Artificial intelligence for hepatitis evaluation

Wei Liu, Xue Liu, Mei Peng, Gong-Quan Chen, Peng-Hua Liu, Xin-Wu Cui, Fan Jiang, Christoph F Dietrich

Abstract

Recently, increasing attention has been paid to the application of artificial intelligence (AI) to the diagnosis of diverse hepatic diseases, which comprises traditional machine learning and deep learning. Recent studies have shown the possible value of AI based data mining in predicting the incidence of hepatitis, classifying the different stages of hepatitis, diagnosing or screening for hepatitis, forecasting the progression of hepatitis, and predicting response to antiviral drugs in chronic hepatitis C patients. More importantly, AI based on radiology has been proven to be useful in predicting hepatitis and liver fibrosis as well as grading hepatocellular carcinoma (HCC) and differentiating it from benign liver tumors. It can predict the risk of vascular invasion of HCC, the risk of hepatic encephalopathy secondary to hepatitis B related cirrhosis, and the risk of liver failure after hepatectomy in HCC patients. In this review, we summarize the application of AI in hepatitis, and identify the challenges and future perspectives.

Key Words: Machine learning; Deep learning; Radiomics; Hepatitis; Fibrosis; Hepatocellular carcinoma

©The Author(s) 2021. Published by Baishideng Publishing Group Inc. All rights reserved.
Artificial intelligence (AI) is a new discipline that aims to simulate, extend, and expand human intelligence and integrates theory, method, and application research and development[1]. According to the different algorithms of AI, AI is usually divided into traditional AI (e.g., traditional machine learning) and deep learning (based on neural network structure) in the medical field[2]. Compared with traditional AI, deep learning can directly apply the image to the learning process without manual feature extraction, while traditional machine learning needs manual recognition and extraction of different features. Therefore, deep learning has the advantage of no manual extraction of various features, which makes its learning process faster, intelligent, and accurate. In addition, deep learning can iterate and improve from past mistakes, but it needs more big data and more result analysis to fully show its robust and precise efficiency[3].

Lately, AI technology has been at the forefront of scientific research. The benefits of the AI model in various fields have triggered an upsurge in the use of this technology for data mining and analysis in more areas, and it has also attracted attention in the field of medicine and biological cognition[4]. For instance, AI has been widely used in the examination of superficial organs such as the thyroid and breast to assist radiologists to evaluate whether nodules are benign or malignant, with a high level of accuracy. The accuracy is lower than that of senior radiologists but higher than that of junior radiologists[5-7]. Some literature has reported the application of AI technology based on data mining or radiology in predicting hepatitis, evaluating fibrosis, and diagnosing benign and malignant liver masses[8-10].

This systematic review is intended to ascertain the clinical utility in routine practice of AI based on data mining, predict the incidence of hepatitis A, B, C, or E, categorize the different stages of hepatitis B, predict the progression of hepatitis C in veterans, and diagnose or screen for hepatitis B. Particularly, this review focuses on the application of deep learning or radiomics based on radiology to stage hepatic fibrosis, and diagnose or screen for hepatitis B. Especially, this review focuses on the application of AI in hepatitis, evaluating fibrosis, and diagnosing benign and malignant liver masses.

Hepatitis Detection Based on Data Mining

Predicting Incidence of Hepatitis

Newborn babies are required to be vaccinated with hepatitis B vaccine in numerous countries, which has reduced the incidence of hepatitis worldwide. However, viral hepatitis is still a serious health problem[11]. Hence, Guan et al[12] designed the AI model of artificial neural network (ANN) and autoregressive integrated moving average (ARIMA) with hepatitis data mined from Liaoning Disease Control and Prevention Center, which aimed to forecast the incidence of hepatitis A. They extracted data concerning the incidence of hepatitis A from 1987 to 2003 in the above-mentioned Disease Control and Prevention Center. The incidence of hepatitis A from

Core Tip: Currently, there is a need of a comprehensive review to introduce the applications of artificial intelligence (AI) in hepatitis. We first discuss the possible value of AI based data mining in predicting the incidence of hepatitis, classifying the different stages of hepatitis, diagnosing or screening for hepatitis, forecasting the progression of hepatitis, and predicting response to antiviral drugs in chronic hepatitis C patients. In addition, we introduce the applications of AI based on radiology in predicting hepatitis and liver fibrosis, grading hepatocellular carcinoma (HCC) and differentiating it from benign liver tumors, and predicting the risk of vascular invasion of HCC as well as the risk of hepatic encephalopathy secondary to hepatitis B related cirrhosis.
1981 to 1997 was taken as the training group and from 1998 to 2001 as the validation group. The forecasting statistical effect value and the correlation coefficient are summarized in Table 1. Intelligently mining hepatitis B-related medical record data from local health websites authorized by the local Health Commission, the AI model of ARIMA (0, 1, 1) and ElmanNN (Elman neural network) with eight neurons were established by Zheng et al[13]. From January 2012 to August 2019, the local health bureau reported 486983 cases of hepatitis B. Hepatitis B incidence from January 2012 to December 2018 was used to build and train ARIMA (0, 1, 1) model and ElmanNN model with eight neurons, and hepatitis B incidence from January 2019 to August 2019 was used to validate ARIMA (0, 1, 1) model and ElmanNN with eight neurons, with root-mean-square error (RMSE) and mean absolute error (MAE) applied to evaluate the prediction effect of the model. The RMSE and MAE of the above-mentioned ARIMA (0, 1, 1) model and ElmanNN with eight neurons are shown in Table 1. Gan et al[14] also tried to forecast the incidence of hepatitis B using a hybrid model combining Grey system and BP-ANN (back propagation artificial neural networks), with trained and validated data collected from National Health Commission. From 2003 to 2012, 10486959 cases of hepatitis B were reported by National Health Commission. Compared with the above data, the prediction effect of a hybrid model, Grey model (1, 1) and Grey model (2, 1), was calculated. The predictive results are summarized in Table 1. The results showed that the hybrid algorithm was superior to the Grey model algorithm in all assessment norms, and a better prediction effect was achieved. Intending to predict the incidence of hepatitis E, Guo et al[15] mined monthly incidence and cases of hepatitis E from the local Center for Disease Control and Prevention as samples. All above-mentioned data were collected from January 2005 to December 2017 in the local Centers for Disease Control. The monthly incidence and cases from January 2005 to June 2015 were used as the input variable to establish an AI model of the ARIMA, supporter vector machine (SVM), and long-short time memory neural network (LSTM), and those from July 2015 to December 2017 as the validating group. The assessment indexes of the ARIMA, SVM, and LSTM models were RMSE and MAE, whose corresponding values are given in Table 1.

