Development and validation of radiomics model built by incorporating machine learning for identifying liver fibrosis and early-stage cirrhosis

Qing-Tao Qiu[1], Jing Zhang[2], Jing-Hao Duan[1], Shi-Zhang Wu[1], Jia-Lin Ding[3], Yong Yin[1]

[1]Department of Radiation Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, Shandong 250117, China;
[2]Department of Radiation Oncology, Cancer Institute and Hospital, Tianjin Medical University, Tianjin 300060, China;
[3]School of Physics and Electronics, Shandong Normal University, Jinan, Shandong 250358, China.

Abstract

Background: Liver fibrosis (LF) continues to develop and eventually progresses to cirrhosis. However, LF and early-stage cirrhosis (ESC) can be reversed in some cases, while advanced cirrhosis is almost impossible to cure. Advances in quantitative imaging techniques have made it possible to replace the gold standard biopsy method with non-invasive imaging, such as radiomics. Therefore, the purpose of this study is to develop a radiomics model to identify LF and ESC.

Methods: Patients with LF ($n = 108$) and ESC ($n = 116$) were enrolled in this study. As a control, patients with healthy livers were involved in the study ($n = 145$). Diffusion-weighted imaging (DWI) data sets with three *b*-values (0, 400, and 800 s/mm^2) of enrolled cases were collected in this study. Then, radiomics features were extracted from manually delineated volumes of interest. Two modeling strategies were performed after univariate analysis and feature selection. Finally, an optimal model was determined by the receiver operating characteristic area under the curve (AUC).

Results: The optimal models were built in plan 1. For model 1 in plan 1, the AUCs of the training and validation cohorts were 0.973 (95% confidence interval [CI] 0.946–1.000) and 0.948 (95% CI 0.903–0.993), respectively. For model 2 in plan 1, the AUCs of the training and validation cohorts were 0.944, 95% CI 0.905 to 0.983, and 0.968, 95% CI 0.940 to 0.996, respectively.

Conclusions: Radiomics analysis of DWI images allows for accurate identification of LF and ESC, and the non-invasive biomarkers extracted from the functional DWI images can serve as a better alternative to biopsy.

Keywords: Diffusion-weighted imaging; Liver fibrosis; Early-stage cirrhosis; Radiomics; Machine learning

Introduction

Hepatic cirrhosis is the main cause of some serious public health problems and has a high prevalence worldwide and poor long-term clinical outcomes in China.[1] In many patients, cirrhosis is caused by the progression of liver fibrosis (LF); however, LF (METAVIR fibrosis stage: F1–F3) and early-stage cirrhosis (ESC) (METAVIR fibrosis stage: F4) can be reversed in some cases, while advanced cirrhosis is almost impossible to cure. Therefore, it is necessary to identify LF and ESC. Although liver biopsy has been considered the gold standard method for diagnosing and staging LF and ESC, it is invasive and cannot probe fibrosis and cirrhosis of the whole liver. Moreover, this method may be associated with high sampling variability and low patient tolerance.[2] Medical imaging, a non-invasive and repeatable technique, plays an important role in the diagnosis and identification of disease.[3] Thus, if clinical and imaging data are more accurate and sensitive, they are necessary.

Magnetic resonance imaging (MRI), which is different from computed tomography, ultrasound, and other anatomical imaging methods, is mainly focused on functional imaging and is widely used in many diseases.[4,5] Especially for diffusion-weighted imaging (DWI), a commonly used functional imaging sequence, it is effective in the detection and staging of liver metastatic tumors[6] and in the evaluation of LF and cirrhosis.[7,8]

Radiomics, also known as texture analysis or quantitative imaging analysis in early published work, is a new application using established techniques to extract amounts of quantitative imaging features from medical

Access this article online

Quick Response Code:

Website: www.cmj.org

DOI: 10.1097/CM9.0000000000001113

Correspondence to: Prof. Yong Yin, Department of Radiation Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, Shandong 250117, China
E-Mail: yinyongsd@126.com

Chinese Medical Journal 2020;133(22)

Received: 05-05-2020 **Edited by:** Qiang Shi

images followed by data mining for clinical use.[9-12] Biomarkers developed by radiomics potentially improve decision support in diagnosis and prognosis.[13-15] For instance, entropy-related features have been developed for distinguishing benign and malignant cervical tissues[16]; a similar conclusion was drawn by Corino *et al*,[17] who proved the feasibility of grading soft tissue sarcomas using DWI-based radiomics features. Some studies have demonstrated that there are strong correlations between textural features and liver disease.[18,19] In these studies mentioned above, something in common is that all the texture analyses were performed on DWI images.

As advanced hepatic fibrosis, cirrhosis is a slow and gradual process. The liver continues to function in the early stage of cirrhosis, whereas the liver will begin to fail as cirrhosis becomes worse.[20,21] Therefore, it is essential to diagnose ESC as early as possible to deliver timely treatment. In this study, we aimed to develop a radiomics model to identify LF and ESC that could be potentially used in quantitative clinical diagnosis of LF and ESC.

