



Original Article



Noninvasive Assessment of Liver Fibrosis and Inflammation in Chronic Hepatitis B: A Dual-task Convolutional Neural Network (DtCNN) Model Based on Ultrasound Shear Wave Elastography

Chengyan Wang^{#1}, Lili Zheng^{#2}, Yan Li³, Shujun Xia⁴, Jun Lv⁵, Xumei Hu¹, Weiwei Zhan⁴, Fuhua Yan³, Ruokun Li^{3*} and Xinping Ren^{4*}

¹Human Phenome Institute, Fudan University, Shanghai, China; ²Ultrasound Department, Ruijin Hospital Wuxi Branch, Shanghai Jiao Tong University School of Medicine, Wuxi, Jiangsu, China; ³Department of Radiology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; ⁴Ultrasound Department, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; ⁵School of Computer and Control Engineering, Yantai University, Yantai, Shandong, China

Received: 4 October 2021 | Revised: 22 February 2022 | Accepted: 3 March 2022 | Published: 29 March 2022

Abstract

Background and Aims: Liver stiffness (LS) measured by shear wave elastography (SWE) is often influenced by hepatic inflammation. The aim was to develop a dual-task convolutional neural network (DtCNN) model for the simultaneous staging of liver fibrosis and inflammation activity using 2D-SWE. **Methods:** A total of 532 patients with chronic hepatitis B (CHB) were included to develop and validate the DtCNN model. An additional 180 consecutive patients between December 2019 and April 2021 were prospectively included for further validation. All patients underwent 2D-SWE examination and serum biomarker assessment. A DtCNN model containing two pathways for the staging of fibrosis and inflammation was used to improve the classification of significant fibrosis ($\geq F2$), advanced fibrosis ($\geq F3$) as well as cirrhosis (F4). **Results:** Both fibrosis and inflammation affected LS measurements by 2D-SWE. The proposed DtCNN performed the best among all the classification models for fibrosis stage [significant fibrosis AUC=0.89 (95% CI: 0.87–0.92), advanced fibrosis AUC=0.87 (95% CI: 0.84–0.90), liver cirrhosis AUC=0.85 (95% CI: 0.81–0.89)]. The DtCNN-based prediction of inflammation activity achieved AUCs of 0.82 (95%

CI: 0.78–0.86) for grade $\geq A1$, 0.88 (95% CI: 0.85–0.90) grade $\geq A2$ and 0.78 (95% CI: 0.75–0.81) for grade $\geq A3$, which were significantly higher than the AUCs of the single-task groups. Similar findings were observed in the prospective study. **Conclusions:** The proposed DtCNN improved diagnostic performance compared with existing fibrosis staging models by including inflammation in the model, which supports its potential clinical application.

Citation of this article: Wang C, Zheng L, Li Y, Xia S, Lv J, Hu X, et al. Noninvasive Assessment of Liver Fibrosis and Inflammation in Chronic Hepatitis B: A Dual-task Convolutional Neural Network (DtCNN) Model Based on Ultrasound Shear Wave Elastography. J Clin Transl Hepatol 2022;10(6):1077–1085. doi: 10.14218/JCTH.2021.00447.

Introduction

Chronic hepatitis B (CHB)-related liver fibrosis has been recognized as a hallmark of progression from mild hepatitis to decompensation manifestations.¹ Inaccurate fibrosis staging may delay antiviral therapy and increase the health-care burden.² It is well known that biopsy is treated as the reference standard for fibrosis staging, but it suffers from problems such as high medical cost, invasiveness, sampling errors, and is often accompanied by complications including infection and hemorrhage.^{3–7} Biomarkers such as aspartate transaminase-to-platelet ratio index (APRI) and fibrosis index based on four factors (FIB-4) remain controversial for the diagnosis of liver fibrosis. Shear wave elastography (SWE) is emerging as a noninvasive tool for fibrosis detection using the biomechanical properties of tissue by applying an external mechanical wave to generate deformation of the liver and then capturing the velocity of shear wave propagation, which is directly associated to stiffness.

Several guidelines have indicated that liver stiffness (LS) assessed by SWE can serve as an alternative to liver biopsy

Keywords: Fibrosis; Inflammation; Shear wave elastography; Chronic hepatitis B; Dual-task convolutional neural network.

Abbreviations: ALT, alanine aminotransferase; APRI, aspartate transaminase-to-platelet ratio index; AST, aspartate transaminase; AUC, area under the curve; CHB, chronic hepatitis B; DtCNN, dual-task convolutional neural network; FIB-4, fibrosis index based on four factors; LS, liver stiffness; LSM, liver stiffness measurement; ROC, receiver operating characteristic; ROI, region of interest; SWE, shear wave elastography.

[#]Contributed equally to this work.

***Correspondence to:** Xinping Ren, Ultrasound Department, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Ruijin 2nd Road, Huangpu District, Shanghai 200020, China. ORCID: <https://orcid.org/0000-0002-7999-4065>. Tel: +86-18930819785, Fax: +86-31265738, E-mail: rxp11946@rjh.com.cn; Ruokun Li, Department of Radiology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Ruijin 2nd Road, Huangpu District, Shanghai 200020, China. ORCID: <https://orcid.org/0000-0002-6929-0013>. Tel: +86-18930819782, Fax: +86-31265443, E-mail: lrk12113@rjh.com.cn

in patients with chronic viral hepatitis.^{8–10} Most of the clinical practice guidelines recommend liver LS measurement for noninvasive fibrosis staging.¹¹ However, there are various potential confounding factors (e.g. liver inflammation), that may influence the LS measurements and lead to false positive LS values.^{12–19} Nakano *et al.* found that the severity of liver fibrosis measured by SWE was influenced by hepatic necroinflammation in chronic hepatitis patients but not in cirrhotic patients.¹³ Ren *et al.* showed that LS on SWE and transient elastography significantly correlated with hepatic inflammation grade on HBV patients who had no or mild (F1) fibrosis.¹⁹ The confounding factors hindered elastography from becoming an ideal alternative to liver biopsy for fibrosis assessment. On the other hand, accurate assessment of liver inflammation in patients who do not have prominent liver fibrosis is important for guiding therapy and improving the prognosis of those patients. Therefore, a model to simultaneously predict the stages of liver fibrosis and inflammation is of clinical significance.

Recently, several studies that applied machine learning algorithms to two-dimensional (2D)-SWE images for chronic liver disease diagnosis reported satisfactory performance.^{20–23} Wang *et al.* described a deep learning radiomics model for assessing liver fibrosis in CHB in a multicenter study that significantly improved diagnostic performance.²⁰ Chen *et al.* employed four existing classification methods (support vector machine, naïve Bayes, random forest, and *k*-nearest neighbors) to build a decision-support system to improve the diagnostic performance for fibrosis staging.²¹ In this study, we developed a dual-task convolutional neural network (DtCNN) for assessment of fibrosis staging and inflammation activity based on 2D-SWE images.