**Classifying different stages of hepatitis**

According to the serological level of transaminase and serological titer of hepatitis B antigen or antibody, hepatitis B can be categorized into five stages: I, HBeAg (+) chronic infection; II, HBeAg (+) chronic hepatitis; III, HBeAg (-) chronic infection; IV, HBeAg (-) chronic hepatitis; V, HBsAg (-) stage. For young doctors, it can be difficult to accurately classify the clinical types of hepatitis. Therefore, with 52 patients included, an automated diagnosis of hepatitis B using multilayer mamdani fuzzy inference system expert system (ADHB-ML-MFIS) was established by Ahmad et al[16] to categorize the different stages of hepatitis B. In the ADHB-ML-MFIS expert system, there were two levels of input variables. The input variables of the first layer were transaminase, and the input variables of the second layer were serological markers of hepatitis B. The diagnostic accuracy of the ADHB-ML-MFIS expert system for differentiating stages of hepatitis B was 92.2% with the detailed results shown in Table 1.

**Diagnosing or screening for hepatitis**

Raman spectroscopy has been used to diagnose hepatitis B, showing a superior performance compared to traditional serological tests. Obtaining 119 confirmed hepatitis B virus (HBV) infected samples from histopathology department, Khan et al [8] combined Raman spectroscopy with pattern recognition technology, and established an AI model based on the SVM algorithm and radial basis function to diagnose hepatitis B. In the study, the AI model achieved high diagnostic performance for hepatitis B, with the data presented in Table 1. Collecting data from 1134 blood samples, Wang et al[17] proposed a method for promptly screening hepatitis B patients and non-hepatitis B patients using serum Raman spectroscopy combined with LSTM. Then, LSTM was used to train the spectral data. Several models can distinguish hepatitis B from non-hepatitis B, with the highest diagnostic efficacy acquired by serum Raman spectroscopy combined with LSTM. The accuracy of serum Raman spectroscopy combined with LSTM for screening hepatitis B from non-hepatitis B was 0.9732. About 50% of hepatitis C virus (HCV) infected persons worldwide have not been diagnosed or treated, and it remains a significant risk factor for human health [18]. Thus, it is vital that clinicians could find these undiagnosed patients with hepatitis C infection, a nowadays-treatable disease. With 120023 HCV patients and 9601900 non-HCV patients used as modeling data, Doyle et al[19] developed five kinds of AI models attempting to find undiagnosed hepatitis C infection: Logistic regression, gradient boosting trees, gradient boosting trees with temporal variables, stacked...
Table 1 Hepatitis detection based on data mining

| No. | Task | Algorithms | Sample size (type) | Evaluation index | Ref. |
|-----|------|------------|-------------------|------------------|------|
| 1   | Predicting incidence of hepatitis A | ANN; ARIMA | N/A (CDC data) | ANN: Correlation coefficient 0.71; ARIMA: Correlation coefficient 0.66 | [12] |
| 2   | Predicting incidence of hepatitis B | ARIMA; ElmanNN | 486983 cases (data from health commission) | ARIMA: RMSE 0.94, MAE 0.81; ElmanNN: RMSE 0.89, MAE 0.70 | [13] |
| 3   | Forecasting incidence of hepatitis B | Hybrid method (combining GM and BP-ANN) | 10486959 cases (data from health ministry) | R 0.9495, RMSE 4.863 x 10^3, MAE 3.9704 x 10^4 | [14] |
| 4   | Prediction of incidence of hepatitis E | ARIMA; SVM; LSTM | N/A (CDC data) | ARIMA: RMSE 0.022, MAE 0.018; SVM: RMSE 0.0204, MAE 0.0167; LSTM: RMSE 0.01, MAE 0.011 | [15] |
| 5   | Automated classification of the different stages of hepatitis B | ADHB-ML-MFIS expert system | 52 patients (serological data) | Overall accuracy: 0.922; No hepatitis accuracy: 1; Due to infection accuracy: 0.75; Acute HBV accuracy: 0.95; Chronic HBV accuracy: 0.91 | [16] |
| 6   | Analyzing HBV infection from normal blood samples | Polynomial function; RBF | 119 serum samples from HBV infected patients (Raman spectroscopy data) | Polynomial kernel (order-2): Quadratic programming/least squares; Accuracy 98%, precision 97%, sensitivity 100%, specificity 95%; RBF kernel (RBF sigma-2): Quadratic programming; accuracy 94%, precision 90%, sensitivity 100%, specificity 87%; RBF kernel (RBF sigma-2): Least squares: Accuracy 95%, precision 92%, sensitivity 100%, specificity 90% | [8] |
| 7   | Rapidly screening hepatitis B from non-hepatitis B | LSTM | 1134 blood samples (Raman spectroscopy data) | Accuracy 97.32%, sensitivity 97.87%, specificity 96.77%, precision 96.84% | [17] |
| 8   | Finding undiagnosed patients with hepatitis C infection | Logistic regression; Gradient boosting trees; Gradient boosting trees with temporal variables; Stacked ensemble; Random forest | 9721923 patients (data from the patient’s medical history) | The stacked ensemble had a specificity of 0.99 and precision of 0.97 at a recall level of 0.50 | [19] |
| 9   | Predicting hepatitis C virus progression among veterans | CS Cox model; longitudinal Cox model; CS boosting model; Longitudinal-boosting model | 72683 CHC individuals (VHA data) | CS Cox model: Concordance 0.746; Longitudinal Cox model: Concordance 0.758; Longitudinal-boosting model: Concordance 0.774 | [20] |
| 10  | Forecasting response to IFN plus RIB treatment in HCV patients | ANN | 300 patients (serological data) | The diagnostic accuracy rose from 52% (ANN 2) to 70% (ANN 6) | [21] |