Methods

Ethical approval

Ethical approval (No. 2019070031) was obtained from the Institutional Review Board of Shandong Cancer Hospital and Institute for this retrospective study, and the need for informed consent was waived.

Study design

The overall workflow of this study consists of five steps, including the collection of patients and images, definition of volumes of interest (VOI), extraction of radiomics features, modeling, and performance evaluation. For the modeling strategy, a brief description is given in Figure 1.

Patients and DWI images

This study included 369 cases at Shandong Cancer Hospital and Institute between January 2014 and June 2019, including 108 patients with LF, 116 patients with ESC, and 145 patients with healthy livers (patients with other malignancies but have healthy livers). The description of LF and ESC is shown in Supplementary file 1, http://links.lww.com/CM9/A340. The inclusion criteria were as follows: (1) normal liver morphology; (2) pathological proved LF and ESC; (3) clear stage of LF and ESC; (4) without space-occupying lesions. While the exclusion criteria were as follows: (1) without DWI scans; (2) abnormal liver morphology or a lot of ascites; (3) insufficient image quality, such as artifacts.

All MR images were acquired with a Philips 3.0 Tesla MR scanner (Philips Medical System, Netherlands) and eight-channel abdominal phased array coils. DWI images were acquired with the following parameters: time of repeatation (TR) = 10,000 ms, time of echo (TE) = 66.1 ms, slice thickness = 7 mm, matrix size = 256 × 256, in-plane resolution = 1.56 × 1.56 mm^2 and b-values = 0, 20, 50, 100, 200, 400, 800, 1000, and 1200 s/mm^2. Based on our previous research, which focused on the impact of b-values on DWI images,[22] similar radiomics features can be extracted within group b = 0, 20, 50, 100, 200 s/mm^2, group 50, 100, 200, and 400 s/mm^2, and group b = 800, 1000, and 1200 s/mm^2. Thus, we selected three b-value images that provided non-redundant information, including b = 0 s/mm^2, b = 400 s/mm^2, and 800 s/mm^2. Additionally, high-resolution T1-weighted sequence images were obtained

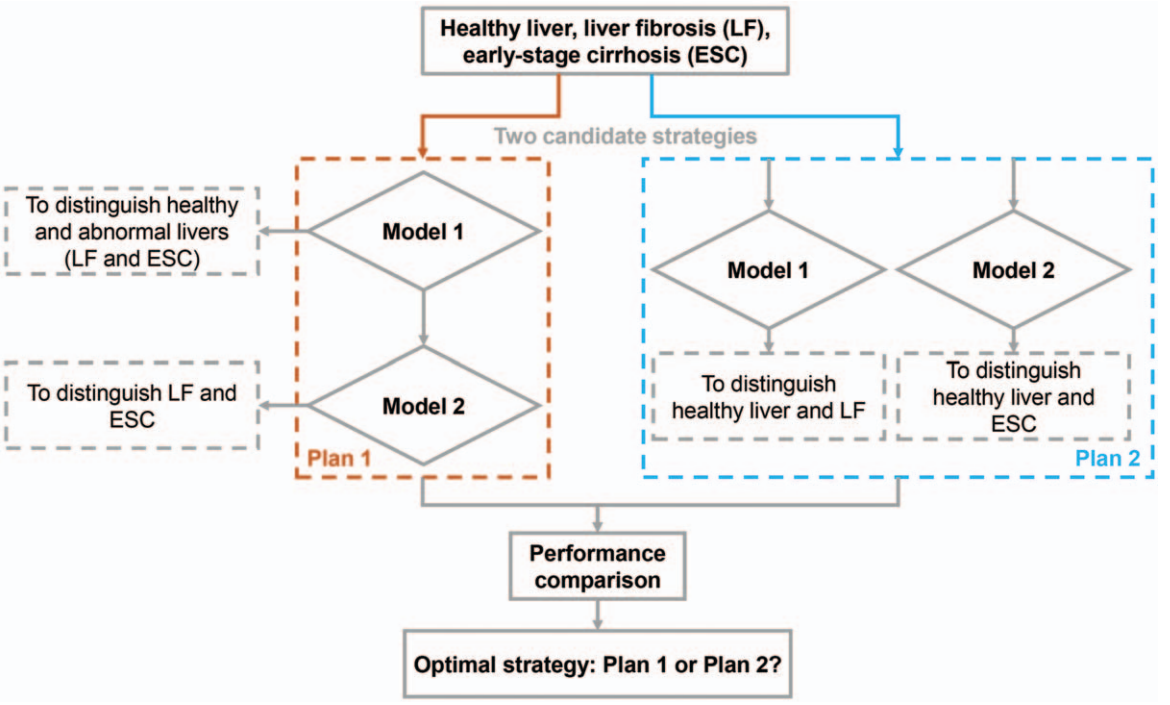

Figure 1: Workflow of the modeling strategy of this study. ESC: Early-stage cirrhosis; LF: Liver fibrosis.

with the parameters as follows: TR = 3.7 ms, TE = 1.7 ms, slice thickness = 5.2 mm, matrix size = 512 × 512, in-plane resolution = 1.56 × 1.56 mm^2.