Methods

Patients

The proposed DtCNN model was developed and validated in a retrospective cohort including a total of 966 consecutive CHB patients who underwent liver biopsy and 2D-SWE between March 2015 and November 2018. The study was in conducted following the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the institutional review board of Ruijin Hospital. The requirement for informed consent was waived for this retrospective study. An additional 180 consecutive patients with known chronic liver disease or who were suspected of having chronic liver disease between December 2019 and April 2021 were prospectively included in this study. All patients in the prospective study were asked to provide informed consent. The inclusion criteria were (1) HBsAg seropositivity for more than 6 months; (2) liver fibrosis and inflammation stages indicating liver biopsy assessment; (3) antiviral treatment-naïve patients. The exclusion criteria were (1) the existence of other hepatitis virus coinfection or autoimmune liver disease, (2) the presence of severe extrahepatic diseases, or (3) pregnancy. A total of 532 CHB patients in the retrospective cohort and 180 in the prospective cohort were included in the analysis. The demographic and clinical data of the patients are shown in Table 1.

Serum samples were collected after the patients had fasted overnight (8 h) for measurement of alanine aminotransferase (ALT), aspartate transaminase (AST), platelet, total cholesterol, triglycerides, alkaline phosphatase, serum glucose, insulin levels, and serum uric acid. The two biomarker models²⁴ were:

$$\text{APRI} = \frac{\text{AST/AST upper limit of normal}}{\text{platelet count (10}^9\text{/L)}} \times 100$$

and

$$\text{FIB-4} = \frac{\text{age (years)} \times \text{AST (U/L)}}{\text{platelet count (109/L)}} \times \text{ALT (U/L)}^{1/2}.$$

SWE examination

SWE was performed with an Re7 ultrasound system (Mindray Medical International Co., Ltd, China, SC5-1 probe) by an ultrasonographer with over 15 years of experience in abdominal ultrasonography. The patients fasted for about 4 hours prior to the scan and lay in the supine position with their right arm maximally raised and abducted during scans. The SWE region of interest was acquired at a location deeper than 2 cm from the hepatic capsule to avoid reverberation artifacts and was kept away from large vessels. The patients were asked to hold their breath for approximately 7 seconds after quiet breathing. A rectangular electronic region of interest (ROI) was shown on the best static SWE image, in which a circular Q-Box ROI with a diameter of about 2 cm and free of large vessels was set for analysis. Five consecutive 2D-SWE images were obtained for each patient, and the median value was reported as the LS. Special attention was paid to avoid focal lesions, vessels, biliary tracts, or artifacts from nearby lung gas or cardiac movement. Measurement failure was defined as inability to obtain any color-coded elasticity images after five trials was considered measurement failure.

Histopathological analysis

Liver biopsy was performed within 1 week of the 2D-SWE examination. The biopsies were performed in the right liver lobe with percutaneous ultrasonographic guidance and processed via formalin fixation, paraffin embedding, hematoxylin-eosin staining, and Masson staining. A pathologist with 12 years of experience in hepatobiliary pathology who are blinded to the laboratory results and LS measurements performed the histopathologic analysis of all liver specimens. Fibrosis staging and inflammation activity were assessed by the Scheuer scoring system.²⁵ The stage of fibrosis was assessed as F0, no fibrosis; F1, periportal fibrosis; F2, few fibrotic septa, F3, numerous septa; and F4, cirrhosis. Activity was graded as A0, no activity; A1, mild activity; A2, moderate activity; and A3, severe activity. Staging of \geq F2, \geq F3, and F4 indicated significant fibrosis, advanced fibrosis and cirrhosis, respectively.

DtCNN model

As shown in Figure 1, the DtCNN model contained two pathways to stage fibrosis (upper) and inflammation staging (below). Each pathway included five convolutional and pooling layers and each layer contained a 3×3 convolution kernel and a rectified linear unit (ReLU) activation function. A 3×3 convolution layer with a stride of 2 was connected immediately after the first convolution layer. A batch normalization layer was used after each convolution for faster convergence. Concatenation connections between the two pathways were applied to provide joint features to the two tasks and to improve the prediction performances. Three fully-connected layers were connected after the convolution layers for binary classification. Dual-task learning was implemented by introducing several cross-feature units between the two pathways. An optimal linear combination of hidden feature layers was learned by the cross-feature units. Given two activation maps x_A^i, x_B^i from the first layer

Table 1. Baseline characteristics of the overall study cohort

	Retrospective cohort			<i>p</i> -value	Prospective cohort
	All patients (<i>n</i> =532)	Training cohort (<i>n</i> =372)	Testing cohort (<i>n</i> =160)		All patients (<i>n</i> =180)
Demographics					
Age, years	59.3±11.6	59.3±11.4	60.3±11.8	0.45	58.6±11.4
Male	290 (54.5)	206 (55.4)	84 (52.5)	0.15	153(85.0)
Anthropometry					
Body mass index (kg/m ²)	23.9±3.2	23.6±3.1	24±3.3	0.41	23.8±3.1
Liver biochemistry					
ALT (U/L)	40.4±55.4	41.5±62.4	43.2±64.8	0.23	38.5±48.7
AST (U/L)	43.6±59.1	45.6±67.1	47.6±69.1	0.37	42.2±53.8
Albumin (mg/L)	39.8±6	39.8±5.7	39.6±6.5	0.53	39.8±5.6
Platelets (×10 ⁹ /L)	143.6±62.6	142.7±64	139.8±64.6	0.19	145.4±60.8
Tbil (μmol/L)	19.5±14.1	18.4±7.8	19.3±15.7	0.31	19.6±13.0
Dbil (μmol/L)	3.9±3.1	3.9±2.6	3.7±2.9	0.22	4.1±3.2
Cr (μmol/L)	76.1±17.4	76.5±18.2	76.9±15.7	0.44	76.0±18.0
PT (s)	12.3±1.3	12.1±1.4	12.5±1.3	0.11	12.2±1.3
INR	1.0±0.1	1.0±0.1	1.1±0.1	0.25	1.0±0.1
Scores					
APRI	1.2±2.4	1.3±2.7	1.4±2.8	0.24	1.1±2.1
FIB-4	3.7±3.8	3.9±4.2	4.1±4.2	0.33	3.5±3.6
Fibrosis stage					
F0-1	186 (35.0)	130 (35.0)	56 (35)		22 (12.2)
F2	146 (27.4)	100 (26.9)	46 (28.8)		36 (20)
F3	118 (22.2)	82 (22.0)	36 (22.5)		27 (15.0)
F4	82 (15.4)	60 (16.1)	22 (13.7)		95 (52.8)
Inflammation					
A0	165 (31.0)	115 (30.9)	50 (31.3)		53 (29.4)
A1	173 (32.5)	121 (32.5)	52 (32.5)		88 (48.9)
A2	132 (24.8)	92 (24.7)	40 (25.0)		23 (12.8)
A3	62 (11.7)	44 (11.8)	18 (11.2)		16 (8.9)

Variables are means ± standard deviation or *n* (%). AFP, alpha fetal protein; ALT, alanine aminotransferase; AST, aspartate transaminase; Cr, Creatine; Dbil, direct bilirubin; INR, international normalized ratio; PT, prothrombin time; Tbil, total bilirubin.

for both tasks, in which *A* corresponds to the task of fibrosis staging and *B* corresponds to the task of inflammation staging, a linear combination matrix was learned to connect the first layer and the second layer ($\tilde{x}_A^{ij}, \tilde{x}_B^{ij}$). Specifically, at the location (*i, j*) in the feature map, the linear combination can be parameterized using *a*:

$$\begin{bmatrix} \tilde{x}_A^{ij} \\ \tilde{x}_B^{ij} \end{bmatrix} = \begin{bmatrix} \alpha_{AA} & \alpha_{AB} \\ \alpha_{BA} & \alpha_{BB} \end{bmatrix} \begin{bmatrix} x_A^{ij} \\ x_B^{ij} \end{bmatrix}. \quad (1)$$

The cross-feature unit was inserted into the two pathways after each convolution. That helped to regularize both tasks by enforcing the shared representations by combining both feature maps. By using this network architecture, the DtCNN model can learn an optimal combination of shared and task-specific representations in a supervised way.