ANN: Artificial neural network; ARIMA: Autoregressive integrated moving average; CDC: Centers for Disease Control; RMSE: Root-mean-square error; MAE: Mean absolute error; GM: Grey model, BP-ANN: Back propagation artificial neural networks; SVM: Supporter vector machine; LSTM: Long-short time memory neural network; ADHB-ML-MFIS: Automated diagnosis of hepatitis B using multilayer Mamdani fuzzy inference; HBV: Hepatitis B virus; RBF: Gaussian radial basis function; CHC: Chronic hepatitis C virus; VHA: National Veterans Health Administration; IFN: Interferon; RIB: Ribavirin, HCV: Hepatitis C virus.

ensemble, and random forest. All these models mined data from demographics, risk factors, symptoms, treatments, and procedures relevant to HCV from the patient’s medical history. At different recall levels, the models had different diagnostic performances. For example, the stacked ensemble had a specificity of 0.99 and precision of 0.97 at a recall level of 0.50.

**Forecasting progression of hepatitis**

Liver biopsy is the gold standard for the evaluation of liver fibrosis and cirrhosis. However, liver biopsy is an invasive examination and carries the risk of causing liver bleeding. In clinical practice, aspartate aminotransferase-to-platelet ratio index (APRI) and radiology are often used to forecast hepatic fibrosis and cirrhosis. In Konerman et al[20]’s study, aiming to predict HCV progression among veterans, they developed the cross-sectional (CS) Cox model, longitudinal Cox model, CS boosting model, and longitudinal-boosting model with 72683 chronic hepatitis C (CHC) individuals and matched control data searched from National Veterans Health Administration between 2000 and 2016. The calculated value of APRI > 2 after time zero for two consecutive times indicated the progression of liver cirrhosis. The predictive
concordance of the above-mentioned four models is given in Table 1.

**Predicting response to antiviral drugs in CHC patients**

The combined administration of interferon-alpha (IFN) and ribavirin (RIB) is the first-line recommended treatment regimen for patients with CHC. However, the cure rate is still low and the incidence of side effects is high. Therefore, there is an urgent need for doctors in the hepatology or infection departments to evaluate the response of patients with CHC to the combinative administration of IFN and RIB. Hence, Maiellaro et al[21] retrospectively analyzed 300 patients treated with IFN plus RIB through an ANN, intending to predict the response to the treatment. The range of positive predictive value and negative predictive value of the ANN model for predicting the treatment response of patients with CHC to the combined administration of IFN and RIB was 57%-75% and 52%-71%, respectively.

**HEPATITIS EVALUATION BASED ON RADIOLOGY**

Radiology plays an important role in evaluating a range of hepatic diseases, such as hepatitis, hepatic fibrosis, HCC, vascular invasion of liver cancer, and even HE. With the rapid development of computer technology and AI, radiomics and deep learning have become hot topics in the field of medical imaging. Compared with traditional AI, deep learning is a more advanced and intelligent learning algorithm, which was based on a neural network structure inspired by the human brain[22,23]. Radiomics has also been developed recently. It can adopt a variety of methods to extract numerous quantitative features from ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI) images, then select the quantitative features related to the task, finally providing clinicians with a model based on radiomics to evaluate diagnostic and prognostic information[24,25].

**Predicting clinical severity in alcohol-associated hepatitis patients**

Tana et al[26] included 34 patients with alcohol-associated hepatitis (AAH) and 35 control subjects in their study to extract texture features from computed tomography images based on random forest and deep learning convolutional neural network algorithms, which evaluated the diagnostic value of CT texture analysis for clinical severity (most notably aspartate aminotransferase) in AAH patients. Recursive feature elimination using random forest identified 23 top features for AAH categorizing, and the accuracy of deep learning for diagnosing AAH was 70% in the validation set.

**Diagnosing or grading liver fibrosis**

According to the METAVIR scoring system[27], hepatic fibrosis can be divided into five categories: F0, no fibrosis; F1, portal fibrosis without septa; F2, portal fibrosis and few septa; F3, numerous septa without cirrhosis; F4, cirrhosis. Significant fibrosis is defined as ≥ F2. Based on radiomics or deep learning algorithms, multifarious imaging modalities have been applied to diagnose or stage liver fibrosis, such as ultrasound, contrast or non-contrast CT, CT texture, MRI, or gadoxetic acid–enhanced hepatobiliary phase MR.