Volume of interest (VOI)

The VOI, defined as the region of LF and ESC, was delineated by two experienced abdominal radiologists in MIM maestro software (version 6.8.2, MIM Software Inc., Cleveland, OH 44122, USA). First, rigid registration was used to align T1WI images to b = 0 s/mm^2 DWI images. Then, using the anatomical information of T1WI as a reference, avoiding the inferior vena cava, portal vein and portal area, three cylindrical VOIs of diameter 20 mm and height 14 mm (thickness of two slices) were defined by one radiologist in the parenchyma of liver segments II/III, V/VI, and VII of b = 0 s/mm^2 DWI images, respectively. After that, all the VOIs were verified by another radiologist. Finally, the VOIs in b = 0 s/mm^2 were mapped to the other nine b-value DWI images. The examples of VOIs are depicted in Figure 2.

Feature extraction

Radiomic features were automatically extracted from each VOI of all DWI scans using SlicerRadomics Extension in 3D Slicer (stable released version 4.8.1, www.slicer.org), open-source, easy to use medical image analysis software.[23] In total, we extracted 93 DWI imaging features, which are divided into six categories: (1) 18 first-order intensity histogram-based features and 75 textural features, including (2) 24 gray-level cooccurrence matrix (GLCM)-based features, (3) 16 gray-level run-length matrix (GLRLM)-based features, (4) 16 gray-level size-zone matrix-based features, (5) five neighboring gray-tone difference matrix (NGTDM)-based features, and (6) 14 gray-level dependence matrix (GLDM)-based features.

To calculate the radiomics features accurately and effectively and to further enhance the inter- and intra-patient reproducibility of feature values, the medical imaging data require discretization of the intensity values.[24] In this study, we adopted a suggestion that was made by Zwanenburg *et al*[25] to rescale the DWI voxel intensity values into fixed-count 100 bins within VOIs before calculating the GLCM, GLRLM, GLSZM, NGTDM, and GLDM features.

Model construction

In this study, the classification model was constructed by following two plans, and each plan contains two models, as shown in Figure 1. For plan 1, model 1 was constructed first to classify healthy and abnormal liver; subsequently, model 2 was built on the basis of the model 1 to identify LF or ESC from abnormal liver. For plan 2, models 1 and 2 were built concurrently to identify the healthy liver and LF and healthy liver and ESC, respectively.

In the process of model construction, two-step feature selection was conducted. First, univariate analysis was used to retain features for model training that were informative and predictive. Second, the feature selection algorithm named RELIEFF, an adaptation of relief feature selection method, was used to avoid model overfitting; the rule of thumb is that the number of predictors should be 1/10 to 1/3 of the sample size in each model of the training cohort.[26] The details of RELIEFF can be found in Supplementary File 1, http://links.lww.com/CM9/A340.

Finally, a support vector machine (SVM) algorithm with a radial basis function kernel was employed to build the classification model. In the training process, 1000 times ten-fold cross-validation was used to avoid overfitting and

