Diagnostic Value of Magnetic Resonance Elastography Radiomics Analysis for the Assessment of Hepatic Fibrosis in Patients With Nonalcoholic Fatty Liver Disease

Ki Choon Sim
Korea University Anam Hospital, Korea University Medicine

Min Ju Kim (✉ dr.minjukim@gmail.com)
Korea University Anam Hospital, Korea University Medicine

Yongwon Cho
Korea University Anam Hospital, Korea University Medicine

Hyun Jin Kim
Korea University Anam Hospital, Korea University Medicine

Beom Jin Park
Korea University Anam Hospital, Korea University Medicine

Deuk Jae Sung
Korea University Anam Hospital, Korea University Medicine

Yeo Eun Han
Korea University Anam Hospital, Korea University Medicine

Na Yeon Han
Korea University Anam Hospital, Korea University Medicine

Tae Hyung Kim
Korea University Ansan Hospital, Korea University Medicine

Yoo-jin Lee
Korea University Anam Hospital, Korea University Medicine

Research Article

Keywords: Radiomics, Magnetic resonance imaging, Elastography, Nonalcoholic fatty liver disease, Liver fibrosis

DOI: https://doi.org/10.21203/rs.3.rs-798827/v1

License: ☺️ This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

**Background:** To investigate the diagnostic performance of radiomics analysis using magnetic resonance elastography (MRE) toward assessing hepatic fibrosis in patients with nonalcoholic fatty liver disease (NAFLD).

**Methods:** A total of 100 patients with suspected NAFLD were retrospectively enrolled. All patients underwent a liver parenchymal biopsy. MRE was performed using a 3.0-T scanner. Following three-dimensional (3D) segmentation of MRE images, 834 radiomic features were analyzed using a commercial program. Radiologic features, such as median and mean values of two-dimensional (2D) or 3D regions of interest (ROIs) and variable clinical features, were analyzed. A random forest regressor was employed to extract important radiomic, radiological, and clinical features. A random forest classifier model was trained to use these features to classify the fibrosis stage. The area under the receiver operating characteristic curve (AUC) was evaluated using a classifier for fibrosis stage diagnosis.

**Results:** The pathological hepatic fibrosis stage was classified as low-grade fibrosis (stages F0–F1, n = 82) or clinically significant fibrosis (stages F2–F4, n = 18). Eight important features were extracted from radiomics analysis, with the two most important being wavelet-HHL gray level dependence matrix (GLDM)-dependence non-uniformity-normalized and wavelet-HHL GLDM-dependence entropy. The median value of the 2D ROI was identified as the most important radiologic feature. Platelet count was identified as an important clinical feature. The AUC of the classifier using radiomics was comparable to that of radiologic measures (0.97 ± 0.07 vs. 0.96 ± 0.06).

**Conclusions:** MRE radiomics analysis provides diagnostic performance comparable to conventional MRE analysis for the assessment of clinically significant hepatic fibrosis in patients with NAFLD.

Background

Chronic liver diseases are a common cause of morbidity and mortality worldwide, although specific diseases differ between regions [1]. The incidence of nonalcoholic fatty liver disease (NAFLD) has steadily increased in the general population [2, 3]. Up to 40% of cases have been reported in the general population, and most patients are asymptomatic and only have simple steatosis. However, a small subset does have significant inflammation with or without fibrosis [3, 4]. Because serious complications, including hepatocellular carcinoma, occur in advanced fibrosis [3, 5], the accurate detection of significant hepatic fibrosis is of utmost importance.

Although percutaneous liver biopsy is the gold standard method for staging liver fibrosis, it has several drawbacks, including procedure-related complications and sampling errors [6, 7]. To overcome the limitations of liver parenchymal biopsies, various noninvasive quantitative methods, including ultrasound and magnetic resonance elastography (MRE), have been introduced and employed. MRE shows better organ coverage, reproducibility, diagnostic accuracy, sensitivity, and specificity as compared to other methods [8, 9]. However, existing image analysis and post-processing MRE methods are limited by the
exclusive use of mean signal intensity values through drawing of region of interest (ROI) [10]. To our knowledge, till date, there have been no studies that have attempted MRE analysis using other methods.

Radiomics analysis is a post-processing method that extracts and analyzes dozens to hundreds of features from various medical images. It extracts additional radiological information from features within images and spatial variations in pixel intensities that are undetectable by human perception [11–13]. Radiomics analysis of hepatic fibrosis using magnetic resonance imaging (MRI) has been examined in previous studies; however, these studies targeted patients with chronic viral hepatitis or used other sequences such as T2-weighted images or the hepatobiliary phase [14–16]. Radiomic analysis of hepatic fibrosis in patients with NAFLD has not been conducted thus far. We hypothesized that radiomic features extracted from MRE images would allow more accurate assessments of hepatic fibrosis than conventional ROI methods or the use of various clinical data. Accordingly, our study investigated the diagnostic performance of MRE radiomics for the assessment of clinically significant hepatic fibrosis in patients with NAFLD.