A study including 144 HBV infected patients, evaluated the performance of multiparametric ultrasonics for the diagnosis of significant hepatic fibrosis[9]. Conventional ultrasonics, original radiofrequency, and contrast-enhanced micro-flow features comprise the multiparametric ultrasonics for clinicians to diagnose diseases and monitor treatment regimens. Multiparametric ultrasonics used Adaboost, random forest, and SVM algorithms for predicting significant liver fibrosis, with results summarized in Table 2. With 466 patients (401 with chronic hepatitis B, 65 without fibrosis) undergoing partial hepatectomy used as the data of the training group and test group of transfer learning (TL) radiomics, Xue et al[28] discussed the diagnostic value of transfer learning radiomics based on multimodal ultrasound in the grading of liver fibrosis. As the name suggests, TL moves a network trained on a large data set on to other related tasks, thus avoiding the over fitting problem caused by insufficient training data in conventional deep learning. Results showed that the area under the curve (AUC) of TL was statistically higher than that of non-TL, and the AUC of multimodal ultrasound imaging was larger than that of single modality ultrasound imaging. Hence, multimodal ultrasound imaging based on TL could usefully be applied to the staging of liver fibrosis.
| No. | Task                                                                 | Algorithms (model)                              | Sample size (type) | Evaluation index                                                                 | Ref. |
|-----|----------------------------------------------------------------------|------------------------------------------------|--------------------|----------------------------------------------------------------------------------|------|
| 1   | Predicting clinical severity in AAH patients                         | Random forest; Convolutional neural network     | 69 cases (CT texture features) | Accuracy: 82.4% of RFE-RF in the test set; Accuracy: 70% of CNN in the test set  | [29] |
| 2   | Assessing significant liver fibrosis by multiparametric ultrasomics data | Adaboost; Random forest; SVM (multiparametric ultrasomics) | 144 HBV infected patients (multiparametric ultrasomics) | AUROC: 0.85 ± 0.01 of Adaboost, random forest, SVM in multiparametric ultrasomics including conventional ultrasomics, ORF and CEMF | [9]  |
| 3   | Grading liver fibrosis                                              | Inception-V3 network (transfer learning)       | 466 patients (multimodal ultrasound) | AUCs of TL in GM + EM reached 0.950, 0.952, and 0.950, respectively, for grading S4, ≥ S3, and ≥ S2an | [28] |
| 4   | Predicting cirrhosis                                                | LASSO (radiomics nomogram)                     | 144 cases of HBV patients (CT features and clinical factors) | AUROC: 0.915 in the training cohort, 0.872 in the validation cohort, overall correctly classified rate of 82.0% | [29] |
| 5   | Differentiating hepatic fibrosis’ grade                              | RFC (CTTA-based models); SVM (CTTA-based models) | 30 fibrosis patients (CT texture features) | Train AUC 0.95 in RFC (model 1); Test AUC 0.90 in RFC (model 1); Train AUC 0.88 in SVM (model 2); Test AUC 0.76 in SVM (model 2) | [30] |
| 6   | Assessing liver fibrosis severity                                   | A prototype convolutional neural network        | 558 cases (CT images) | AUCs were 0.82, 0.85, and 0.88 of Vol/L, VolS in diagnosing advanced fibrosis, cirrhosis, and decompensated cirrhosis in the whole study population | [31] |
| 7   | Staging liver fibrosis                                              | Convolutional neural network                   | 634 fibrosis patients (MR images and MR/virus) | AUCs were 0.84, 0.84, and 0.85 of the model full for diagnosing F4, ≥ F3, and ≥ F2 in the test set, respectively | [34] |
| 8   | Assessing liver fibrosis in chronic hepatitis B                     | Convolutional neural network (DLRE)             | 398 HBV patients (shear wave elastography) | AUROC of DLRE 1.00, 0.99, and 0.99 for classifying F4, ≥ F3, and ≥ F2 in the training set and 0.97, 0.98, and 0.85 in the validation set | [35] |
| 9   | Diagnosing FNH from HCC in the non-cirrhotic liver                   | LASSO (radiomics nomogram)                     | 156 patients (CT images and clinical factors) | Accuracy: 92.4% in the training set, 89.2% in the validation set | [38] |
| 10  | Diagnosing HCC                                                      | LASSO (radiomics signature)                    | 211 patients (MR images) | AUROC: 0.861 in the training set, 0.810 in the validation set | [39] |
| 11  | Preoperative prediction of HCC grade                                 | LASSO (combined model with clinical factors and radiomics signature) | 170 HCC patients (MR images and clinical factors) | AUROC: 0.742, 0.786, and 0.800 based on T1WI images, T2WI images, and combined T1WI and T2WI images in the combined model | [41] |
| 12  | Predicting MVI risk in HBV-related HCC preoperatively               | LASSO (radiomics nomogram)                     | 304 HCC patients (CT images and AFP) | AUROC: 0.846 in the training set, 0.844 in the validation set | [43] |
| 13  | Preoperative prediction of MVI in HCC patients                      | LASSO (combined model)                         | 157 HCC patients (CT images and clinical factors) | AUROC: 0.835 in the training dataset, 0.801 in the validation dataset | [44] |
| 14  | Predicting risk of HE complicated by hepatitis B related cirrhosis | LASSO (integrated model of radiomics and clinical factors) | 304 cirrhosis patients (CT images and clinical factors) | Accuracy: 0.93 in the training cohort, 0.83 in the testing cohort | [45] |
| 15  | Predicting liver failure in cirrhotic patients with HCC after major heptatectomy | LASSO (integrated radiomics-based mode)         | 101 HCC patients (MR images and clinical factors) | Accuracy: 0.802 in radiomics-based model | [47] |

AAH: Alcohol-associated hepatitis; CT: Computed tomography; RFE-RF: Recursive feature elimination using random forest; CNN: Convolutional neural network; SVM: Supporter vector machine; HBV: Hepatitis B virus; AUROC: Area under the receiver operating curve; ORF: Original radiofrequency; CEMF: Contrast-enhanced micro-flow; AUC: Area under curve; TL: Transfer learning; GM: Gray scale modality; EM: Elastogram modality; LASSO: Least absolute shrinkage and selection operator; RFC: Random forest classifier; CTTA: Computed tomography texture analysis; VolL: Liver volume; VolS: Spleen volume; MR: Magnetic resonance; DLRE: Deep learning radiomics of shear wave elastography; FNH: Focal nodular hyperplasia; HCC: Hepatocellular carcinoma; MVI: Microvascular invasion; AFP: Alpha-fetoprotein; HE: Hepatic encephalopathy.

