Whole-liver histogram analysis of hepatocyte-specific contrast-enhanced magnetic resonance imaging for predicting progression in patients with cirrhosis

Xu Qiao[1#], Zirui Wang[2#], Xianru Zhang[1], Wei Chen[3], Li Wang[4], Yen-Wei Chen[5]

[1]Department of Biomedical Engineering, School of Control Science and Engineering, Shandong University, Jinan, China; [2]Zhongtai Securities Institute for Financial Studies, Shandong University, Jinan, China; [3]School of Radiology, Shandong First Medical University and Shandong Academy of Medical Sciences, Tai'an, China; [4]Department of Health Management Center, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China; [5]Graduate School of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga, Japan

Contributions: (I) Conception and design: L Wang, X Qiao, YW Chen; (II) Administrative support: X Qiao; (III) Provision of study materials or patients: W Chen, L Wang; (IV) Collection and assembly of data: X Zhang, L Wang, X Qiao; (V) Data analysis and interpretation: Z Wang, X Zhang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Xu Qiao, PhD. Department of Biomedical Engineering, School of Control Science and Engineering, Shandong University, No. 17923, Jingshi Road, Jinan 250061, China. Email: qiaoxu@sdu.edu.cn; Li Wang, PhD. Department of Health Management Center, Shandong Provincial Hospital Affiliated to Shandong First Medical University, No. 324, Jingwuweiqi Road, Huaiyin Distinct, Jinan 250021, China. Email: wangli72977297@126.com.

Background: Liver cirrhosis, as the terminal phase of chronic liver disease fibrosis, is associated with high morbidity and mortality. Traditional methods for assessing liver function, such as clinical scoring systems, offer only a global evaluation and may not accurately reflect regional liver function variations. This study aimed at evaluating the diagnostic potential of whole-liver histogram analysis of gadobenate dimeglumine (Gd-BOPTA)-enhanced magnetic resonance imaging (MRI) for predicting the progression of cirrhosis.

Methods: In this retrospective study, 265 consecutive patients with cirrhosis admitted to the Department of Radiology, Shandong Provincial Hospital Affiliated to Shandong First Medical University from August 2012 to September 2019 were enrolled. After the exclusion criteria were applied, 117 patients (84 males and 33 females) were divided into Child-Pugh A cirrhosis (n=43), Child-Pugh B cirrhosis (n=49), and Child-Pugh C cirrhosis (n=25). After correction for liver signal intensity with the spleen was completed, 19 histogram features of the whole liver were extracted and modeled to evaluate liver function, with the Child-Pugh class being incorporated as a clinical parameter. Receiver operating characteristic (ROC) curves were used to assess the diagnosis capability and determine the optimal cutoffs after a mean follow-up of 42.3±19.1 (range, 8–93) months. The association between significant histogram features and the cumulative incidence of hepatic insufficiency was analyzed with the adjusted Kaplan-Meier curve model.

Results: Among 117 patients (12%), 14 developed hepatic insufficiency through a period of follow-up. Five features, including the median (P<0.01), 90th percentile (P<0.01), root mean squared (P<0.01), mean (P<0.01), and 10th percentile (P<0.05), were significantly different between the groups with and without hepatic insufficiency according to the Kruskal-Wallis test; in the ROC curve analysis, the area under the curve (AUC) of these features was 0.723 [95% confidence interval (CI): 0.653–0.793], 0.722 (95% CI: 0.652–0.792), 0.722 (95% CI: 0.652–0.792), 0.721 (95% CI: 0.651–0.791), and 0.674 (95% CI: 0.600–0.748) after correction, respectively (all P values <0.05). Median, 90th percentile, root mean squared, and mean were found to be significant factors in predicting liver insufficiency. The adjusted Kaplan-Meier curves revealed that patients with a feature level less than the cutoff, as compared to those with a level above the cutoff, showed

a statistically shorter progression-free survival and higher incidences of hepatic insufficiency for significant features of median (cutoff =26.001; 21.28% versus 5.71%; P=0.02), 90th percentile (cutoff =86.263; 20.41% versus 5.88%; P<0.01), root mean squared (cutoff =1,028.477; 19.15% versus 7.14%; P=0.049), and mean (cutoff =27.484; 19.15% versus 7.14%; P=0.049). Patients with a 10th percentile less than −39.811 also showed a higher cumulative incidence of hepatic insufficiency than did those with a value higher than the cutoff (0.18% versus 7.46%; P=0.22).