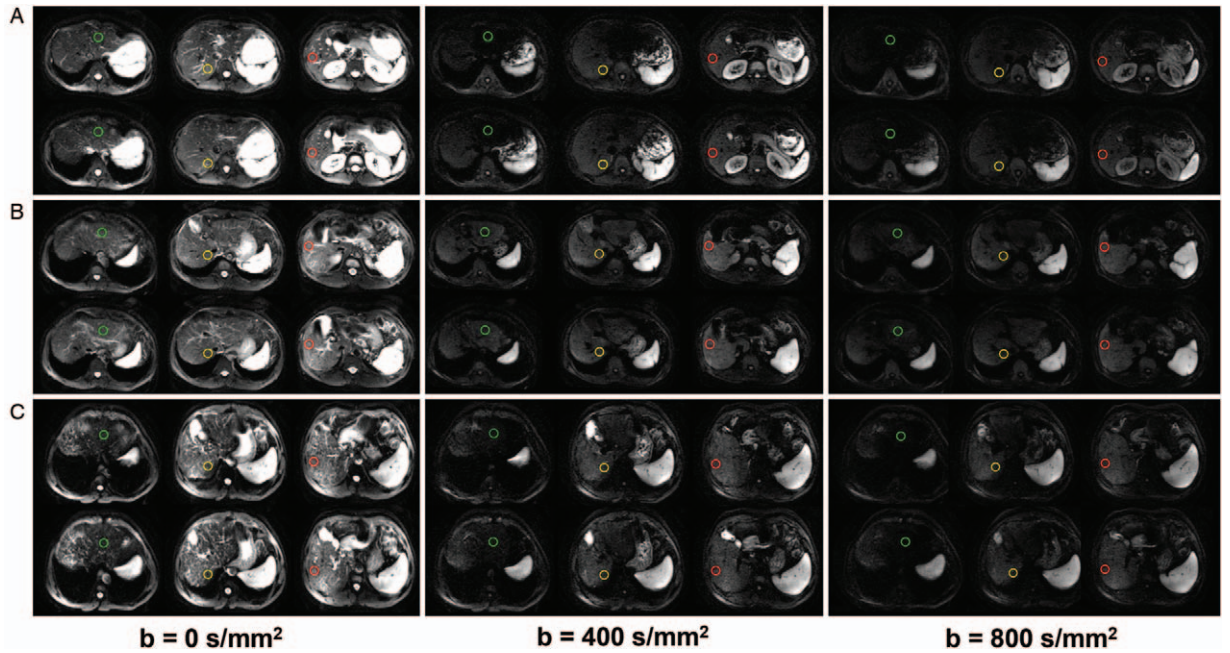

Figure 2: Examples of volume of interest in this study. (A) healthy liver, (B) liver fibrosis, (C) early-stage cirrhosis. Columns from left to right are diffusion-weighted images with b = 0, 400, and 800 s/mm^2, respectively. Green, yellow, and red contours are liver segments II/III, V/VI, and VII, respectively.

Table 1: The sample size of the training and validation cohorts of models in each plan (*n*).

	Plan 1				Plan 2			
	Model 1		Model 2		Model 1		Model 2	
Groups	Training	Validation	Training	Validation	Training	Validation	Training	Validation
Healthy	110	35	N/A	N/A	100	45	100	45
LF	74	31	68	40	70	38	N/A	N/A
ESC	82	34	76	40	N/A	N/A	70	46

LF: Liver fibrosis; ESC: Early stage cirrhosis; N/A: Not applicable.

to ensure the model with optimal performance can be obtained. After training, a validation cohort was employed to test the performance. In this study, the training and validation cohorts were divided randomly. The sample size of the training and validation cohorts in each model is listed in Table 1.

Statistical analysis

For univariate analysis, features with P values < 0.1 were considered to be associated with the response variable.[27] In this study, univariate analysis and model performance assessment were implemented in R software (version 3.3.1, www.r-project.org) with the "stats" package and the "pROC" package, respectively. The RELIEFF algorithm was implemented in Matlab software (version R2018a, The MathWorks Inc., Natick, MA 01760, USA) with the feature selection library (version 6.2.1) toolbox. Machine learning modeling was performed in Matlab with in-house code.

Results

Univariate analysis and RELIEFF feature selection

For initial univariate analysis, 75 and 63 of 279 features showed potential predictive power on corresponding response variables, that is, healthy versus abnormal in plan 1 model 1 and LF *vs.* ESC in plan 1 model 2. Meanwhile, 62 and 59 of 279 features showed predictive power on corresponding variables, that is, healthy versus LF in plan 2 model 1 and healthy versus ESC in plan 2 model 2. Subsequently, the RELIEFF algorithm was used to sort the predictive features selected by single-factor analysis. RELIEFF feature selection was implemented to select three top-ranking features in terms of its weight in each feature category. The results of RELIEFF feature selection are shown in Supplementary File 1, http://links.lww.com/CM9/A340. In the training process of each model, there are 18 features as input variables for the SVM model. The P values of univariate analysis and RELIEFF weights of plans 1 and 2 are shown in Supplementary File 1, http://links.lww.com/CM9/A340.

Performance of constructed models

The SVM model incorporates three top-ranking features of each feature category, in total 18 features as input, with 1000 times ten-fold cross-validation. Plan 1 contains two series models, while plan 2 contains two parallel models. For model 1 in plan 1, the SVM model for differentiating healthy and abnormal liver, yielded an accuracy of 91.5% (95% confidence interval [CI], 89.3%–93.7%) and a receiver operating characteristic (ROC) area under the curve (AUC) of 0.973 (95% CI, 0.946–1.000) in the training cohort and an accuracy of 89.1% (95% CI, 86.4%–91.8%) and an AUC of 0.948 (95% CI, 0.903–0.993) in the validation cohort. For model 2 in plan 1, the SVM model for identifying LF and ESC, yielded an accuracy of 88.9% (95% CI, 87.3%–90.5%) and an AUC of 0.944 (95% CI, 0.905–0.983) in the training cohort and an accuracy of 92.6% (95% CI, 90.4%–94.8%) and an AUC of 0.968 (95% CI, 0.940–0.996) in the validation cohort. The ROC curves of models 1 and 2 in plan 1 are illustrated in Figure 3A and 3B, respectively. For model 1 in plan 2, the SVM model for differentiating healthy liver and LF, yielded an accuracy of 82.5% (95% CI, 79.3%–85.7%) and an AUC of 0.882 (95% CI, 0.845–0.919) in the training cohort and an accuracy of 82.1% (95% CI, 77.6%–86.6%) and an AUC of 0.857 (95% CI, 0.808–0.906) in the validation cohort. For model 2 in plan 2, the SVM model for identifying healthy liver and ESC, yielded an accuracy of 74.3% (95% CI, 70.2%–78.4%) and an AUC of 0.843 (95% CI, 0.793–0.899) in the training cohort and an accuracy of 79.3% (95% CI, 77.8%–80.8%) and an AUC of 0.863 (95% CI, 0.804–0.922) in the validation cohort. The ROC curves of models 1 and 2 in plan 2 are plotted in Figure 3C and 3D, respectively.