Methods

Study population

Approval for this retrospective investigation study was obtained from the Institutional review board (approval No. 2020AN0387), and the requirement for informed consent was waived. Between November 2017 to May 2020, 190 MREs were acquired. The inclusion criteria were as follows: (a) age ≥ 18 years, (b) no known history of liver disease, (c) suspected NAFLD on screening based on ultrasound and laboratory studies, and (d) ≤ 60 days between MRE and liver biopsy. The exclusion criteria were as follows: (a) history of alcohol consumption; (b) history of chronic liver diseases such as chronic hepatitis B or C infection, autoimmune hepatitis, and primary sclerosing cholangitis; and (c) history of major liver surgery such as liver transplantation and hemihepatectomy. Based on these criteria, 100 consecutive patients were enrolled during the study period (Fig. 1).

Demographic, laboratory, and clinical features, including age, sex, weight, height, and blood test results (alanine aminotransferase [ALT], aspartate aminotransferase [AST], triglyceride, low-density lipoprotein, and platelet count) obtained within one month of MRE were evaluated. The body mass index and aspartate aminotransferase-to-platelet ratio index (APRI) were calculated.

Reference standard for liver fibrosis

Pathologic examination of the liver served as the reference standard for liver fibrosis. All patients underwent percutaneous liver biopsy using an 18-gauge needle with a 20-mm penetration depth targeting segment 5/6. Histology preparations from liver biopsies were retrospectively reviewed by one pathologist (7 years of experience) who was blinded to the clinical data and MRE results. The classification by Kleiner et al. [17] was used to grade and stage NAFLD. Fibrosis was staged between 0 and 4 as follows: F0, absence of fibrosis; F1, perisinusoidal or portal; F2, perisinusoidal and portal/periportal; F3, septal or
bridging fibrosis; and F4, cirrhosis. The pathologic hepatic fibrosis stage was classified as mild fibrosis (F0 or F1) or clinically significant fibrosis (F2–F4) [18, 19].

MRI examination

MRI examinations were performed using 3.0-T scanners (Magnetom Skyra; Siemens Healthineers, Erlangen, Germany) with a 30-channel body coil. MRE was performed according to a previously described protocol [20, 21] using commercial hardware and software (Resoundant Inc., Rochester, MN, USA; Syngo MR E11, Siemens Healthineers). In the supine position, the passive acoustic driver was placed on the right chest wall and upper abdominal wall, with its center at the xiphoid process level. An elastic strap was used to secure it to the patient's body. MRE was performed using a phase-contrast gradient-recalled echo sequence, which applies motion-encoding gradients in synchronization with a 60-Hz external shear wave induced in the abdomen. The total acquisition time was approximately 20 s per slice, and four contiguous slices were obtained for each patient. The acquisition parameters were as follows: axial plane, field of view, 380 mm × 285 mm; acquisition matrix, 128×77; flip angle, 25°; number of excitations, 1; repetition time ms/echo time ms, 50/17; and slice thickness and gap, 5 mm. After the magnitude and phase images were obtained, an inversion algorithm installed in the MRI unit automatically processed raw data images to create several additional images and maps [22].

MRE analysis processing

Three-dimensional (3D) segmentation of MRE images was performed by three radiologists (two abdominal radiologists with 22 years [n = 21] and 10 years [n = 28] of clinical experience, respectively, and a 2-year resident [n = 51]) using a commercial program, AVIEW (version 1.0.32.12, Coreline Soft, Seoul, Korea). As it was a time-consuming task, the three radiologists were randomly assigned patients to perform 3D segmentation. The segmentation was conducted on four contiguous grayscale elastograms with a 95% confidence map (Fig. 2). All radiologists were trained by a software applicator to improve segmentation accuracy before they started the process. After completing the 3D segmentation, the mean and median values of liver fibrosis obtained from AVIEW were organized as 3D ROI values.

Using the grayscale elastogram with a 95% confidence map, two circular ROIs were defined per slice, and up to eight fibrosis values were obtained for each patient (Fig. 3). The ROI area was maintained at approximately 300–350 mm². ROIs were drawn at two separate sites or at one site while avoiding the edges of the liver [23]. All postprocessing was performed using a commercial workstation by a single abdominal radiologist who had 10 years of experience and was blinded to the clinicopathological data. The obtained liver fibrosis values were organized using mean and median values (2D ROI values). The 2D and 3D ROI values were used for the analysis of radiologic features.