Conclusions: Whole-liver histogram analysis of Gd-BOPTA-enhanced MRI may serve as a noninvasive analytical method to predict hepatic insufficiency in patients with cirrhosis.

Keywords: Liver cirrhosis; gadobenic acid; magnetic resonance imaging (MRI); spleen-guided normalization; histogram analysis

Introduction

Liver cirrhosis, as the terminal phase of chronic liver disease fibrosis, is associated with high morbidity and mortality (1). The advancement of cirrhosis from a compensated state to a decompensated state progresses at an annual rate of 5–11% and is marked with severe complications, including variceal bleeding, ascites, jaundice, and encephalopathy (2,3). Compared with the compensated state, decompensated cirrhosis is characterized by a lower quality of life and a poorer median survival of ~2 versus >12 years (4). Fortunately, cirrhosis progression can be reversed in a variety of liver diseases with close follow-up and the administration of effective therapies—specifically antiviral therapy—at the early stage (5,6). Identifying patients at a high risk of disease progression may be beneficial for clinical practice in developing varied monitoring schedules and tailored therapeutic strategies. Thus, it is critical to evaluate liver function and predict the progression of adverse outcomes for patients with cirrhosis (7,8).

Traditional methods for measuring liver function include the indocyanine green clearance test and clinical scoring systems, such as Child-Pugh score and the Model for End-Stage Liver Disease score (9-12). However, these measurements only offer a global assessment, and regional hepatic function may differ substantially across individuals with chronic hepatic disorders, which could lead to potentially inaccurate assessments of liver-related conditions. Noninvasive methods based on imaging have been investigated and proven to be comparable with conventional clinical measurements (13,14). Therefore, the quantitative assessment of regional liver function is being increasingly studied (15-17).

Gadobenate dimeglumine (Gd-BOPTA) is widely accepted and commonly used as a cost-effective hepatocyte-specific contrast agent in both academic research and clinical practice (18). Several studies have confirmed that hepatobiliary-specific contrast agent-enhanced hepatic parenchyma magnetic resonance imaging (MRI) is an effective noninvasive method for assessing liver function and predicting the prognosis of patients with conditions such as liver fibrosis (18-22). One study reported that the liver ratio enhancement calculated using Gd-BOPTA-enhanced MRI was clinically effective in evaluating liver function in a noninvasive manner, as the intracellular uptake of Gd-BOPTA decreased with impaired liver function (23). Therefore, hepatocyte-specific contrast-enhanced MRI is widely regarded as providing superior estimation of liver function from both global and regional perspectives (24). However, the routine Gd-BOPTA-enhanced liver MRI measurements have focused predominantly on calculating the mean value of signal intensity (SI)—a method that fails to consider the inherent spatial heterogeneity of the tissue (25).