The patients were randomly divided into a training cohort and a testing cohort. 2D-SWE images in the training cohort were augmented through flipping, mirroring and rotating to reduce the overfitting of the proposed DtCNN model. All five images acquired from each patient were employed in this model. As the training was performed on a per-image basis, the class with the highest average probability was chosen to be the most likely score from multiple images. Their corresponding histological results (F scores and G scores) were used as the training labels. Three DtCNN models were trained and tested for binary classification of fibrosis stages in different subgroups, i.e. ≥F2 (significant fibrosis), ≥F3 (advanced fibrosis), and F4 (cirrhosis). Similarly, the inflammatory activity was classified as A0, ≥A1, ≥A2 and A3. For comparison, two independent CNN models were tested: the fibrosis model and inflammation model. The diagnostic performances of using LS measurements for binary classifications were also compared.

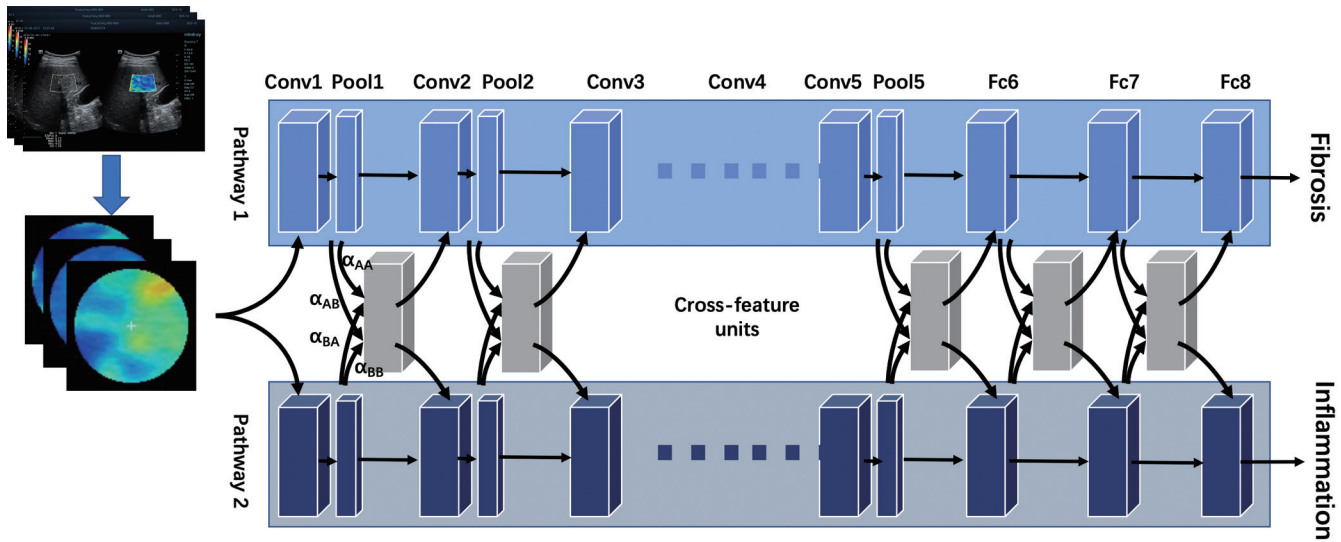


Fig. 1. Illustration of the dual-task convolutional neural network (DtCNN) model for liver fibrosis and inflammation staging. Cross-feature units between the two pathways were applied to provide joint features to the two tasks and to improve the prediction performances.

The DtCNN model was performed on a computer equipped with an Intel® Xeon® Processor E5-2640, 16 GB of memory, and 4 NVIDIA V100 graphic processing units with 16 gigabytes of memory each.

Statistical analysis

Variables were reported as means±SD. Mann-Whitney U-tests and χ^2 -tests were performed for the comparative analysis. Receiver operating characteristic (ROC) curves were plotted to evaluate the discriminating fibrosis stages as $\geq F2$ (significant fibrosis), $\geq F3$ (advanced fibrosis), and F4 (cirrhosis). The area under the curve (AUC), accuracy, sensitivity, specificity, positive predictive value and negative predictive value were calculated to evaluate the performance of the DtCNN model. Comparisons between the AUCs were performed with the DeLong test.²⁶ All statistical analyses were performed using SPSS v.20.0 (IBM Corp., Armonk, NY, USA). *P* value <0.05 was considered statistically significant.

Results

A typical case of liver fibrosis and inflammation is shown in Figure 2. The LS values increased significantly with the progression of liver fibrosis and inflammation (overall *p* < 0.001). The patients with inflammation activities >2 had higher 2D-SWE LS measurement than those < 2. It can be observed that both fibrosis and inflammation influenced the LS measurement, especially for patients with higher fibrosis stages or severe inflammation activities. The diagnostic performance of the DtCNN in classifying liver fibrosis and inflammation using 2D-SWE images in the retrospective cohort is shown in Figure 3 and Table 2. For classification of significant fibrosis ($\geq F2$), the AUC of the liver stiffness measurement (LSM) was 0.78 [95% confidence interval (CI): 0.75–0.82], sensitivity was 0.68, and specificity was 0.81. For classification of advanced fibrosis ($\geq F3$), LSM had an AUC of 0.77 (95% CI: 0.72–0.81), sensitivity of 0.60, and specificity of 0.85. For classification of liver cirrhosis (F4), LSM had an AUC of 0.76 (95% CI:

