic adenomas. *J Gastrointest Surg*. 2019;23(1):199-209. Erratum in: *J Gastrointest Surg*. 2020;24(1):232. [\[Crossref\]](#)

19. Schütte K, Schulz C, Poranzke J, et al. Characterization and prognosis of patients with hepatocellular carcinoma (HCC) in the non-cirrhotic liver. *BMC Gastroenterol.* 2014;14:117. [\[Crossref\]](#)
20. Jansen MJA, Kuijff HJ, Veldhuis WB, Wessels FJ, Viergever MA, Pluim JPW. Automatic classification of focal liver lesions based on MRI and risk factors. *PLoS One.* 2019;14(5):e0217053. [\[Crossref\]](#)
21. Pettit RW, Marlatt BB, Corr SJ, Havelka J, Rana A. nnU-Net deep learning method for segmenting parenchyma and determining liver volume from computed tomography images. *Ann Surg Open.* 2022;3(2):e155. [\[Crossref\]](#)
22. Duc VT, Chien PC, Huyen LDM, et al. Deep learning model with convolutional neural network for detecting and segmenting hepatocellular carcinoma in CT: a preliminary study. *Cureus.* 2022;14(1):e21347. [\[Crossref\]](#)
23. Goehler A, Harry Hsu TM, Lacson R, et al. Three-dimensional neural network to automatically assess liver tumor burden change on consecutive liver MRIs. *J Am Coll Radiol.* 2020;17(11):1475-1484. [\[Crossref\]](#)
24. Zhang X, Kanematsu M, Fujita H, et al. Application of an artificial neural network to the computer-aided differentiation of focal liver disease in MR imaging. *Radiol Phys Technol.* 2009;2(2):175-182. [\[Crossref\]](#)
25. Hamm CA, Wang CJ, Savic LJ, et al. Deep learning for liver tumor diagnosis part I: development of a convolutional neural network classifier for multi-phasic MRI. *Eur Radiol.* 2019;29(7):3338-3347. [\[Crossref\]](#)
26. Zhen SH, Cheng M, Tao YB, et al. Deep learning for accurate diagnosis of liver tumor based on magnetic resonance imaging and clinical data. *Front Oncol.* 2020;10:680. [\[Crossref\]](#)
27. Stollmayer R, Budai BK, Tóth A, et al. Diagnosis of focal liver lesions with deep learning-based multi-channel analysis of hepatocyte-specific contrast-enhanced magnetic resonance imaging. *World J Gastroenterol.* 2021;27(35):5978-5988. [\[Crossref\]](#)
28. Wang SH, Han XJ, Du J, et al. Saliency-based 3D convolutional neural network for categorising common focal liver lesions on multisequence MRI. *Insights Imaging.* 2021;12(1):173. [\[Crossref\]](#)
29. Popa SL, Grad S, Chiarioni G, et al. Applications of artificial intelligence in the automatic diagnosis of focal liver lesions: a systematic review. *J Gastrointest Liver Dis.* 2023;32(1):77-85. [\[Crossref\]](#)