Histogram analysis using MRI reflects the SI distribution and quantifies the heterogeneity of pathological changes (26). Histogram parameters, such as skewness and kurtosis, mitigate the loss of diagnostic information that often occurs with the use of a single index or two-dimensional (2D) imaging methods. Given the uneven distribution of liver function (27), histogram analysis of whole-liver MRI SI has great potential to characterize both the overall and regional characteristics and better depict the liver function. To our
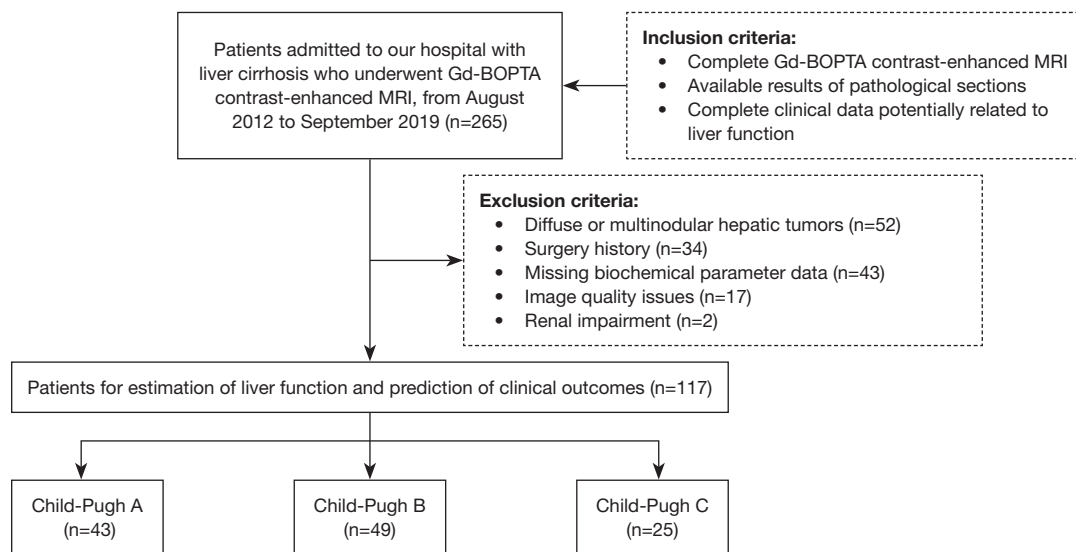
Figure 1 Flow diagram of the study population. Gd-BOPTA, gadobenate dimeglumine; MRI, magnetic resonance imaging.

knowledge, the utility of whole-liver histogram analysis of hepatocyte-specific contrast-enhanced MRI in patients with liver cirrhosis has not been extensively reported. Therefore, this study aimed to investigate the diagnostic value of whole-liver histogram analysis of Gd-BOPTA-enhanced MRI in evaluating liver function and predicting progression in patients with cirrhosis. We present this article in accordance with the STARD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-24-109/rc).

Methods

Study population

This study was registered in the Chinese Clinical Trials Registry Center (https://www.chictr.org.cn/indexEN.html; registration number: ChiCTR2300072395) and was approved by the ethics committee of Shandong Provincial Hospital Affiliated to Shandong First Medical University (SWYX: No. 2021-399). This study was conducted in accordance with the guidelines of the Declaration of Helsinki (as revised in 2013). Due to the retrospective nature of this study, the requirement for informed consent was waived by the ethics committee. In this retrospective study, 265 consecutive patients with cirrhosis who underwent Gd-BOPTA-enhanced MRI admitted to the Department of Radiology, Shandong Provincial Hospital Affiliated to Shandong First Medical University from

August 2012 to September 2019 were enrolled (the study execution time was from October 10, 2021, to May 1, 2022). The inclusion criteria were as follows: (I) complete Gd-BOPTA-enhanced MR images; (II) results of pathological sections available [hematoxylin and eosin (HE) staining and immunohistochemical staining]; and (III) complete clinical data potentially related to liver function, including age, gender, Child-Pugh scores, and other clinical indexes. Meanwhile, the exclusion criteria were the following: (I) patients with hepatic tumors (n=52), (II) a history of surgery (n=34), (III) renal impairment (n=2), (IV) s lack of biochemical metric information (n=43), and (V) excessive motion artifact or absence of hepatobiliary phase images (n=17). Finally, a cohort of 117 patients (84 males and 33 females) were included in the study (*Figure 1*). Clinical characteristics potentially related to liver function were systematically retrieved from the electronic medical records. According to surgical resection or an amalgam of clinical manifestations and radiological characteristics, individuals with hepatic cirrhosis were categorized as Child-Pugh A cirrhosis (n=43), Child-Pugh B cirrhosis (n=49), or Child-Pugh C cirrhosis (n=25). The clinical data for all patients are summarized in *Table 1*.

Patients were monitored for the development of hepatic insufficiency with a period of follow-up. The onset of hepatic insufficiency was defined as the presence of gross ascites, uncontrolled hepatic encephalopathy, jaundice, or death.