Radiomic feature extraction

Based on MRE data with 3D segmentation applied, several hundreds of radiomic features were analyzed by an artificial intelligence research professor using PyRadiomics (version 3.0, PyRadiomics Community)
The features included were related to shape, first-order statistical, second-order statistical (including so-called textural features such as the gray-level co-occurrence matrix, gray-level run-length matrix, gray-level dependence matrix, GLDM, gray level size zone matrix, and neighboring gray-tone difference matrix), and higher-order statistical using wavelet filters.

**Data analysis**

*Feature selection and classification method using machine learning*

Although many quantitative features (radiomics, radiologic, and clinical) can be extracted from medical datasets, these may be highly correlated with each other or simply considered as noise. Thus, it is important to reduce features to select a subset of specific features, enhance the performance, and minimize the computational cost. Among radiomics, radiologic, and clinical features, important features for predicting low-grade or clinically significant hepatic fibrosis in patients with NAFLD were selected using a random forest regressor in Python (Python Software Foundation, version 3.6) with the Scikit-learn package (https://github.com/scikit-learn/scikit-learn). A random forest classifier model [25] was trained to use these important features to classify the fibrosis stage. The 20 repeated 10-fold stratified cross-validations verified the stability of the results. We evaluated the area under the receiver operating characteristic curve (AUC) and classifier accuracy. The classifier diagnosed the fibrosis stage based on radiomic, radiologic, or clinical features, or a combination of all features. Statistical differences in the AUC according to each classifier were compared using a machine learning model with DeLong’s test. \( P \) values <0.05 were considered statistically significant.

*Conventional statistical analyses*

The demographic and clinical data in the low-grade fibrosis and clinically significant fibrosis groups were compared using the Mann–Whitney U test, chi-squared test, or Fisher’s exact test. The liver stiffness determined by the measurements was compared using a paired t-test. All statistical analyses were performed using SPSS Statistics for Windows (version 20.0; IBM Corp., NY, USA). \( P \) values <0.05 were considered statistically significant.

**Results**

*Demographic and clinical characteristics*

As mentioned earlier, 100 patients were included in this study (Fig. 1). The mean interval between MRE and liver biopsies was 1.29 (range, 0–33) days. In 97 patients, the interval between MRE and liver biopsies was <3 days. Among the 100 patients, 16 were classified as having F0, 66 as having F1, five as having F2, six as having F3, and seven as having F4. Thus, the low-grade fibrosis group included 82 patients, while the clinically significant fibrosis group included 18 patients. Iron overload was suspected in only one patient in the MRE; however, this was not demonstrated in the pathology review. The demographic features of patients with NAFLD are shown in Table 1.
Liver stiffness measurement and MRE evaluation

We aimed to obtain eight 2D ROIs for each patient. However, because of some inappropriately acquired slices, one patient had five ROIs, two had six ROIs, and four had seven ROIs. Liver stiffness values were organized according to the measurement method used in the two patient groups (Table 2). There was a significant difference in liver stiffness values between the two groups, regardless of the ROI measurement method ($P<0.001$). There was no remarkable difference in 2D or 3D ROI values (approximately 2.5–2.7 kPa) in the mild fibrosis group. However, in the significant fibrosis group, the 3D ROI value was lower than the 2D value (approximately 4.67–4.83 kPa vs. 5.3 kPa, $P<0.001–0.002$).

Comparison of diagnostic performances for classifying liver stiffness using machine learning

The diagnostic performance for classifying liver stiffness and the important features extracted by each method are shown in Table 3.

Classifying liver stiffness using radiomics

The AUC and accuracy to discriminate between mild and significant fibrosis using radiomics were 0.97 and 0.94, respectively. A total of eight radiomic features were extracted: wavelet-HHL GLDM-dependence non-uniformity normalized, wavelet-HHL GLDM dependence entropy, and original first-order median were determined to be the three most important features.

Classifying liver stiffness using clinical features

The AUC and accuracy of clinical features were 0.91 and 0.86, respectively. A total of nine features were extracted; platelet count, APRI, and age were extracted as the three most important features.

Classifying liver stiffness using radiologic ROI measures

The AUC and accuracy of radiologic ROI measurements alone were 0.96 and 0.92, respectively. The median of the 2D ROIs was extracted as the most important feature.