In another study, Wang et al.[29] included 144 HBV patients with hepatic fibrosis.
confirmed by liver biopsy who underwent unenhanced CT examinations in their study and successfully extracted 25 cirrhosis-related CT features to construct a radiomics signature using the SVM algorithm. In addition to the CT features that were closely related to liver cirrhosis, the specific clinical factors (alanine transaminase, aspartate aminotransferase, globulin, and international normalized ratio) were also closely related to liver cirrhosis. Combining the radiomics signature and clinical factors, a radiomics-based nomogram was established, and the results demonstrated that the area under the receiver operating curve (AUROC) of the radiomics nomogram was 0.915 and 0.872 in the training and validation cohorts, respectively. CT texture analysis (CTTA) revealed the biological and pathological features of diverse liver diseases by analyzing the characteristics and distribution of pixels in medical images, and extracted texture features that were difficult to recognize by the naked eye. To explore the diagnostic value of CTTA in low- and high-grade hepatic fibrosis, Budai et al.[30] included 30 patients with liver fibrosis in the study. They extracted 354 CT texture features and constructed a CTTA-based model using a random forest classifier (RFC) algorithm and SVM algorithm, respectively. The AUROC of the CTTA-based model based on the RFC algorithm in predicting fibrosis grade was 0.90 in both the first and second analyses. The AUC of the CTTA-based model using the SVM algorithm in predicting liver fibrosis in analysis I was 0.76, and 0.91 in analysis II that was the highest prediction rate. Son et al.[31] also evaluated liver fibrosis severity using splenic volume (measured on portal venous phase CT images). They proposed a convolution neural network (CNN) based on a three-dimensional U-net, with 513 chronic liver disease patients and 45 healthy liver subjects taken as modeled data, dividing the liver and spleen into many segments. By using the calculus method, the volume of hepatic and splenic segments in each thin slice was first calculated, then the volume of the liver and spleen in each thin slice was calculated, and finally, the volume of the liver and spleen amounted to the volume of the liver and spleen in each thin slice multiplied by the thickness of each thin slice. Then, they evaluated the correlation between liver and spleen volume indexes [Vol, (spleen volume), Vol, (liver volume), and Vol/ Vol,] and liver fibrosis severity and their diagnostic value in liver fibrosis severity. The results demonstrated that the AUCs were 0.82, 0.85, and 0.88 of Vol/ Vol, for the diagnosis of advanced fibrosis, cirrhosis, and decompensated cirrhosis in the entire study population, which outperformed the Vol,.

A variety of methods have been used to detect the stage of liver fibrosis, such as serological markers and elasticity scores based on ultrasound[32,33]. Yasaka et al[34] specifically included 534 patients with hepatic fibrosis into training datasets and 100 fibrosis patients into test sets, analyzing the diagnostic value of deep CNN using gadoxetic acid-enhanced hepatobiliary phase MR images and MR/virus model for liver fibrosis. The results showed that AUCs were 0.84, 0.84, and 0.85 of the model full (combined MR images and MR/virus model) for diagnosing F4, ≥ F3, and ≥ F2 in test sets, respectively, while 0.81, 0.84, and 0.83 of MR/virus model for diagnosing F4, ≥ F3, and ≥ F2 in test sets.

Previous studies have been limited to single-center studies to explore the diagnostic value of ultrasound, CT, or MRI images-based radiomics or deep learning model for liver fibrosis. Wang et al[35] subsequently carried out a prospective multicenter study based on deep learning radiomics of shear wave elastography (DLRE) to assess the diagnostic performance for evaluating liver fibrosis in chronic hepatitis B, and compare it with 2D-SWE (shear wave elastography) and serological biomarkers [APRI and fibrosis index based on four factors (FIB-4)]. A total of 398 patients from 12 hospitals in China with 1990 2D-SWE images were included in the study, and a DLRE model was established using the CNN algorithm. The results showed that the AUCs of DLRE were 0.97, 0.98, 0.85 for classifying F4, ≥ F3, and ≥ F2 in the validation set, which statistically outperformed 2D-SWE and serological biomarkers (APRI and FIB-4) except 2D-SWE in ≥ F2. Simplified structural flowchart of convolutional neural network architecture for staging liver fibrosis can be seen in Figure 1.

**Diagnosing HCC**

Typical focal nodular hyperplasia (FNH) has central cicatricial foci, which is enhanced in the arterial phase (AP), and the enhancement degree gradually decreases to low density in the portal venous phase (PVP). Atypical FNH has no central cicatricial foci and shows low perfusion, even hemorrhage, necrosis, and calcification on enhanced CT, which makes it difficult to differentiate atypical FNH from HCC[36,37]. Hence, Nie et al[38] aimed to establish a radiomics nomogram model based on CT for differentiating FNH from HCC. A total of 156 patients with FNH and HCC were divided into a training group and test group of the above-mentioned radiomics nomogram model. Initially, 4227 radiomics features were extracted from the regions of interest (ROIs) on
hepatic CT scan. After selection for inter- and intra-class correlation coefficients (ICCs) > 0.75 and statistical difference between the two groups, 764 radiomics features were entered into least absolute shrinkage and selection operator (LASSO) regression, and finally, ten features were used for the construction of a radiomics signature. After multivariate logistic regression calculation, old age, HBV infection, and early enhancement (with a washout pattern) were used to establish the clinical factors model. Three models were then established: Radiomics signature model, clinical factors model, and radiomics nomogram model (clinical factors and radiomics signature combined). The predictive performance of the radiomics nomogram model was superior to those of the other constructed models. The accuracy of the radiomics nomogram model for diagnosing HCC was 92.4% and 89.2% in the training and validation datasets, respectively.