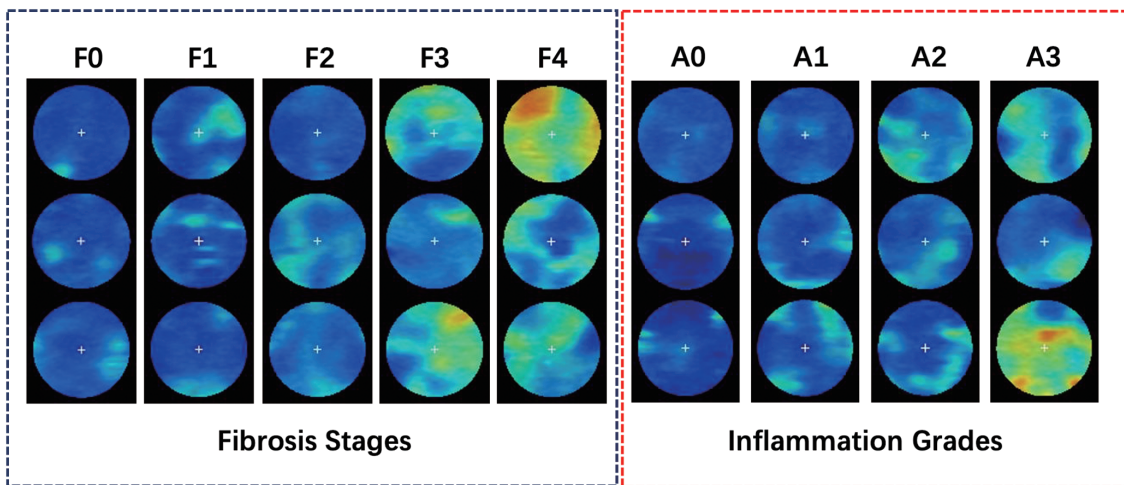


Fig. 2. Typical cases of liver fibrosis and inflammation. It can be observed that both fibrosis and inflammation influence the liver stiffness (LS) measurement.

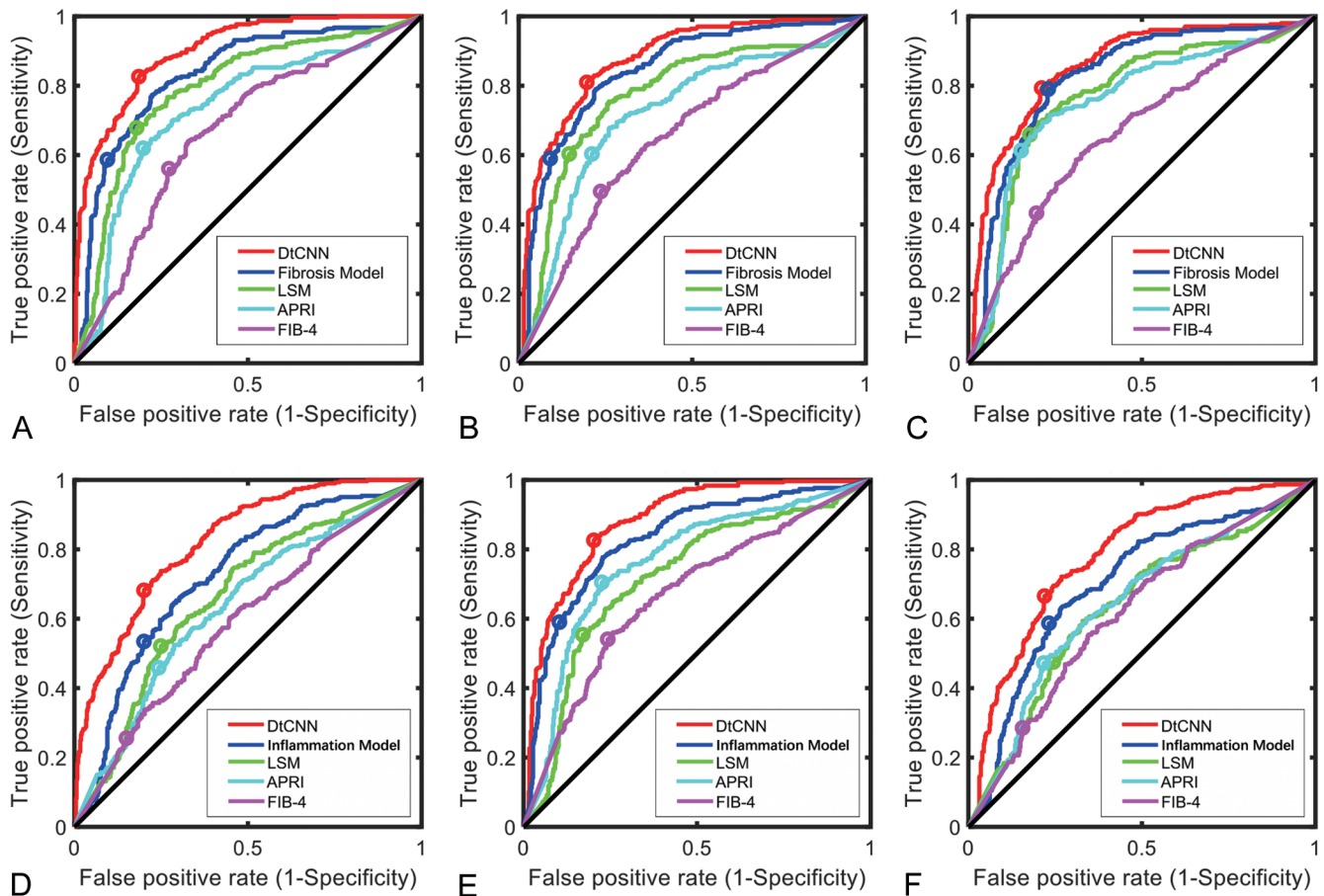


Fig. 3. Comparison of the receiver operating characteristic (ROC) curves of the dual-task convolutional neural network (DtCNN); fibrosis model; inflammation model; LSM; APRI and FIB-4 for the assessment of (A) significant fibrosis ($\geq F2$); (B) advanced fibrosis ($\geq F3$); (C) cirrhosis (F4); (D) $\geq A2$; (E) $\geq A2$; and (F) A3 inflammation using retrospective data.

0.72–0.80), sensitivity of 0.66, and specificity of 0.82. The fibrosis model outperformed the AUCs for different Scheuer fibrosis stages, for significant fibrosis the AUC=0.82 (95% CI: 0.80–0.85), advanced fibrosis AUC=0.84 (95% CI: 0.81–0.87), liver cirrhosis AUC=0.82 (95% CI: 0.80–0.84)). The DtCNN clearly outperformed the other classification models of Scheuer fibrosis stage [(significant fibrosis AUC=0.89 (95% CI: 0.87–0.92), advanced fibrosis AUC=0.87 (95% CI: 0.84–0.90), liver cirrhosis AUC=0.85 (95% CI: 0.81–0.89)]. DtCNN-based inflammation staging had AUCs of 0.82 (95% CI: 0.78–0.86) for $\geq A1$, 0.88 (95% CI: 0.85–0.90) for $\geq A2$, and 0.78 (95% CI: 0.75–0.81) for $\geq A3$, which were significantly higher than the single-task groups (Table 3).

For the prospective dataset, the DtCNN had the best diagnostic performance compared with the other classification methods. The AUCs were 0.88 (95% CI: 0.83–0.91) for significant fibrosis, 0.83 (95% CI: 0.80–0.85) for advanced fibrosis, and 0.88 (95% CI: 0.82–0.93) for liver cirrhosis (Fig. 4 and Supplementary Table 1). DtCNN-based inflammation staging had AUCs of 0.83 (95% CI: 0.79–0.86) for $\geq A1$, 0.88 (95% CI: 0.86–0.91) for $\geq A2$, and 0.77 (95% CI: 0.75–0.79) for $\geq A3$, which were significantly higher than those for the single-task groups (Fig. 4 and Supplementary Table 2). The main findings in the prospective validation dataset were consistent with those obtained with the retrospective data. The dual-task DtCNN model increased the diagnostic performance for both fibrosis and inflamma-

tion staging.