Optimal plan for identifying LF and ESC

The optimal plan for differentiating LF and ESC was determined by a robust validation. For plan 1, the robust validation is the same as the validation procedure of model 2. The excellent AUC values (0.968, 95% CI, 0.940–0.996) of the validation cohort indicated that good agreement existed between predicted class and ground-truth. For plan 2, the ESC as validation cases input into model 1 yielded false predicted LF. The AUC of this false predicted LF was 0.774 (95% CI, 0.720–0.828). Similar to the false prediction of model 1, the AUC of false predicted ESC was 0.698 (95% CI, 0.635–0.761). The ROC curve of the robust validation of plan 2 is illustrated in Figure 3E. The bias results indicated that plan 2 was not suitable for clinical use. Therefore, plan 1 was optimal for identifying LF and ESC.

Discussion

For patients with hepatic cirrhosis, it usually occurs following previous liver disease, such as LF, and may result

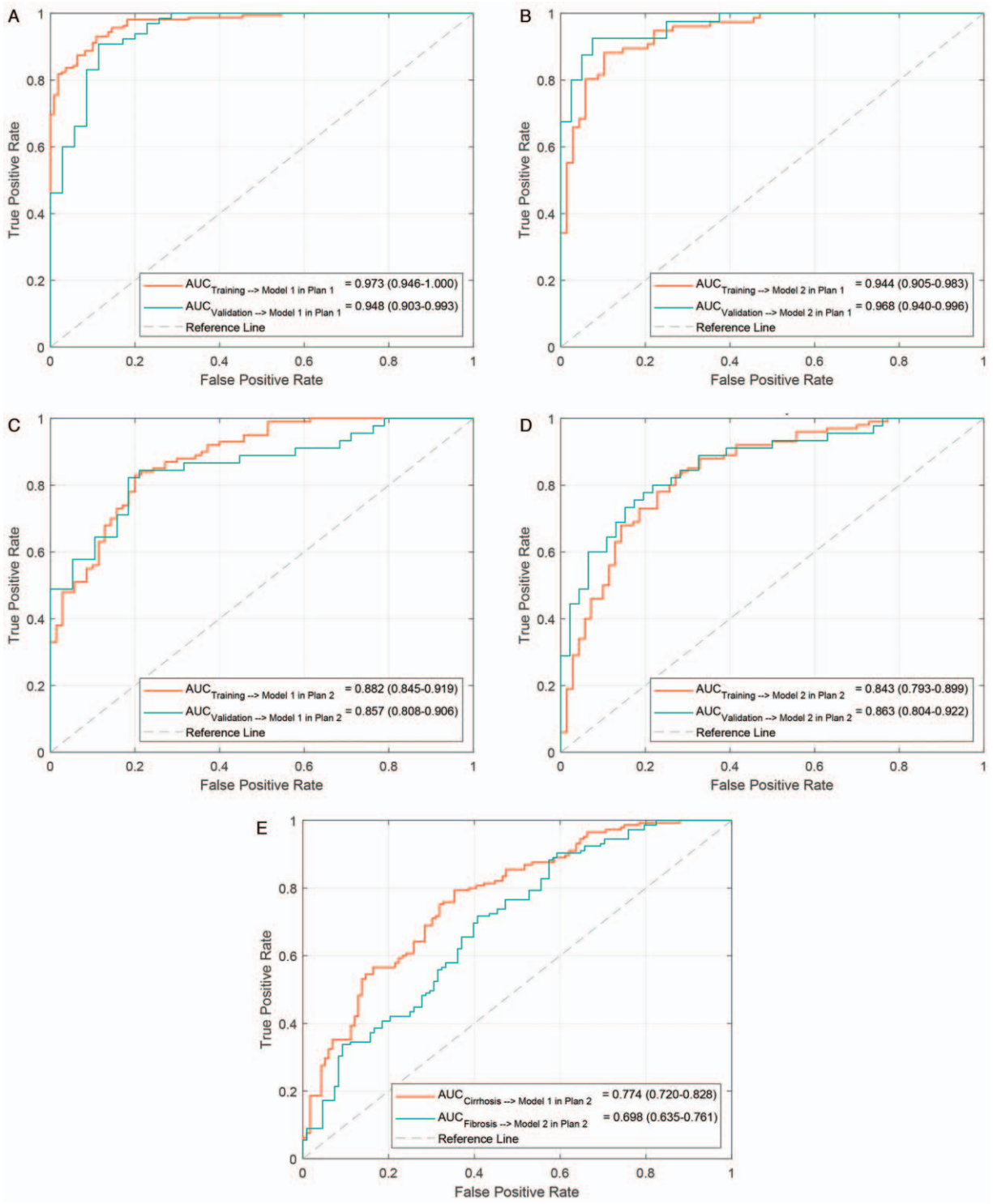

Figure 3: The performances of models developed in this study. (A and B) The ROC curves of models 1 and 2 in plan 1, respectively. (C and D) The ROC curves of models 1 and 2 in plan 2, respectively. (E) The ROC curve of the robust validation of plan 2. For robust validation of plan 2, false predictions occurred, including false predicted LF and ESC, when ESC and LF as robust test cohorts were input models 1 and 2 of plan 2, respectively. AUC: Area under the curve; ESC: Early-stage cirrhosis; LF: Liver fibrosis; ROC: Receiver operating characteristic.