Classifying liver stiffness using a combination of radiomic, clinical, and radiologic measures

The AUC and accuracy of all combinations of variable features were 0.98 and 0.95, respectively. These values were higher than those for each of the features mentioned above. A total of 11 important features were extracted: wavelet-HHL GLDM dependence entropy, wavelet-HHL GLDM-dependence non-uniformity normalized, and the median of 2D ROIs were extracted as the three most important features.

Comparisons of AUCs of machine learning models with DeLong’s test

The AUC of the classifier determined using a combination of radiomics, radiologic measures, and clinical features (AUC = 0.98, accuracy = 0.95) was slightly higher than that using radiomics (AUC = 0.97, accuracy = 0.94) or radiologic measures (AUC = 0.96, accuracy = 0.92) alone (Fig. 5). Clinical features
alone showed the lowest diagnostic performance (AUC = 0.91, accuracy = 0.86). There was a significant difference between usage of clinical features alone with that of the usage of combination of all features (P = 0.011), and use of radiomics alone (P = 0.039) (Table 4). However, there was no significant difference between using a combination of all features and radiomics alone (P = 0.960) or the combination of all features and radiologic measures (P = 0.254). Similarly, the virtually identical AUCs from radiologic measures and radiomics suggest that the added complexity of radiomics analysis does not afford a significant benefit (P = 0.694).

Discussion

Our study demonstrated that radiomics analysis of MRE (AUC = 0.97 ± 0.07) could accurately differentiate clinically significant hepatic fibrosis (F2–F4) from low-grade fibrosis (F0 or F1) in patients with NAFLD. The AUC for MRE radiomics features (0.97 ± 0.07) was comparable with that for radiologic measures (0.96 ± 0.06). Therefore, the results of our study suggest that radiomics features of MRE analyzed using a machine learning approach can accurately stage liver fibrosis. The purpose of our study was to investigate the difference between MRE and ROI measurement, which is widely used as a gold standard, with radiomics analysis, having recently been in the spotlight. However, the AUC of MRE radiomics analysis showed no significant improvement compared to the conventional radiologic measures (P = 0.694). Our results suggest that MRE radiomics analysis is not as effective as the conventional measurement method in terms of time and effort invested. In future, these disadvantages may be resolved by further development of software such as auto segmentation tools. This in turn would permit easy application of radiomics analysis in routine practice.

Previous studies using radiomics analysis and machine learning have attempted to differentiate the fibrosis stages of chronic liver disease using various MRI sequences. Lan et al. conducted a study to discriminate clinically significant fibrosis based on MRE [14]; however, because their study included patients with hepatitis B/C and analyzed the AUC for each radiomic feature, directly comparison of the results of the two studies is not trivial. Other studies have attempted to predict liver stiffness by performing radiomics analysis using the hepatobiliary phase [15], or T1-weighted or T2-weighed images [16, 26] instead of MRE. Moreover, these three previous studies evaluated heterogeneous chronic liver diseases. Thus, it is challenging to directly compare the AUCs obtained in our study with these studies (AUC = 0.84–0.934) because of differences in patient groups, MR equipment, and MRI sequences.

Considering each of the important features extracted in the MRE radiomics analysis and the combination of all features, we found two of them: GLDM-dependence non-uniformity-normalized and GLDM-dependence entropy with a wavelet-HHL filter, to be more important than the others. Because the grayscale elastogram with a 95% confidence map reflects liver stiffness, a specific radiomic feature related to the gray level in the image is considered an important feature. Several radiomic features were presented by Lan et al. [14], who used MRE radiomics to diagnose clinically significant fibrosis in patients with chronic viral hepatitis; however, none of the radiomic features matched those identified in our study. A standard model for radiomic feature analysis has not been established, and the use of MR equipment
with variable parameters due to vendor differences limits its clinical application. Thus, additional research is needed to identify the standard model and the most meaningful radiomic features.

In general, the shape features of radiomics analysis are reported to be of importance since they reflect the aggressiveness of several malignant tumors [27]. In radiologic imaging of diffuse liver disease, depending on the cause and stage of the disease, morphological changes distinguishable from normal liver are observed [28]; however, in the current study, shape features were not identified to be of importance in any analysis. There are several possible reasons for this observation. First, few patients progressed to advanced fibrosis, accompanied by morphological alterations. Second, subtle surface irregularities were not reflected during 3D segmentation. Finally, because the grayscale elastogram with a 95% confidence map does not accurately reflect the liver surface or whole volume, only four contiguous slices were obtained.