To confirm the value of the AI model for predicting HCC, Jiang et al.[39] implemented a prospective study to compare the diagnostic differences of the 2018 European Association for the Study of the Liver (EASL) criteria, Liver Imaging Reporting and Data System (LI-RADS) criteria, and a radiomics signature model for diagnosing HCC. They included 211 patients with surgically confirmed focal liver lesions (165 HCC patients, 30 non-HCC malignancies patients, and 16 non-HCC benign lesions patients) who were randomly divided into a training cohort (n = 133) and a validation cohort (n = 78). The results demonstrated that the AUROCs of HCC risk in all nodules calculated by EASL v2018 criteria, LR5/LR5V in LI-RADS v2018 criteria, and radiomics signature model were 0.811, 0.841, and 0.810, respectively. The benignity and malignancy of liver tumors are usually based on Edmondson grade. Generally speaking, low-grade liver tumors conform to Edmondson I/II, while high-grade liver tumors conform to Edmondson III/IV.[40] Wu et al.[41] included 125 HCC patients in the training group and 45 HCC patients in the validation group to predict the grade of HCC patients based on unenhanced MRI radiomics. The MRI image was uploaded to ITK-SNAP software, and the ROI was depicted along the edge of the tumor on each TIWI and T2WI segment with radiomics features extracted from ROI. A total of 656 radiomics features were extracted. After LASSO regression selection, 14, 18, and 20 features were selected based on T1WI, T2WI, and combined T1WI and T2WI, respectively. Finally, the combined model was established with results summarized in Table 2.

**Predicting microvascular invasion of HCC**

The occurrence of microvascular invasion (MVI) often reduces the 5-year survival rate of HCC patients, which is closely related to the poor prognosis of HCC patients.[42] Peng et al.[43] randomly assigned 304 cases of eligible HCC patients to the training group and validation group. IBEX software was used to extract radiomics features from the ROI of enhanced CT images in each of the HCC patients. In this study, 980 candidate radiomics features were generated from each patient, and after screening by LASSO with λ chosen smallest cross-validation error, eight radiomics features were selected to establish the radiomics signature. Radiomics signature, together with radiologic features (non-smooth tumor margin, internal arteries, and hypoattenuating halos) and alpha-fetoprotein > 20 ng/mL, was used to construct the radiomics nomogram, whose AUROC values for predicting MVI status were 0.846 and 0.844 in the training and validation sets, respectively. Ma et al.[44] also discussed the value of preoperative radiomics nomogram for diagnosing MVI in HCC patients based on enhanced CT; 110 HCC patients were randomly assigned to the training cohort (MVI+ group 57 cases, MVI-group 73 cases), and 47 HCC patients to the validation cohort (MVI+ group 18 cases, MVI-group 29 cases). The ROI was drawn by an experienced
radiologist using ITK-SNAP software in AP, PVP, and delayed phase (DP) images on CT along the visible borders of the lesion. Initially, in the AP, PVP, and DP images, 647 radiomics features were extracted. After screening by ICC and concordance correlation coefficient ≥ 0.75 and LASSO method, finally, five AP features, seven PVP features, and nine DP features were selected. In predicting HCC invasion of MVI, PVP and radiomics signature outperformed AP and DP, hence, the PVP radiomics signature with the effective clinical factors constituted the combined model, whose AUCs were 0.835 and 0.801 in the training and validation groups, respectively for diagnosing MVI status in HCC.

**Predicting HE secondary to hepatitis B related cirrhosis**
Decompensated liver cirrhosis is often accompanied by multiple complications, of which HE is the most serious complication. Therefore, it is worthwhile to try to predict hepatic fibrosis patients who are susceptible to HE to provide early prevention and treatment[45]. A total of 304 cases of first-diagnosed hepatitis B-related cirrhosis were included in their study, 212 cases in the training group (HE = 38, non-HE = 174) and 92 cases in the validation group (HE = 21, non-HE = 71). Initially, 356 radiomics features were picked from the ROI of a patient’s liver enhanced CT. After three times selection, 19 radiomics features were applied to construct a radiomics model. Three clinical factors related to HE, including serum albumin, ascites, and collateral circulation, were used to establish a clinical factor model. The results showed that, in the training and validation cohorts, the model combining radiomics and clinical features had an accuracy rate of 0.93 and 0.83 in predicting the risk of HE complicated by hepatitis B cirrhosis, respectively, while the corresponding values were 0.89 and 0.84 for radiomics model, and 0.83 and 0.77 for the clinical model.

**Predicting liver failure in cirrhotic patients with HCC after major hepatectomy**
Postoperative hepatic failure is the most serious complication in HCC patients after hepatectomy, which not only prolongs the hospitalization time of patients but also increases the mortality of patients[46]. Zhu et al[47] included 101 eligible HCC patients who had a major liver resection, to use preoperative gadoxetic acid-enhanced MRI for predicting liver failure based on the radiomics model. Removing the necessary structures, on the hepatobiliary phase image of the above-mentioned patient, experienced radiologists drew ROIs of the liver along the edge of the whole hepatic parenchyma. Through screening for ICC > 0.75 and LASSO regression, five features were employed to establish a radiomics signature. Then, the radiomics signature and ICG-R15 (a variable in the clinical factors model) were incorporated to establish a radiomics-based model. The results showed that the accuracy of the radiomics-based model in predicting postoperative liver failure was 0.802.