Table 1 Characteristics of the study population

Characteristic	Value (N=117)
Demographic data	
Gender: male/female	84/33
Age (years)*	
Male	49.7±9.4 (30–83)
Female	54.5±11.3 (33–72)
Causes of liver cirrhosis, n (%)	
Hepatitis B virus	95 (81)
Hepatitis C virus	3 (3)
Alcohol intake	26 (22)
Autoimmune hepatitis	6 (5)
NAFLD	2 (2)
Cryptogenic	10 (9)
Antiviral treatment	94 (80)
Serum markers*	
Total bilirubin (µmol/L)	68.1±95.6 (8.5–581.5)
Albumin (g/L)	35.4±6.0 (16.4–50.0)
Prothrombin time (s)	15.4±2.8 (10.7–23.9)
Prothrombin time (INR)	1.3±0.3 (0.9–2.5)
Creatinine (µmol/L)	68.0±18.2 (30.0–146.0)
Aspartate aminotransferase (U/L)	63.7±40.4 (15.0–272.0)
Alanine aminotransferase (U/L)	60.0±46.4 (10.0–269.0)
Glutamyltransferase (U/L)	76.6±65.7 (10.0–335.0)
Child-Pugh class, n (%)	
A	43 (36.8)
B	49 (41.9)
C	25 (21.4)
MELD score, n (%)	
≤10	52 (44.4)
11–20	55 (47.0)
>20	10 (8.5)
Hepatic insufficiency, n (%)	14 (12)
Intractable ascites, n (%)	5 (4)
Hepatic encephalopathy, n (%)	6 (5)
Mortality, n (%)	8 (7)
Liver transplantation, n (%)	1 (1)

*, data are the mean ± standard deviation with ranges in parentheses. NAFLD, nonalcoholic fatty liver disease; INR, international normalized ratio; MELD, model for end-stage liver disease.

MRI acquisition

MRI was performed using a 3T MR imaging scanner (MAGNETOM Verio or Prisma, Siemens Healthineers, Erlangen, Germany) outfitted with a phased-array body coil. The axial fat-saturated T1-weighted volumetric interpolated breath-hold examination sequence was obtained prior to and 90 minutes following the administration of contrast material during the hepatobiliary phase. The MRI settings were as follows: repetition time, 3.31 ms; echo time, 1.3 ms; slice thickness, 3 mm; number of partitions, 72; field of view, 380×308 mm; acceleration factor, 1; flip angle, 9°; and bandwidth, 450 Hz/pixel. For enhanced MR imaging, Gd-BOPTA was injected in a single dose of 0.05 mmol/kg (0.1 mL/kg) body weight intravenously, which was followed by a saline flush of 20 mL.

Image correction process

(I)   Data loading: all original MRI images of the hepatobiliary phase were loaded onto ITK-SNAP software (http://www.itksnap.org/) (28).

(II) Region of interest drawing: the hepatic Gd-BOPTA-enhanced MR images were read independently by two experienced radiologists (observer 1 and observer 2, with 10 and 15 years of abdominal imaging experience, respectively) who were blinded to the clinical information. The regions of interest (ROIs), including the whole liver and spleen, were manually delineated slice by slice in the axial plane. The visible blood vessels and imaging artifacts were carefully excluded, and thus the ROI markers did not extend to the edge of the areas. The final three-dimensional (3D)-segmented volumes created on the original images were automatically propagated (*Figure 2*).