The median of the 2D ROIs was extracted as the third most important feature in the combination of all features. Moreover, it was identified as the most important radiologic measure. The finding that the median of multiple small ROIs is more important than generally used radiologic measures, such as the mean ROI of a whole volume or mean of multiple small ROIs [10, 23], is significant. Thus, MRE analysis can yield sufficient accuracy with only two appropriate ROIs per image, rather than using whole volume segmentation. Similarly, in liver stiffness measurements using FibroScan® (Echosens, France), which is widely used, the median value, and not the mean value of several repeated measurements, has been used [29].

Liver stiffness in the significant fibrosis group differed depending on whether it was a 2D or 3D ROI measurement; the overall liver stiffness was lower as per the 3D ROI measurement (Table 3). As the 2D ROI measurement was selected as a feature that is more important than the 3D measurement in the diagnostic performance analysis; it can be assumed that liver stiffness is generally lower with 3D measurements than with 2D ROI measurements (median 4.67 kPa vs, 5.33 kPa, \( P < 0.001 \)). A possible reason could be that peripheral portions or areas showing relatively low stiffness in patients with advanced fibrosis were included in the 3D segmentation (Fig. 6). However, when 2D ROIs were drawn for the same patient, the drawings were only in the central portion, which showed relatively high stiffness. Another possible reason may be that the overall stiffness was lower because of the included vessel area, considering that it is impossible to segment a whole volume excluding vessels in a 95% map with a low resolution. With manual ROI drawing, the possibility of liver stiffness masking by blood vessels may be lowered because blood vessels can be avoided following correlations with other imaging sequences. As shown in the present study, to avoid inclusion of peripheral liver portions, which can affect significant fibrosis diagnosis, FibroScan® recommends avoiding vessels and obtaining measurements at a depth of at least 25–65 mm from the skin [30].

Several demographics demonstrated significant differences between the mild and significant fibrosis groups (Table 1). Platelet count and AST showed the most significant differences, and there were significant differences in the APRI and AST/ALT ratio calculated using these two features. In both groups,
ALT was elevated to a similar degree; however, in the mild fibrosis group, the AST value was relatively less elevated (48.8 ± 31.2 IU/L vs. 78.1 ± 60.1 IU/L, \( P < 0.05 \)). A high AST/ALT ratio in patients with elevated AST and ALT levels is considered an important clinical finding suggestive of clinically significant fibrosis. Diabetes and old age were also found to differ between the groups, consistent with known risk factors for NAFLD [31]. In the analysis of clinical features using machine learning, platelet count, APRI, old age, and diabetes were identified as the top four important features (Table 3).

Our study had some limitations. First, this retrospective study might have had a selection bias. The pathologic fibrosis stage in our study population was not equivalent, and there were more patients in stages F0 and F1 than in stages F2–F4. This may indicate a spectrum bias and may have led to an overestimation of diagnostic performance. In addition, since this was a retrospective study, the reproducibility of the MRE examination could not be assessed. Second, the study population size was not large enough to be divided into training and validation sets to be used with machine learning. To overcome this limitation, we analyzed the results using 20 repeated 10-fold stratified cross-validation. This method, as shown in previous studies [32–35], was the best way to increase the reliability of the evaluation using limited data. Third, external validation for machine learning was not performed, limiting the generalizability of our results. Finally, the inter- and intraobserver reliabilities of the 2D and 3D ROI measures could not be assessed. However, these shortcomings would have been overcome because 2D ROI drawings were performed up to eight times, and 3D segmentation was performed after sufficient training by a software applicator.

Conclusions

MRE radiomics analysis provides diagnostic performance comparable to that of conventional MRE analysis for the assessment of clinically significant hepatic fibrosis in patients with NAFLD.

Abbreviations

MRE, magnetic resonance elastography; NAFLD, nonalcoholic fatty liver disease; 3D, three-dimensional; 2D, two-dimensional; ROI(s), region-of-interest(s); AUC, area under the receiver operating characteristic curve; ALT, alanine aminotransferase; AST, aspartate aminotransferase; APRI, aspartate aminotransferase-to-platelet ratio index; GLDM, gray level dependence matrix.

Declarations

Ethics approval and consent to participate

The study was approved by the institutional review board of the Korea University Anam Hospital (approval No. 2020AN0387), we retrospectively enrolled 190 cases of magnetic resonance elastography of the liver from November 2017 to May 2020, and the requirement for informed consent was waived.

Consent for publication
Not applicable.

Availability of data and materials

The datasets used or analyzed during the current study are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no conflicts of interest.

Funding

Not applicable.