**CONCLUSION**

AI can mine data from patients’ medical records, the Centers for Disease Control and Prevention in countries or regions, and serological markers. Raman spectroscopy can predict the incidence of hepatitis, classify the different stages of hepatitis, diagnose or screen hepatitis, forecast the progression of hepatitis, and predict response to antiviral drugs in CHC patients with a high level of diagnostic performance. AI, especially deep learning or radiomics, can predict hepatitis, predict liver fibrosis and its grade, differentiate HCC from the benign liver tumor, predict the risk of vascular invasion of HCC, predict the risk of HE complicated by hepatitis B related cirrhosis, and predict the risk of liver failure after hepatectomy in HCC patients, with good levels of accuracy.

**Limitations and future perspectives**
At present, all reviewed studies using AI are mostly single-center research, and the amount of data in the training sets may be insufficient. In the future, people should build an internet database or hospital, and disease-related information from various countries or regions could be uploaded to the internet database or hospital so that AI researchers can obtain more disease-related information with different demographics, geographic areas, etc., to carry out multicenter research, and establish a more robust and inclusive AI model for clinical use.
REFERENCES

1 Angermueller C, Parnamaa T, Parts L, Stegle O. Deep learning for computational biology. Mol Syst Biol 2016; 12: 878 [PMID: 27474269 DOI: 10.15252/msb.20156651]

2 Zhou LQ, Wang JY, Yu SY, Wu GG, Wei Q, Deng YB, Wu XL, Cui XW, Dietrich CF. Artificial intelligence in medical imaging of the liver. World J Gastroenterol 2019; 25: 672-682 [PMID: 30783371 DOI: 10.3748/wjg.v25.i6.6762]

3 Chan HP, Sanalla RK, Hadijiski LM, Zhou C. Deep learning in Medical Image Analysis. Adv Exp Med Biol 2020; 1213: 3-21 [PMID: 32030660 DOI: 10.1007/978-3-030-33128-3_1]

4 Gore JC. Artificial intelligence in medical imaging. Magn Reson Imaging 2020; 68: A1-A4 [PMID: 31857130 DOI: 10.1016/j.mri.2019.12.006]

5 Liang X, Yu J, Liao J, Chen Z. Convolutional Neural Network for Breast and Thyroid Nodules Diagnosis in Ultrasound Imaging. Biomed Res Int 2020; 2020: 1763803 [PMID: 32420322 DOI: 10.1155/2020/1763803]

6 Wang F, Liu X, Yuan N, Qian B, Ruan L, Yin C, Jin C. Study on automatic detection and classification of breast nodule using deep convolutional neural network system. J Thorac Dis 2020; 12: 4690-4701 [PMID: 33145042 DOI: 10.21037/jtd-19-3013]

7 Dong F, She R, Cui C, Shi S, Hu X, Zeng J, Wu H, Xu J, Zhang Y. One step further into the blackbox: a pilot study of how to build more confidence around an AI-based decision system of breast nodule assessment in 2D ultrasound. Eur Radiol 2021; 31: 4991-5000 [PMID: 33404698 DOI: 10.1007/s00330-020-07561-7]

8 Khan S, Ullah R, Khan A, Ashraf R, Ali H, Bilal M, Saleem M. Analysis of hepatitis B virus infection in blood sera using Raman spectroscopy and machine learning. Photodiagnosis Photodyn Ther 2018; 23: 89-93 [PMID: 29787817 DOI: 10.1016/j.pdpdt.2018.05.010]

9 Li W, Huang Y, Zhang BW, Liu JG, Hu HT, Li X, Liang JY, Wang Z, Huang XW, Zhang CQ, Ruan SM, Xie XY, Kuang M, Lu MD, Chen LD, Wang W. Multiparametric ultrasonics of significant liver fibrosis: A machine learning-based analysis. Eur Radiol 2019; 29: 1496-1506 [PMID: 30178143 DOI: 10.1007/s00330-019-5680-z]

10 Zhen S, Cai X. Letter to the Editor: Predicting Survival After Hepatocellular Carcinoma Resection Using Deep-Learning on Histological Slides. Hepatology 2021; 73: 2077-2078 [PMID: 32894573 DOI: 10.1002/hep.31543]

11 Lavanchy D. Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. J Viral Hepat 2004; 11: 97-107 [PMID: 14996343 DOI: 10.1046/j.1365-2893.2003.00487.x]

12 Guan P, Huang DS, Zhou BS. Forecasting model for the incidence of hepatitis A based on artificial neural network. World J Gastroenterol 2004; 10: 3579-3582 [PMID: 15553410 DOI: 10.3748/wjg.v10.i24.3579]

13 Zheng Y, Zhang L, Zhuo G. A comparative study of two methods to predict the incidence of hepatitis B in Guangxi, China. PLoS One 2020; 15: e0234660 [PMID: 32579590 DOI: 10.1371/journal.pone.0234660]

14 Gan R, Chen X, Yan Y, Huang D. Application of a hybrid method combining grey model and back propagation artificial neural networks to forecast hepatitis B in China. Comput Math Methods Med 2015; 2015: 328273 [PMID: 25815044 DOI: 10.1155/2015/328273]

15 Guo Y, Feng Y, Qu F, Zhang L, Yan B, Lv J. Prediction of hepatitis E using machine learning models. PLoS One 2020; 15: e0237750 [PMID: 32941452 DOI: 10.1371/journal.pone.0237750]

16 Ahmad G, Khan MA, Abbas S, Athar A, Khan BS, Aslam MS. Automated Diagnosis of Hepatitis B Using Multilayer Mamdani Fuzzy Inference System. J Healthc Eng 2019; 2019: 6361318 [PMID: 30867895 DOI: 10.1155/2019/6361318]

17 Wang X, Tian S, Yu L, Lv X, Zhang Z. Rapid screening of hepatitis B using Raman spectroscopy and long short-term memory neural network. Lasers Med Sci 2020; 35: 1791-1799 [PMID: 32285292 DOI: 10.1007/s11010-020-03003-4]