The 2D-SWE results for Scheuer fibrosis and inflammation stages are shown in Figure 5. With the DtCNN model, the diagnostic concordance rates of fibrosis staging were 77.3% for F1, 51.3% for F2, 51.4% for F3, and 55.7% for F4; and the rates for inflammation staging were 58.9% for A0, 60.3% for A1, 48.9% for A2, and 33.0% for A3. When both Scheuer fibrosis and inflammation stages were taken into consideration in the proposed DtCNN model, the discordance was significantly less than that observed with the single-task models.

Discussion

This study explored the influence of liver inflammation on fibrosis staging using 2D-SWE. The dual-task predictive model significantly improved staging accuracy. The DtCNN model achieved better AUCs than those of the single-task (i.e. fibrosis or inflammation) models. Although the diagnostic performance of SWE has been well validated for liver fibrosis staging, the effect of hepatic inflammation on tissue stiffness has not been evaluated with the existence of liver fibrosis. To the best of our knowledge, this is the first deep learning model to assess liver fibrosis and inflammation simultaneously.

The LS measurement-based fibrosis staging outper-

Table 2. Comparison of diagnostic performance of the models in assessment of significant fibrosis, advanced fibrosis, and cirrhosis using retrospective data

Classification	Model	AUC	Accuracy	Sensitivity	Specificity
Significant fibrosis ($\geq F2$)	DtCNN model	0.89	0.81	0.82	0.80
	Fibrosis model	0.82	0.78	0.59	0.90
	LSM	0.78	0.76	0.68	0.81
	APRI	0.72	0.73	0.62	0.80
	FIB-4	0.66	0.66	0.56	0.72
Advanced fibrosis ($\geq F3$)	DtCNN model	0.87	0.80	0.81	0.79
	Fibrosis model	0.84	0.78	0.59	0.90
	LSM	0.77	0.75	0.60	0.85
	APRI	0.72	0.71	0.60	0.78
	FIB-4	0.65	0.66	0.49	0.76
Cirrhosis (F4)	DtCNN Model	0.85	0.78	0.62	0.89
	Fibrosis Model	0.82	0.77	0.79	0.76
	LSM	0.76	0.76	0.66	0.82
	APRI	0.75	0.75	0.61	0.84
	FIB-4	0.66	0.66	0.43	0.80

APRI, aspartate transaminase-to-platelet ratio index; AUC, area under the curve; DtCNN, dual-task convolutional neural network; FIB-4, fibrosis index based on four factors; LSM, liver stiffness measurement.

formed APRI and FIB-4 for all groups; and with the introduction of the deep learning model, the diagnostic performance further improved in all study groups. Several studies have reported that the diagnostic performance of 2D-SWE was better than that of transient elastography and point SWE in assessing liver fibrosis.²⁷⁻²⁹ Although the measurement of stiffness and thresholding is the most widely used method of fibrosis staging, deep learning has advantages because it can extract texture and pattern.¹²⁻¹⁴ Previous studies have reported that the hepatic surface nodularity,³⁰

coarseness of the hepatic parenchyma,³¹ and caudate lobe hypertrophy³² were significant predictors of liver cirrhosis. Deep learning can capture the heterogeneity of intensity and texture from the images.

The assessment of significant and advanced fibrosis was significantly improved by the DtCNN compared with single-task models, which means the bias in 2D-SWE measurement caused by liver inflammation was significant. The findings are consistent with previous studies¹⁶⁻¹⁸ that reported a significant association between the extent of inflamma-

Table 3. Comparison of diagnostic performance of the models in assessment of liver inflammation on retrospective data

Classification	Model	AUC	Accuracy	Sensitivity	Specificity
$\geq A1$	DtCNN model	0.82	0.75	0.69	0.80
	Inflammation Model	0.72	0.70	0.54	0.80
	LSM	0.66	0.67	0.53	0.75
	APRI	0.63	0.65	0.46	0.76
	FIB-4	0.59	0.63	0.26	0.85
$\geq A2$	DtCNN model	0.88	0.80	0.82	0.79
	Inflammation model	0.82	0.77	0.59	0.89
	LSM	0.71	0.72	0.55	0.82
	APRI	0.76	0.74	0.57	0.85
	FIB-4	0.67	0.67	0.53	0.76
A3	DtCNN model	0.78	0.74	0.67	0.78
	Inflammation model	0.70	0.70	0.59	0.76
	LSM	0.63	0.65	0.48	0.76
	APRI	0.64	0.67	0.48	0.78
	FIB-4	0.62	0.63	0.29	0.85

APRI, aspartate transaminase-to-platelet ratio index; AUC, area under the curve; DtCNN, dual-task convolutional neural network; FIB-4, fibrosis index based on four factors; LSM, liver stiffness measurement.