Authors’ contributions

KCS proposed methods, performed the computations, and wrote the paper. MJK devised and supervised the project. MJK guaranteed the integrity of the entire study. KCS, MJK, YWC, BJP, DJS, YEH, NYH, THK, and YJL contributed to the interpretation of the data. KCS, YWC, and HJK contributed to the analysis and interpretation of the data. YWC solved some detailed problems in a deep learning experiment. All authors read and approved the final manuscript.

Acknowledgments

Not applicable.

References


Tables

Table 1. Demographic characteristics of patients with nonalcoholic fatty liver disease (NAFLD)
<table>
<thead>
<tr>
<th>Patients</th>
<th>Low-grade fibrosis (n = 82)</th>
<th>Clinically significant fibrosis (n = 18)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (male:female)</td>
<td>57:25</td>
<td>9:9</td>
<td>0.114</td>
</tr>
<tr>
<td>Hypertension (+ve)</td>
<td>24 (29.3%)</td>
<td>8 (44.4%)</td>
<td>0.211</td>
</tr>
<tr>
<td>Diabetes mellitus (+ve)</td>
<td>11 (13.4%)</td>
<td>11 (61.1%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age (years)</td>
<td>45.4 ± 13.3</td>
<td>56.6 ± 10.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>28.6 ± 5.2</td>
<td>27.5 ± 4.3</td>
<td>0.501</td>
</tr>
<tr>
<td>AST (IU/L)</td>
<td>48.8 ± 31.2</td>
<td>78.1 ± 60.1</td>
<td>0.015</td>
</tr>
<tr>
<td>ALT (IU/L)</td>
<td>85.3 ± 65.7</td>
<td>71.6 ± 45.7</td>
<td>0.726</td>
</tr>
<tr>
<td>AST/ALT ratio</td>
<td>0.7 ± 0.4</td>
<td>1.4 ± 1.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total TG (mg/dL)</td>
<td>187.8 ± 141.6</td>
<td>184.3 ± 89.5</td>
<td>0.615</td>
</tr>
<tr>
<td>LDL</td>
<td>121.7 ± 30.6</td>
<td>113.0 ± 33.4</td>
<td>0.24</td>
</tr>
<tr>
<td>Platelet count (x 10⁶/L)</td>
<td>257.9 ± 55.6</td>
<td>171.0 ± 35.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>APRI</td>
<td>0.50 ± 0.34</td>
<td>1.21 ± 0.95</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NAS</td>
<td>2.89 ± 1.51</td>
<td>3.60 ± 1.77</td>
<td>0.145</td>
</tr>
</tbody>
</table>

P values were derived using the independent t-test, Mann–Whitney U test, chi-squared test, or Fisher’s exact test.

NAFLD, nonalcoholic fatty liver disease; BMI, body mass index; AST, aspartate transaminase; ALT, alanine aminotransferase; TG, triglyceride; LDL, low-density lipoprotein; APRI, aspartate aminotransferase-to-platelet ratio index; NAFLD, nonalcoholic fatty liver disease.

Table 2. Measurement of liver stiffness using magnetic resonance elastography in patients with nonalcoholic fatty liver disease
<table>
<thead>
<tr>
<th>Liver stiffness measurement methods</th>
<th>Low-grade fibrosis, n = 82</th>
<th>Clinically significant fibrosis, n = 18</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD (kPa)</td>
<td>Range (kPa)</td>
<td>Mean ± SD (kPa)</td>
</tr>
<tr>
<td>2D ROI measurement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>2.53 ± 0.62</td>
<td>1.56–4.25</td>
<td>5.33 ± 1.65</td>
</tr>
<tr>
<td>Mean</td>
<td>2.58 ± 0.63</td>
<td>1.59–4.25</td>
<td>5.34 ± 1.69</td>
</tr>
<tr>
<td>3D ROI measurement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>2.51 ± 0.58</td>
<td>1.81–4.13</td>
<td>4.67 ± 1.23</td>
</tr>
<tr>
<td>Mean</td>
<td>2.69 ± 0.54</td>
<td>1.92–4.36</td>
<td>4.83 ± 1.24</td>
</tr>
</tbody>
</table>

SD, standard deviation; ROI, region of interest; 2D, two-dimensional; 3D, three-dimensional