18 Yehia BR, Schranz AJ, Umscheid CA, Lo Re V 3rd. The treatment cascade for chronic hepatitis C virus infection in the United States: a systematic review and meta-analysis. PLoS One 2014; 9: e101554 [PMID: 24988388 DOI: 10.1371/journal.pone.0101554]

19 Doyle OM, Leavitt N, Riggi JA. Finding undiagnosed patients with hepatitis C infection: an application of artificial intelligence to patient claims data. Sci Rep 2020; 10: 10521 [PMID: 32601354 DOI: 10.1038/s41598-020-76713-6]

20 Konerman MA, Boste AL, van T, Liu B, Zhang X, Zhu J, Saini SD, Su GL, Nallamothu BK, Joannou GN, Waljee AK. Machine learning models to predict disease progression among veterans with hepatitis C virus. PLoS One 2019; 14: e0208141 [PMID: 30608929 DOI: 10.1371/journal.pone.0208141]

21 Maielloaro PA, Cozzolongo R, Marino P. Artificial neural networks for the prediction of response to interferon plus ribavirin treatment in patients with chronic hepatitis C. Curr Pharm Des 2004; 10: 2101-2109 [PMID: 15279549 DOI: 10.2174/138161204384240]

22 Chartrand G, Cheng PM, Vorontsov E, Drozdzal M, Turcotte S, Pal CJ, Kadoury S, Tang A. Deep Learning: A Primer for Radiologists. Radiographics 2017; 37: 2113-2131 [PMID: 29131760 DOI: 10.1148/rg.2017170077]

23 Lee JG, Jun S, Cho YW, Lee H, Kim GB, See JB, Kim N. Deep Learning in Medical Imaging: General Overview. Korean J Radiol 2017; 18: 570-584 [PMID: 28670152 DOI: ]
B related cirrhosis.

CT.
omogram for microvascular invasion prediction in hepatocellular carcinoma using contrast-enhanced
Ma X
S59-S63 [PMID: 10.1002/1097-0142(195405)7:3<462::aid-cncr2820070308>3.0.co;2-e
necropsies.
Edmondson HA
Talwalkar JA
dimeglumine.
liver lesions: dynamic and late enhancement patterns with the dual phase contrast agent gadobenate
1660-1668 [PMID: 10.1136/gutjnl-2018-316204
hepatitis B: a prospective multicentre study.
Ai H, Xie XY, Yin LP, Liang P, Tian J, Zheng R. Deep learning Radiomics of shear wave
time.
Yasaka K
134
Friedrich-Rust M
elastography for the detection of hepatic fibrosis: systematic review and meta-analysis.
Talwalkar JA, Kurtz DM, Schoenleber SJ, West CP, Montori VM. Ultrasound-based transverse
elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic
grade and high-grade fibrosis.
Biomark Res
2020; 8: 47 [PMID: 32963787 DOI: 10.1038/s41436-020-00219-y]
B related cirrhosis.

CT.
omogram for microvascular invasion prediction in hepatocellular carcinoma using contrast-enhanced
Ma X
S59-S63 [PMID: 10.1002/1097-0142(195405)7:3<462::aid-cncr2820070308>3.0.co;2-e
necropsies.
Edmondson HA
Talwalkar JA
dimeglumine.
liver lesions: dynamic and late enhancement patterns with the dual phase contrast agent gadobenate
1660-1668 [PMID: 10.1136/gutjnl-2018-316204
hepatitis B: a prospective multicentre study.
Ai H, Xie XY, Yin LP, Liang P, Tian J, Zheng R. Deep learning Radiomics of shear wave
time.
Yasaka K
134
Friedrich-Rust M
elastography for the detection of hepatic fibrosis: systematic review and meta-analysis.
Talwalkar JA, Kurtz DM, Schoenleber SJ, West CP, Montori VM. Ultrasound-based transverse
elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic
grade and high-grade fibrosis.
Biomark Res
2020; 8: 47 [PMID: 32963787 DOI: 10.1038/s41436-020-00219-y]
B related cirrhosis.

CT.
omogram for microvascular invasion prediction in hepatocellular carcinoma using contrast-enhanced
Ma X
S59-S63 [PMID: 10.1002/1097-0142(195405)7:3<462::aid-cncr2820070308>3.0.co;2-e
necropsies.
Edmondson HA
Talwalkar JA
dimeglumine.
liver lesions: dynamic and late enhancement patterns with the dual phase contrast agent gadobenate
1660-1668 [PMID: 10.1136/gutjnl-2018-316204
hepatitis B: a prospective multicentre study.
Ai H, Xie XY, Yin LP, Liang P, Tian J, Zheng R. Deep learning Radiomics of shear wave
time.
Yasaka K
134
Friedrich-Rust M
elastography for the detection of hepatic fibrosis: systematic review and meta-analysis.
Talwalkar JA, Kurtz DM, Schoenleber SJ, West CP, Montori VM. Ultrasound-based transverse
elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic
grade and high-grade fibrosis.
Biomark Res
2020; 8: 47 [PMID: 32963787 DOI: 10.1038/s41436-020-00219-y]
postoperative liver failure and the economic impact of liver function capacity after extended liver resection—a single-center experience. Langenbecks Arch Surg 2009; 394: 1047-1056 [PMID: 19533168 DOI: 10.1007/s00423-009-0518-4]

Zhu WS, Shi SY, Yang ZH, Song C, Shen J. Radiomics model based on preoperative gadoxetic acid-enhanced MRI for predicting liver failure. World J Gastroenterol 2020; 26: 1208-1220 [PMID: 32231424 DOI: 10.3748/wjg.v26.i11.1208]