in hepatocellular carcinoma.[28] Liver fibrosis continues to develop and eventually progress to cirrhosis. However, LF and ESC can be reversed in some cases, while advanced cirrhosis is almost impossible to cure. With the rise of quantitative imaging analysis in precision medicine, magnetic resonance DWI radiomics has been widely studied. In this study, the proposed DWI-based radiomics classification model for identifying LF shows remarkable AUC in plan 1, a series strategy. The results demonstrated that DWI is useful for predicting the LF or ESC of patients with liver disease and thus have the power to aid in the determination of subsequent treatment strategies.

Previous studies have tried to investigate the stage of LF using computerized texture analyses based on iron oxide and gadolinium chelate-enhanced MRI,[29,30] T2-weighted images,[31] and gadoxetic acid-enhanced MRI.[32,33] To our best knowledge, DWI radiomics studies for staging LF have been rarely reported. A radiomics model based on gadoxetic acid-enhanced MRI was recently developed for staging fibrosis (AUC $\leq$ 0.910, accuracy $\leq$ 82.1%).[32] The AUCs and accuracies are significantly lower than our plan 1 based on DWI radiomics for diagnosis of LF and ESC. Another paper published a model incorporating deep neural network with gadoxetic acid-enhanced MRI to stage LF (AUC $\leq$ 0.850)[34]; the training process was implemented by extracting and analyzing features from small cropped images of the liver that are not adequate to reflect all characteristics of LF, which may be why the performance of state-of-the-art deep learning is lower than the radiomics model of reference[32] and ours, although a further comparison is needed to compare these results head-to-head.

Three predecessor works and independent validations have been performed to ensure the robustness of constructed models. First, since the multiple *b*-values were used in DWI images, the impact of *b*-value on radiomics features needs to be explored thoroughly to build a robust model. Research conducted by Becker *et al*[35] reported that several features were correlated to *b*-values in multiple organs. Our previous studies have also proved that the similar radiomics features can be extracted within nearby *b*-value DWI images in patients with cirrhosis[22] and hepatocellular carcinoma.[36] Thus, we selected three 400-equally spaced *b*-value (b = 0 s/mm^2, b = 400, and 800 s/mm^2) images that provide non-redundant, multi-dimensional information. Second, to guarantee the reproducibility and robustness of feature extraction, we tried to discretize the voxel intensity with DWI VOIs. According to the experience of Tixier *et al*[37] and the guidelines of the Imaging Biomarker Standardization Initiative team,[25] a fixed number of bins ranging from 30 to 130 was adopted in this study, which can achieve a good reproducibility and performance. Meanwhile, this also allows for differing ranges of intensity within VOIs, while still keeping the texture features informative and comparable between VOIs or patients. However, the robustness of the model training process also needs to be considered. In the training process of a classification model, high-dimensional features are prone to overfitting, which may lead to optimistic results. To avoid the problem, univariate analysis and RELIEFF feature selection were used to remove the unreliable irrelevant features and to reduce the dimension of predictors. Finally, each model was validated in an independent validation cohort. The results of the training cohort in plan 1 showed excellent performance, and similar results were observed in the validation cohort, which suggests that our strategy can reliably minimize overfitting problems.

To determine the optimal plan, robust validation has been implemented. As a series plan, all cases were involved in the training phase of plan 1. For this reason, the distinctive features between healthy and abnormal liver and homogeneous features between LF and ESC were considered in the feature selection phase. Thus, the results were robust and less influenced by other factors. However, there is a lack of cirrhosis and fibrosis information in the training phase of models 1 and 2 in plan 2, respectively. Meanwhile, in the advanced stage of fibrosis, cirrhosis also has pathological characteristics similar to LF. Medical images are the macroscopic manifestation of the microscopic pathology, and the radiomics features are the quantitative descriptions of the medical images. Therefore, the radiomics features show similarities between fibrosis and cirrhosis, and our feature selection results have proven it. For models 1 and 2 in plan 2, four features were the same after feature selection, including 0_GLDM_DE, 400_GLDM_DE, 800_GLCM_Idn, and 0_GLSZM_SAE, which may be the reason why the false LF and false ESC occurred.

A limitation is that because this study was retrospective and preliminary, the ADC maps were unavailable when DWI images were acquired. We believe that the results will benefit greatly from the radiomics research of ADC maps. This work will be improved when we have prospectively collected large cirrhosis patient cohorts with DWI and ADC. Another limitation is that some clinical features are unavailable, such as hepatic biological parameters. The use of multiple parameters to build a holistic model to predict LF and ESC is a future research work and must be meaningful to improve the diagnosis of patients with hepatic disease.