Table 3. Diagnostic performance for classifying liver stiffness and several important features of extracted for each method for in patients with nonalcoholic fatty liver disease
<table>
<thead>
<tr>
<th>Method of classifying liver stiffness</th>
<th>AUC</th>
<th>Accuracy</th>
<th>Feature</th>
<th>Feature importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiomics</td>
<td>0.97 ± 0.07</td>
<td>0.94 ± 0.13</td>
<td>Wavelet-HHL GLDM dependence non-uniformity normalized</td>
<td>0.31005 ± 0.02728</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wavelet-HHL GLDM dependence entropy</td>
<td>0.24375 ± 0.01775</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Original first order median</td>
<td>0.11797 ± 0.01811</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wavelet-LLL first order median</td>
<td>0.09658 ± 0.02037</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wavelet-HHL GLDM dependence variance</td>
<td>0.08890 ± 0.01732</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wavelet-HHL GLCM joint energy</td>
<td>0.06421 ± 0.02168</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wavelet-HHH GLRML low gray level run emphasis</td>
<td>0.05277 ± 0.01339</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wavelet-HHL GLCM maximum probability</td>
<td>0.02576 ± 0.01374</td>
</tr>
<tr>
<td>Clinical features</td>
<td>0.91 ± 0.11</td>
<td>0.86 ± 0.20</td>
<td>Platelet count</td>
<td>0.33412 ± 0.03177</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>APRI</td>
<td>0.13672 ± 0.01692</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Age</td>
<td>0.11368 ± 0.01720</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Diabetes mellitus</td>
<td>0.08113 ± 0.01571</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Triglycerides</td>
<td>0.07980 ± 0.00745</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Low-density lipoprotein</td>
<td>0.07809 ± 0.01098</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Aspartate aminotransferase</td>
<td>0.07480 ± 0.01358</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Alanine aminotransferase</td>
<td>0.05398 ± 0.00850</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Body mass index</td>
<td>0.04768 ± 0.00611</td>
</tr>
<tr>
<td>Radiologic measures</td>
<td>0.96 ± 0.17</td>
<td>0.92 ± 0.17</td>
<td>Median of 2D ROIs</td>
<td>0.30721 ± 0.06659</td>
</tr>
<tr>
<td>Combination of radiomic, clinical, and radiologic measures</td>
<td>0.98 ± 0.04</td>
<td>0.95 ± 0.10</td>
<td>Wavelet-HHL GLDM dependence entropy</td>
<td>0.17780 ± 0.02365</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wavelet-HHL GLDM dependence non-uniformity normalized</td>
<td>0.15993 ± 0.02046</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Median of 2D ROIs</td>
<td>0.15427 ± 0.02729</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean of 2D ROIs</td>
<td>0.09536 ± 0.01701</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Median of 3D ROIs</td>
<td>0.09281 ± 0.02347</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean of 3D ROIs</td>
<td>0.08871 ± 0.01810</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wavelet-LLL first order median</td>
<td>0.04850 ± 0.01805</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wavelet-HHL GLCM joint energy</td>
<td>0.04764 ± 0.02058</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Platelet count</td>
<td>0.04722 ± 0.01561</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wavelet-HHL GLDM dependence variance</td>
<td>0.04595 ± 0.01581</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wavelet-HHH GLRML low gray level run emphasis</td>
<td>0.04182 ± 0.01147</td>
</tr>
</tbody>
</table>

AUC, area under the receiver operating characteristic curve; ROIs, regions of interest; 2D, two-dimensional; 3D, three-dimensional; APRI, aspartate aminotransferase-to-platelet ratio index; GLCM, gray-level co-occurrence matrix; GLRML, gray-level run length matrix; GLDM, gray-level dependence matrix

Table 4. Pairwise comparisons of areas under the receiver operating characteristic curve using Delong’s test
<table>
<thead>
<tr>
<th>Variables</th>
<th>AUC</th>
<th>95% CI</th>
<th>(P) value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Combination of radiomic, radiologic measures, and clinical features</strong></td>
<td>0.98 ± 0.04</td>
<td>0.951–1</td>
<td>0.960</td>
</tr>
<tr>
<td>Radiomics</td>
<td>0.97 ± 0.07</td>
<td>0.934–1</td>
<td></td>
</tr>
<tr>
<td><strong>Combination of radiomic, radiologic measures, and clinical features</strong></td>
<td>0.98 ± 0.04</td>
<td>0.951–1</td>
<td>0.254</td>
</tr>
<tr>
<td>Radiologic measures</td>
<td>0.96 ± 0.06</td>
<td>0.928–1</td>
<td></td>
</tr>
<tr>
<td><strong>Combination of radiomic, radiologic measures, and clinical features</strong></td>
<td>0.98 ± 0.04</td>
<td>0.951–1</td>
<td>0.011</td>
</tr>
<tr>
<td>Clinical features</td>
<td>0.91 ± 0.11</td>
<td>0.811–0.969</td>
<td></td>
</tr>
<tr>
<td>Radiomics</td>
<td>0.97 ± 0.07</td>
<td>0.934–1</td>
<td>0.694</td>
</tr>
<tr>
<td>Radiologic measures</td>
<td>0.96 ± 0.06</td>
<td>0.928–1</td>
<td></td>
</tr>
<tr>
<td>Radiomics</td>
<td>0.97 ± 0.07</td>
<td>0.934–1</td>
<td>0.039</td>
</tr>
<tr>
<td>Clinical features</td>
<td>0.91 ± 0.11</td>
<td>0.811–0.969</td>
<td></td>
</tr>
<tr>
<td>Radiologic measures</td>
<td>0.96 ± 0.06</td>
<td>0.928–1</td>
<td>0.062</td>
</tr>
<tr>
<td>Clinical features</td>
<td>0.91 ± 0.11</td>
<td>0.811–0.969</td>
<td></td>
</tr>
</tbody>
</table>