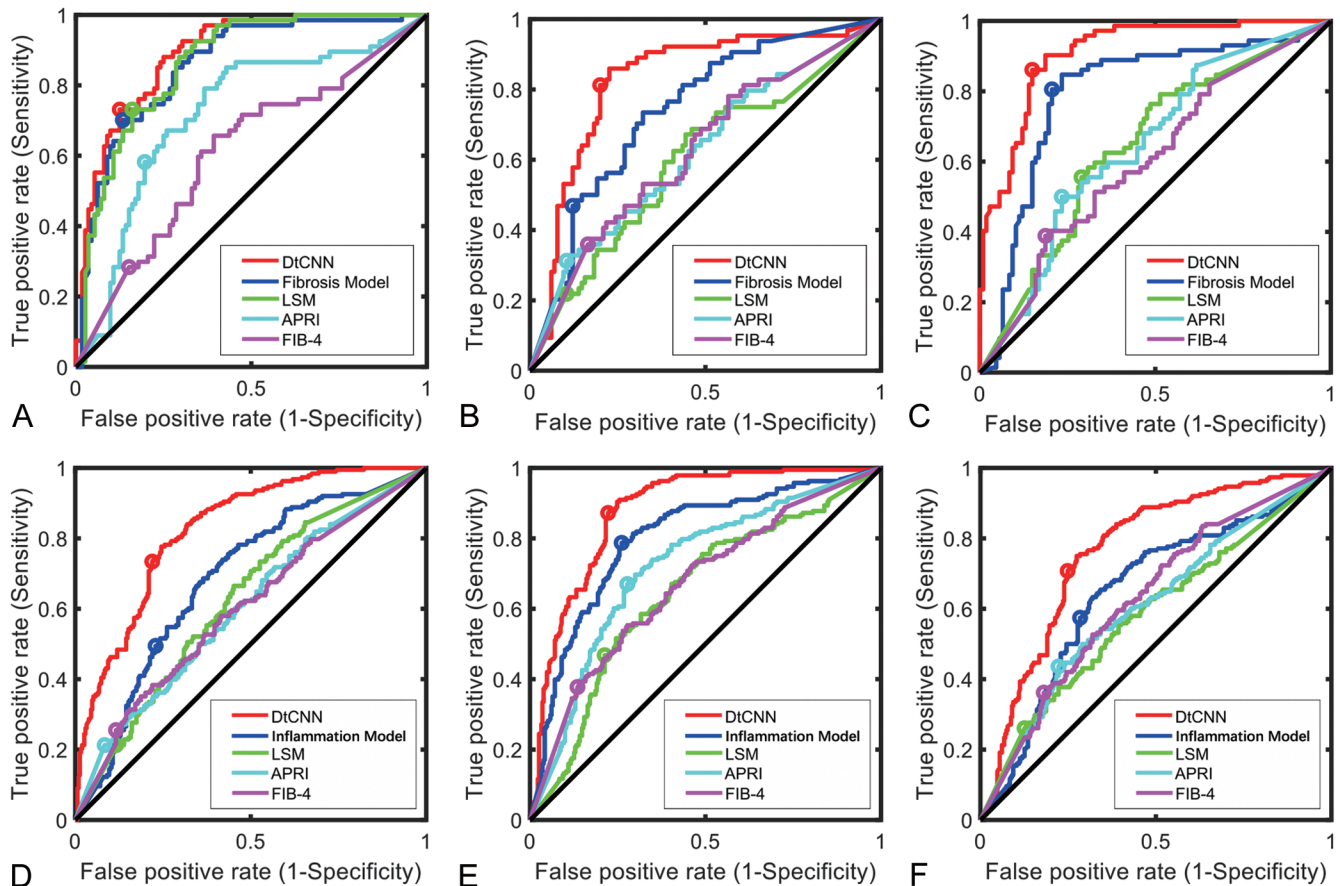


Fig. 4. Comparison of receiver operating characteristic (ROC) curves of dual-task convolutional neural network (DtCNN); fibrosis model; inflammation model; LSM; APRI and FIB-4 for the assessment of (A) significant fibrosis ($\geq F2$); (B) advanced fibrosis ($\geq F3$) and (C) cirrhosis (F4); and (D) $\geq A1$; (E) $\geq A2$ and (F) A3 inflammation using prospective data.

tion and 2D-SWE-based stiffness and that both measurements should be evaluated and considered in the predictive model. It is known that the velocity of shear waves can be affected by inflammation because tissues are composed of viscoelastic materials.³³ However, the extent of inflammation increases the viscosity of liver tissue.³⁴ Our DtCNN model performed two tasks simultaneously and shared the features of the layers for joint learning. It thus overcame the interference of liver inflammation on fibrosis staging in CHB patients, which is essential for elastography diagnosis. In addition to liver inflammation, many other confounding factors for fibrosis staging exist (e.g. steatosis, siderosis, obesity, cholestasis, and ascites). It has been reported that in patients with moderate to severe steatosis, the thicker abdominal walls attenuated ultrasound reflection and affected the LS measurements.³⁵ Fatty liver was associated with a significant decrease in the AUC of 2D-SWE.³⁵ Our DtCNN model can easily be extended to a multitask model that includes more confounding factors in the prediction, thus further improving the diagnostic performance in real scenarios. In addition, CNN models have been reported to perform well in the assessment of liver fibrosis using computed tomography (CT)^{36–38} and magnetic resonance imaging (MRI).^{39,40} With mild modification, the DtCNN model can also be used with those imaging modalities.

Several limitations exist in this study. First, the study was validated only in a single center. It is necessary to test the performance of the DtCNN using US scanners and systems from different manufacturers. Standard algorithms are also

necessary to reduce the variability of US images from different centers. Unlike Wang *et al.*,¹² we developed an automatic tool to select the circular ROI inside the Q-Box as the input of the network, as the quality of those areas were well controlled by the system. This technique can be easily applied to existing US scanners without specific operations for the selection of the scanning planes. Second, the distribution of fibrosis and inflammation stages was uneven, which may bias the training procedure. Resampling algorithms should be introduced into the data augmentation step to improve performance.

Conclusions

The proposed DtCNN improved the diagnostic performance of existing fibrosis staging models by introducing liver inflammation into the model. DtCNN provided more accurate assessments of liver fibrosis and inflammation stages than serum biomarkers in patients with CHB, which supports its potential for clinical application.

Funding

This study was funded by the National Natural Science Foundation of China (No. 62001120) and the Shanghai Sailing Program (No. 20YF1402400).