In conclusion, radiomics analysis of DWI images allows for accurate identification of LF and ESC, and the non-invasive biomarkers extracted from functional DWI images might serve as a better alternative to biopsy.

Funding

This work was supported by grants from the Key Support Program of Natural Science Foundation of Shandong Province (No. ZR2019LZL017), the Taishan Scholars Project of Shandong Province (No. ts201712098), the National Natural Science Foundation of China (No. 81901743), the WBE Liver Fibrosis Foundation (No. CFHPC2019027), and the Key Research and Development Program of Shandong Province (No. 2019GSF108134).

Conflicts of interest

None.

References

1. Wang FS, Fan JG, Zhang Z, Gao B, Wang HY. The global burden of liver disease: the major impact of China. Hepatology 2014;60:2099–2108. doi: 10.1002/hep.27406.
2. Harada TL, Saito K, Araki Y, Matsubayashi J, Nagao T, Sugimoto K, *et al*. Prediction of high-stage liver fibrosis using ADC value on diffusion-weighted imaging and quantitative enhancement ratio at the hepatobiliary phase of Gd-EOB-DTPA-enhanced MRI at 1.5 T. Acta Radiol 2018;59:509–516. doi: 10.1177/0284185117725778.
3. O'Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, *et al*. Imaging biomarker roadmap for cancer studies. Nat Rev Clin Oncol 2017;14:169–186. doi: 10.1038/nrclinonc.2016.162.
4. Shetty M. Imaging and differential diagnosis of ovarian cancer. Semin Ultrasound CT MR 2019;40:302–318. doi: 10.1053/j.sult.2019.04.002.
5. Boesch C. Quantitative MR imaging is increasingly important in liver disease. Radiology 2018;286:557–559. doi: 10.1148/radiol.2017172312.