AUC, areas under the receiver operating characteristic curve; CI, confidence interval

**Figures**
Patients who underwent MR elastography from November 2017 to May 2020 (n = 190)

Excluded 90 patients
- Unspecified liver disease (n=42)
- Chronic viral hepatitis (HBV, n=11, HCV, n=1)
- None biopsy (n=25)
- ≥ 60 days between MRE and liver biopsy (n = 2)
- Prior major liver surgery (n = 2)
- Alcohol consumption (n = 1)
- Liver donor evaluation (n= 6)

Eligible patients (n = 100)
- Age ≥ 18 years
- No known history of liver disease
- Suspected NAFLD on screening ultrasound and laboratory studies
- ≤ 60 days between MRE and liver biopsy

Low-grade fibrosis (n= 82)
Clinically significant fibrosis (n= 18)

Figure 1
Flow chart depicting patient selection NAFLD, nonalcoholic fatty liver disease
Figure 2

A 36-year-old man with NAFLD without hepatic fibrosis (stage F0; same as the patient in Fig. 1) The work screen shows the completion of the 3D segmentation of the grayscale elastogram with a 95% confidence map using a commercial program. The median value of the whole volume ROI was 2.10 kPa, and the mean value was 2.24 kPa. NAFLD, nonalcoholic fatty liver disease; 3D, three-dimensional; ROI, region of interest.
Figure 3

A 36-year-old man with NAFLD without hepatic fibrosis (Stage F0) (a) The axial grayscale elastogram with a 95% confidence map shows an example of two separate circular ROIs in one slice, avoiding the peripheral portion. The MRE measured liver stiffness at the two sites was approximately 1.57 kPa and 2.43 kPa, respectively. The median value of multiple ROIs was 1.75 kPa, and the mean value was 1.78 kPa, which is appropriate for stage F0. (b) The color scale shows shear stiffness values ranging from 0 to 8 kPa, confirming the lack of significant fibrosis. NAFLD, nonalcoholic fatty liver disease; ROI, region of interest.

Figure 4
Machine learning approach for classifying hepatic fibrosis according to MR elastography, clinical, and radiologic features 3D segmentation was performed using a grayscale elastogram with a 95% confidence map of the MR elastography. 2D and 3D ROI values from radiologic measurements and variable clinical data were also used in the analysis. Feature extraction was performed using 3D segmentation and clinical and radiologic data using PyRadiomics. A random forest regressor was used to extract the features based on their importance. The 20 repeated 10-fold stratified cross-validations verified the stability of the results, and the area under the receiver operating characteristic curve of the classifier was evaluated. 2D, two-dimensional; 3D, three-dimensional; ROI, region of interest.

**Figure 5**

A random forest classifier model ROC curve. This is used for diagnosing significant hepatic fibrosis using radiomics alone, clinical features alone, radiologic measures alone, and all features combined. Use of all combined features shows the highest AUC and accuracy. ROC, receiver operating characteristic; AUC, area under the ROC curve.
A 63-year-old woman with NAFLD with clinically significant hepatic fibrosis (Stage F3) The liver stiffness on MR elastography was as follows: 2D median value, 5.72 kPa; 2D mean value, 5.87 kPa; 3D median value, 5.14 kPa; and 3D mean value, 5.32 kPa. The axial grayscale elastogram with a 95% confidence map (a) and its corresponding RGB image (b) shows three separate circular ROIs in one slice. The ROI of the most fibrotic area (asterisk) was 8.11 kPa, and that of the area just inside the grid was approximately 3–3.22 kPa. The peripheral portion or some area (blue area in RGB image) showing relatively low stiffness was included in the whole volume segmentation process, which may have lowered the overall stiffness value. NAFLD, nonalcoholic fatty liver disease; 2D, two-dimensional; 3D, three-dimensional; ROI, region of interest.