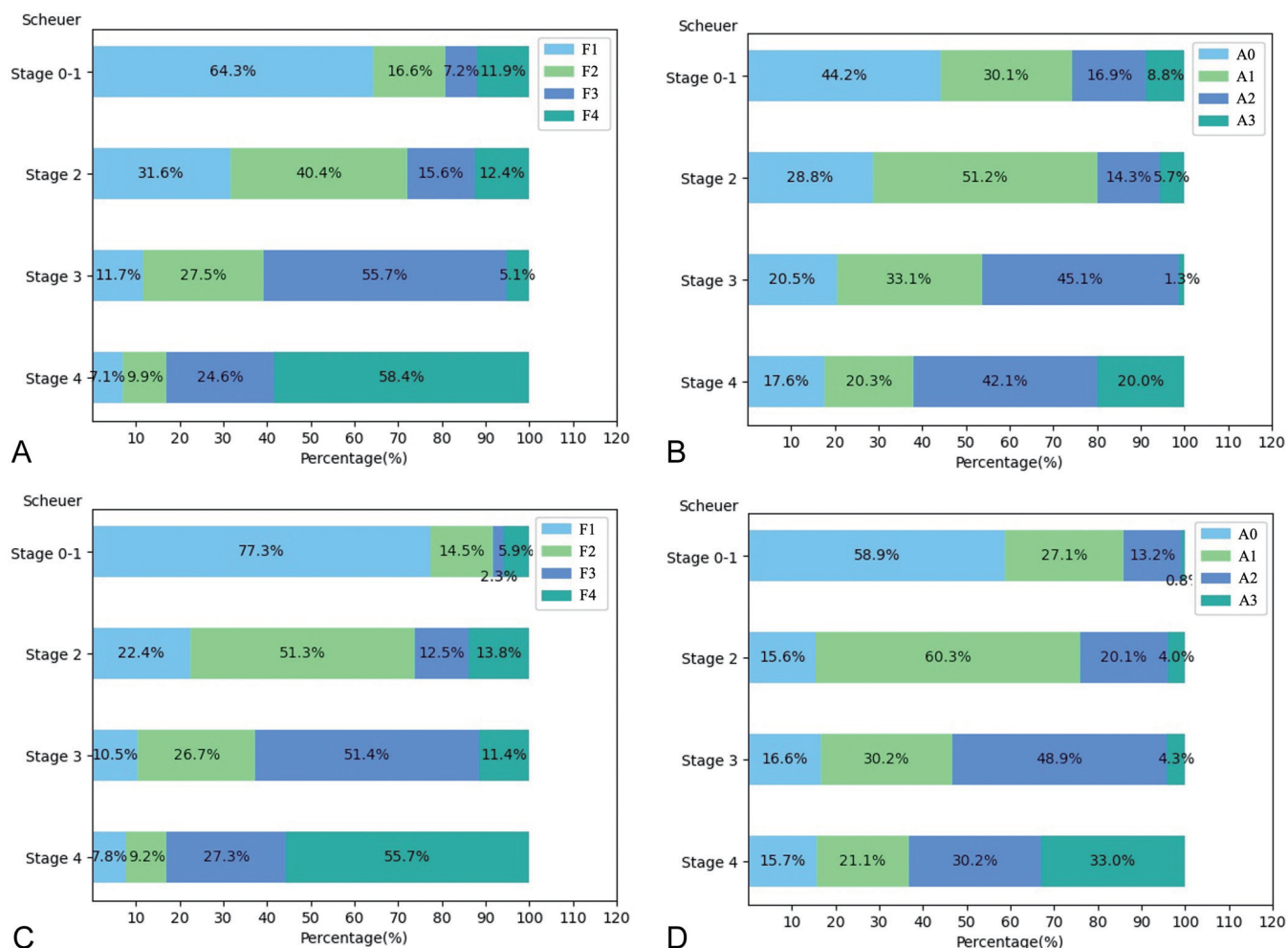


Fig. 5. Distribution of predictions based on 2D-shear wave elastography (SWE) and pathology scores. (A) Single fibrosis model; (B) Single inflammation model; (C) DtCNN-based fibrosis model; (D) DtCNN-based inflammation model.

Conflict of interest

The authors have no conflict of interests related to this publication.

Author contributions

Conceptualization (CW, RL, XR), methodology (CW, LZ, YL), software (JL), validation (CW, SX), formal analysis (CW, XH), investigation (YL), resources (FY, XR), writing-original draft preparation (CW), writing-review and editing (XR, RL), visualization (CW), supervision (RL, XR), project administration (CW, LZ), and funding acquisition (XR). All authors have read and agreed to the published version of the manuscript.

Ethical statement

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of Shanghai Jiao Tong University School of Medicine.

Data sharing statement

The data supporting the findings of this study can be requested from the corresponding author.

References

- [1] Schweitzer A, Horn J, Mikolajczyk RT, Krause G, Ott JJ. Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. *Lancet* 2015;386(10003):1546–1555. doi:10.1016/S0140-6736(15)61412-X, PMID:26231459.
- [2] Wu JF, Song SH, Lee CS, Chen HL, Ni YH, Hsu HY, *et al*. Clinical Predictors of Liver Fibrosis in Patients With Chronic Hepatitis B Virus Infection From Children to Adults. *J Infect Dis* 2018;217(9):1408–1416. doi:10.1093/infdis/jiy048, PMID:29390144.
- [3] Martínez SM, Crespo G, Navasa M, Forns X. Noninvasive assessment of liver fibrosis. *Hepatology* 2011;53(1):325–335. doi:10.1002/hep.24013, PMID:21254180.
- [4] Cadranel JF, Rufat P, Degos F. Practices of liver biopsy in France: results of a prospective nationwide survey. For the Group of Epidemiology of the French Association for the Study of the Liver (AFEF). *Hepatology* 2000;32(3):477–481. doi:10.1053/jhep.2000.16602, PMID:10960438.
- [5] Colloredo G, Guido M, Sonzogni A, Leandro G. Impact of liver biopsy size on histological evaluation of chronic viral hepatitis: the smaller the sample, the milder the disease. *J Hepatol* 2003;39(2):239–244. doi:10.1016/s0168-8278(03)00191-0, PMID:12873821.
- [6] Bedossa P, Dargère D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 2003;38(6):1449–1457. doi:10.1016/j.