6. Lincke T, Zech CJ. Liver metastases: detection and staging. Eur J Radiol 2017;97:76–82. doi: 10.1016/j.ejrad.2017.10.016.
7. Petitclerc L, Gilbert G, Nguyen BN, Tang A. Liver fibrosis quantification by magnetic resonance imaging. Top Magn Reson Imaging 2017;26:229–241. doi: 10.1097/rmr.0000000000000149.
8. Donato H, França M, Candelária I, Caseiro-Alves F. Liver MRI: from basic protocol to advanced techniques. Eur J Radiol 2017;93:30–39. doi: 10.1016/j.ejrad.2017.05.028.
9. Qiu Q, Duan J, Duan Z, Meng X, Ma C, Zhu J, *et al*. Reproducibility and non-redundancy of radiomic features extracted from arterial phase CT scans in hepatocellular carcinoma patients: impact of tumor segmentation variability. Quant Imaging Med Surg 2019;9:453–464. doi: 10.21037/qims.2019.03.02.
10. Qiu Q, Duan J, Yin Y. Radiomics in radiotherapy: applications and future challenges. Precision Radiat Oncol 2020;4:29–33. doi: 10.1002/pro6.1087.
11. Parekh V, Jacobs MA. Radiomics: a new application from established techniques. Expert Rev Precis Med Drug Dev 2016;1:207–226. doi: 10.1080/23808993.2016.1164013.
12. Wang Y, Jin ZY. Radiomics approaches in gastric cancer: a frontier in clinical decision making. Chin Med J 2019;132:1983–1989. doi: 10.1097/CM9.0000000000000360.
13. Wáng YXJ, Wang X, Wu P, Wang Y, Chen W, Chen H, *et al*. Topics on quantitative liver magnetic resonance imaging. Quant Imaging Med Surg 2019;9:1840–1890. doi: 10.21037/qims.2019.09.18.
14. Padhani AR, Liu G, Koh DM, Chenevert TL, Thoeny HC, Takahara T, *et al*. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. Neoplasia 2009;11:102–125. doi: 10.1593/neo.81328.
15. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, *et al*. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–446. doi: 10.1016/j.ejca.2011.11.036.
16. Guan Y, Li W, Jiang Z, Chen Y, Liu S, He J, *et al*. Whole-lesion apparent diffusion coefficient-based entropy-related parameters for characterizing cervical cancers: initial findings. Acad Radiol 2016;23:1559–1567. doi: 10.1016/j.acra.2016.08.010.
17. Corino VDA, Montin E, Messina A, Casali PG, Gronchi A, Marchianò A, *et al*. Radiomic analysis of soft tissues sarcomas can distinguish intermediate from high-grade lesions. J Magn Reson Imaging 2018;47:829–840. doi: 10.1002/jmri.25791.
18. Barry B, Buch K, Soto JA, Jara H, Nakhmani A, Anderson SW. Quantifying liver fibrosis through the application of texture analysis to diffusion weighted imaging. Magn Reson Imaging 2014;32:84–90. doi: 10.1016/j.mri.2013.04.006.
19. Horowitz JM, Venkatesh SK, Ehman RL, Jhaveri K, Kamath P, Ohliger MA, *et al*. Evaluation of hepatic fibrosis: a review from the society of abdominal radiology disease focus panel. Abdom Radiol (NY) 2017;42:2037–2053. doi: 10.1007/s00261-017-1211-7.
20. Gyamfi MA, Wan YJ. Pathogenesis of alcoholic liver disease: the role of nuclear receptors. Exp Biol Med (Maywood) 2010;235:547–560. doi: 10.1258/ebm.2009.009249.
21. Makol A, Watt KD, Chowdhary VR. Autoimmune hepatitis: a review of current diagnosis and treatment. Hepat Res Treat 2011;2011:390916. doi: 10.1155/2011/390916.
22. Qiu Q, Zhang J, Duan J, Yin Y. EP-2129: the impact of b-values on radiomic features of diffusion-weighted imaging in hepatic cirrhosis. Radiot Oncol 2018;127:S1172–S1173. doi: 10.1016/s0167-8140(18)32438-1.
23. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, *et al*. Computational radiomics system to decode the radiographic phenotype. Cancer Res 2017;77:e104–e107. doi: 10.1158/0008-5472.Can-17-0339.
24. Larue RT, Defraene G, De Ruysscher D, Lambin P, van Elmpt W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. Br J Radiol 2017;90:20160665. doi: 10.1259/bjr.20160665.
25. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, *et al*. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology 2020;295:191145. doi: 10.1148/radiol.2020191145.
26. Shen C, Liu Z, Wang Z, Guo J, Zhang H, Wang Y, *et al*. Building CT radiomics based nomogram for preoperative esophageal cancer patients lymph node metastasis prediction. Transl Oncol 2018;11:815–824. doi: 10.1016/j.tranon.2018.04.005.
27. Boonstra K, Weersma RK, van Erpecum KJ, Rauws EA, Spanier BW, Poen AC, *et al*. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. Hepatology 2013;58:2045–2055. doi: 10.1002/hep.26565.
28. Tayob N, Richardson P, White DL, Yu X, Davila JA, Kanwal F, *et al*. Evaluating screening approaches for hepatocellular carcinoma in a cohort of HCV related cirrhosis patients from the Veteran's Affairs Health Care System. BMC Med Res Methodol 2018;18:1. doi: 10.1186/s12874-017-0458-6.
29. Bahl G, Cruite I, Wolfson T, Gamst AC, Collins JM, Chavez AD, *et al*. Noninvasive classification of hepatic fibrosis based on texture parameters from double contrast-enhanced magnetic resonance images. J Magn Reson Imaging 2012;36:1154–1161. doi: 10.1002/jmri.23759.
30. Yokoo T, Wolfson T, Iwaisako K, Peterson MR, Mani H, Goodman Z, *et al*. Evaluation of liver fibrosis using texture analysis on combined-contrast-enhanced magnetic resonance images at 3.0T. Biomed Res Int 2015;2015:387653. doi: 10.1155/2015/387653.
31. House MJ, Bangma SJ, Thomas M, Gan EK, Ayonrinde OT, Adams LA, *et al*. Texture-based classification of liver fibrosis using MRI. J Magn Reson Imaging 2015;41:322–328. doi: 10.1002/jmri.24536.
32. Park HJ, Lee SS, Park B, Yun J, Sung YS, Shim WH, *et al*. Radiomics analysis of gadoxetic acid-enhanced MRI for staging liver fibrosis. Radiology 2019;290:380–387. doi: 10.1148/radiol.2018181197.
33. Wu Z, Matsui O, Kitao A, Kozaka K, Koda W, Kobayashi S, *et al*. Hepatitis C related chronic liver cirrhosis: feasibility of texture analysis of MR images for classification of fibrosis stage and necroinflammatory activity grade. PLoS One 2015;10:e0118297. doi: 10.1371/journal.pone.0118297.
34. Yasaka K, Akai H, Kunimatsu A, Abe O, Kiryu S. Liver fibrosis: deep convolutional neural network for staging by using gadoxetic acid-enhanced hepatobiliary phase MR images. Radiology 2018;287:146–155. doi: 10.1148/radiol.2017171928.
35. Becker AS, Wagner MW, Wurning MC, Boss A. Diffusion-weighted imaging of the abdomen: impact of b-values on texture analysis features. NMR Biomed 2017;30:e3669. doi: 10.1002/nbm.3669.
36. Zhang J, Qiu Q, Duan J, Gong G, Jiang Q, Sun G, *et al*. Variability of radiomic features extracted from multi-b-value diffusion-weighted images in hepatocellular carcinoma. Transl Cancer Res 2019;8:130–140. doi: 10.21037/tcr.2019.01.14.
37. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges JP, *et al*. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. J Nucl Med 2011;52:369–378. doi: 10.2967/jnumed.110.082404.

How to cite this article: Qiu QT, Zhang J, Duan JH, Wu SZ, Ding JL, Yin Y. Development and validation of radiomics model built by incorporating machine learning for identifying liver fibrosis and early-stage cirrhosis. Chin Med J 2020;133:2653–2659. doi: 10.1097/CM9.0000000000001113