- hep.2003.09.022, PMID:14647056.
- [7] Rousselet MC, Michalak S, Dupré F, Croué A, Bedossa P, Saint-André JP, *et al*. Sources of variability in histological scoring of chronic viral hepatitis. *Hepatology* 2005;41(2):257–264. doi:10.1002/hep.20535, PMID:15660389.
- [8] Ferraioli G, Wong VW, Castera L, Berzigotti A, Sporea I, Dietrich CF, *et al*. Liver Ultrasound Elastography: An Update to the World Federation for Ultrasound in Medicine and Biology Guidelines and Recommendations. *Ultrasound Med Biol* 2018;44(12):2419–2440. doi:10.1016/j.ultrasmedbio.2018.07.008, PMID:30209008.
- [9] Dietrich CF, Bamber J, Berzigotti A, Bota S, Cantisani V, Castera L, *et al*. EFSUMB Guidelines and Recommendations on the Clinical Use of Liver Ultrasound Elastography, Update 2017 (Short Version). *Ultraschall Med* 2017;38(4):377–394. doi:10.1055/s-0043-103955, PMID:28407654.
- [10] European Association for Study of Liver; Asociacion Latinoamericana para el Estudio del Hígado. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. *J Hepatol* 2015;63(1):237–264. doi:10.1016/j.jhep.2015.04.006, PMID:25911335.
- [11] Singh S, Muir AJ, Dieterich DT, Falck-Ytter YT. American Gastroenterological Association Institute Technical Review on the Role of Elastography in Chronic Liver Diseases. *Gastroenterology* 2017;152(6):1544–1577. doi:10.1053/j.gastro.2017.03.016, PMID:28442120.
- [12] Ye J, Wang W, Feng S, Huang Y, Liao X, Kuang M, *et al*. Precise fibrosis staging with shear wave elastography in chronic hepatitis B depends on liver inflammation and steatosis. *Hepatol Int* 2020;14(2):190–201. doi:10.1007/s12072-020-10017-1, PMID:32078141.
- [13] Nakano C, Nishimura T, Tada T, Yoshida M, Takashima T, Aizawa N, *et al*. Severity of liver fibrosis using shear wave elastography is influenced by hepatic necroinflammation in chronic hepatitis patients, but not in cirrhotic patients. *Hepatol Res* 2021;51(4):436–444. doi:10.1111/hepr.13617, PMID:33462941.
- [14] Perazzo H, Veloso VG, Grinsztejn B, Hyde C, Castro R. Factors That Could Impact on Liver Fibrosis Staging by Transient Elastography. *Int J Hepatol* 2015;2015:624596. doi:10.1155/2015/624596, PMID:26770833.
- [15] Nishikawa T, Hashimoto S, Kawabe N, Harata M, Nitta Y, Muraio M, *et al*. Factors correlating with acoustic radiation force impulse elastography in chronic hepatitis C. *World J Gastroenterol* 2014;20(5):1289–1297. doi:10.3748/wjg.v20.i5.1289, PMID:24574802.
- [16] Choi M, Kwon H, Cho J, Oh J, Nam K, Kang M, *et al*. Serial change of liver stiffness measured by acoustic radiation force impulse imaging in chronic liver disease: correlation with biochemical markers. *J Med Ultrason* (2001) 2014;41(3):311–317. doi:10.1007/s10396-013-0508-6, PMID:27277904.
- [17] Wu M, Wu L, Jin J, Wang J, Li S, Zeng J, *et al*. Liver Stiffness Measured with Two-dimensional Shear-Wave Elastography Is Predictive of Liver-related Events in Patients with Chronic Liver Disease Due to Hepatitis B Viral Infection. *Radiology* 2020;295(2):353–360. doi:10.1148/radiol.2020191481, PMID:32125254.
- [18] Huang Z, Zheng J, Zeng J, Wang X, Wu T, Zheng R. Normal liver stiffness in healthy adults assessed by real-time shear wave elastography and factors that influence this method. *Ultrasound Med Biol* 2014;40(11):2549–2555. doi:10.1016/j.ultrasmedbio.2014.05.008, PMID:25282481.
- [19] Ren X, Zhang L, Xia S, Zhou J, Lin Y, Zhijie C, *et al*. Diagnostic Performance of Shear Wave Elastography in the Noninvasive Evaluation of Liver Inflammation of Chronic Hepatitis B Patients. *Ultrasound Q* 2021;37(2):111–117. doi:10.1097/RUQ.0000000000000522, PMID:34009924.
- [20] Wang K, Lu X, Zhou H, Gao Y, Zheng J, Tong M, *et al*. Deep learning Radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study. *Gut* 2019;68(4):729–741. doi:10.1136/gutjnl-2018-316204, PMID:29730602.
- [21] Chen Y, Luo Y, Huang W, Hu D, Zheng RQ, Cong SZ, *et al*. Machine-learning-based classification of real-time tissue elastography for hepatic fibrosis in patients with chronic hepatitis B. *Comput Biol Med* 2017;89:18–23. doi:10.1016/j.compbiomed.2017.07.012, PMID:28779596.
- [22] Gatos I, Tsantis S, Spiliopoulos S, Karnabatidis D, Theotokas I, Zoumpoulis P, *et al*. A Machine-Learning Algorithm Toward Color Analysis for Chronic Liver Disease Classification, Employing Ultrasound Shear Wave Elastography. *Ultrasound Med Biol* 2017;43(9):1797–1810. doi:10.1016/j.ultrasmedbio.2017.05.002, PMID:28634041.
- [23] Liu X, Song JL, Wang SH, Zhao JW, Chen YQ. Learning to Diagnose Cirrhosis with Liver Capsule Guided Ultrasound Image Classification. *Sensors (Basel)* 2017;17(1):E149. doi:10.3390/s17010149, PMID:28098774.
- [24] Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS, *et al*. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 2003;38(2):518–526. doi:10.1053/jhep.2003.50346, PMID:12883497.
- [25] Scheuer P. Primary biliary cirrhosis. *Proc R Soc Med* 1967;60:1257–1260.
- [26] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44(3):837–845. PMID:3203132.
- [27] Cassinotto C, Lapuyade B, Mouries A, Hiriart JB, Vergnion J, Gaye D, *et al*. Non-invasive assessment of liver fibrosis with impulse elastography: comparison of Supersonic Shear Imaging with ARFI and FibroScan®. *J Hepatol* 2014;61(3):550–557. doi:10.1016/j.jhep.2014.04.044, PMID:24815876.
- [28] Ferraioli G, Tinelli C, Dal Bello B, Zicchetti M, Filice G, Filice C, *et al*. Accuracy of real-time shear wave elastography for assessing liver fibrosis in chronic hepatitis C: a pilot study. *Hepatology* 2012;56(6):2125–2133. doi:10.1002/hep.25936, PMID:22767302.
- [29] Zhuang Y, Ding H, Zhang Y, Sun H, Xu C, Wang W. Two-dimensional Shear-Wave Elastography Performance in the Noninvasive Evaluation of Liver Fibrosis in Patients with Chronic Hepatitis B: Comparison with Serum Fibrosis Indices. *Radiology* 2017;283(3):873–882. doi:10.1148/radiol.2016160131, PMID:27982760.
- [30] Viganò M, Visentin S, Aghemo A, Rumi MG, Ronchi G. US features of liver surface nodularity as a predictor of severe fibrosis in chronic hepatitis C. *Radiology* 2005;234(2):641. doi:10.1148/radiol.2342041267, PMID:15671013.
- [31] Lee CH, Choi JW, Kim KA, Seo TS, Lee JM, Park CM. Usefulness of standard deviation on the histogram of ultrasound as a quantitative value for hepatic parenchymal echo texture; preliminary study. *Ultrasound Med Biol* 2006;32(12):1817–1826. doi:10.1016/j.ultrasmedbio.2006.06.014, PMID:17169693.
- [32] Colli A, Fraquelli M, Andreoletti M, Marino B, Zucconi E, Conte D. Severe liver fibrosis or cirrhosis: accuracy of US for detection—analysis of 300 cases. *Radiology* 2003;227(1):89–94. doi:10.1148/radiol.2272020193, PMID:12601199.
- [33] Deffieux T, Montaldo G, Tanter M, Fink M. Shear wave spectroscopy for in vivo quantification of human soft tissues visco-elasticity. *IEEE Trans Med Imaging* 2009;28(3):313–322. doi:10.1109/TMI.2008.925077, PMID:19244004.
- [34] Sugimoto K, Moriyasu F, Oshiro H, Takeuchi H, Yoshimasu Y, Kasai Y, *et al*. Clinical utilization of shear wave dispersion imaging in diffuse liver disease. *Ultrasonography* 2020;39(1):3–10. doi:10.14366/usg.19031, PMID:31645092.
- [35] Sogabe M, Okahisa T, Hibino S, Yamanoi A. Usefulness of differentiating metabolic syndrome into visceral fat type and subcutaneous fat type using ultrasonography in Japanese males. *J Gastroenterol* 2012;47(3):293–299. doi:10.1007/s00535-011-0489-4, PMID:22065161.
- [36] Hectors SJ, Kennedy P, Huang KH, Stocker D, Carbonell G, Greenspan H, *et al*. Fully automated prediction of liver fibrosis using deep learning analysis of gadoxetic acid-enhanced MRI. *Eur Radiol* 2021;31(6):3805–3814. doi:10.1007/s00330-020-07475-4, PMID:33201285.
- [37] Choi KJ, Jang JK, Lee SS, Sung YS, Shim WH, Kim HS, *et al*. Development and Validation of a Deep Learning System for Staging Liver Fibrosis by Using Contrast Agent-enhanced CT Images in the Liver. *Radiology* 2018;289(3):688–697. doi:10.1148/radiol.2018180763, PMID:30179104.
- [38] Yasaka K, Akai H, Kunimatsu A, Abe O, Kiryu S. Deep learning for staging liver fibrosis on CT: a pilot study. *Eur Radiol* 2018;28(11):4578–4585. doi:10.1007/s00330-018-5499-7, PMID:29761358.
- [39] Yasaka K, Akai H, Kunimatsu A, Abe O, Kiryu S. Liver Fibrosis: Deep Convolutional Neural Network for Staging by Using Gadaxetic Acid-enhanced Hepatobiliary Phase MR Images. *Radiology* 2018;287(1):146–155. doi:10.1148/radiol.2017171928, PMID:29239710.
- [40] Zheng R, Shi C, Wang C, Shi N, Qiu T, Chen W, *et al*. Imaging-Based Staging of Hepatic Fibrosis in Patients with Hepatitis B: A Dynamic Radiomics Model Based on Gd-EOB-DTPA-Enhanced MRI. *Biomolecules* 2021;11(2):307. doi:10.3390/biom11020307, PMID:33670596.