(III) Image correction: to minimize the errors resulting from scanners, three approaches for correcting the SI of each patient's liver were assessed, with the patient's spleen being used as an internal reference. First, we regarded each patient's liver and spleen SI values read from the 3D MRI image as two 3D tensors, denoted as SI_{liver} and SI_{spleen}. The mean and standard deviation (SD) were calculated as the global mean and global SD values. $\boldsymbol{J}$ was defined as a 3D tensor the same size as SI_{liver}, with all elements equal to 1. The widely adopted liver-to-spleen contrast ratio (LSC) is denoted in the format of Eq. [1] (29). Expanding upon this concept,
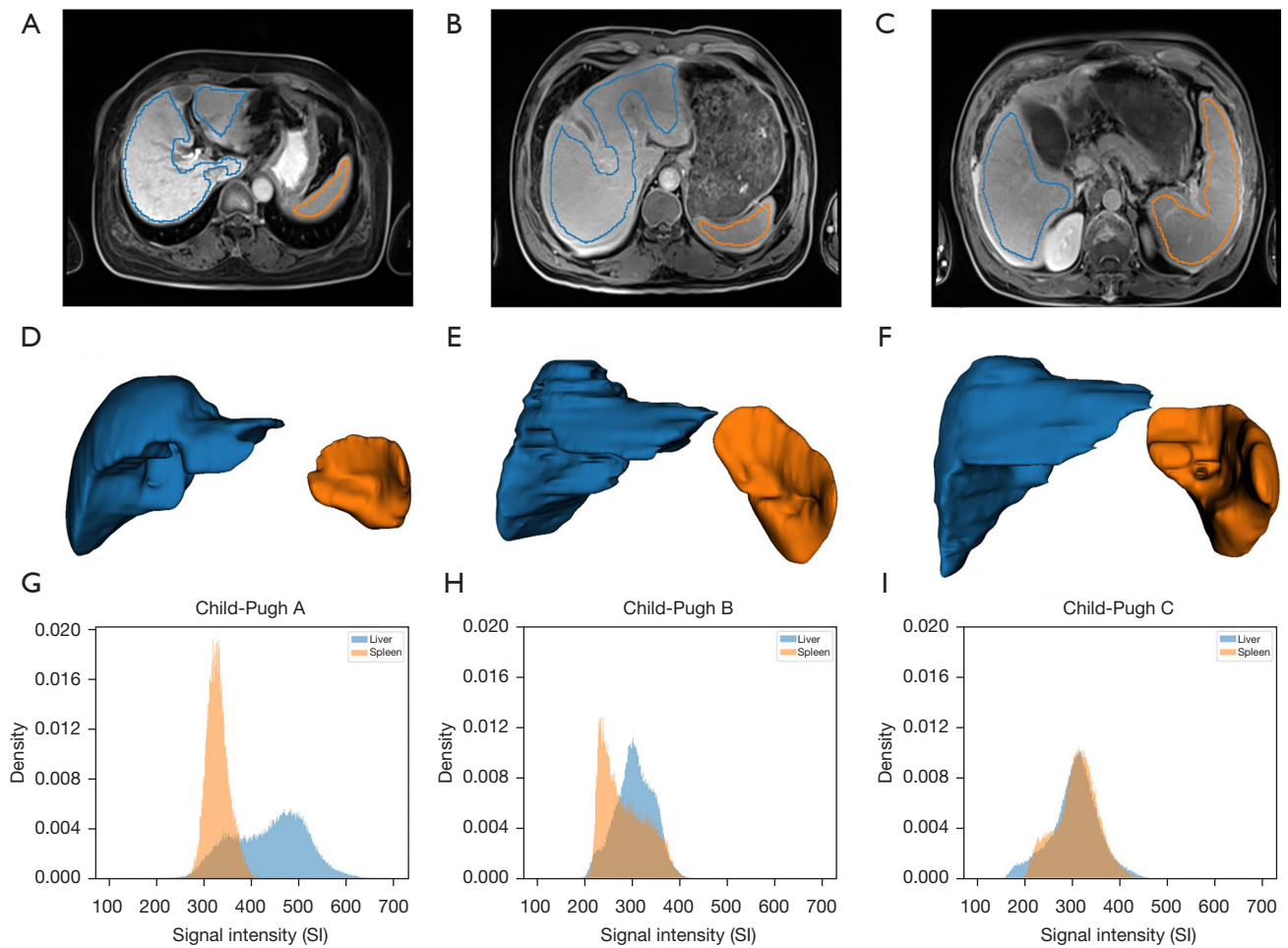
6076

Qiao et al. MRI histogram analysis in patients with cirrhosis

Figure 2 Representative MRI of patients with different Child-Pugh grades. (A-C) One slice of the hepatocyte-specific contrast-enhanced MRI with manual labels. (D-F) Corresponding 3D segmentation. (G-I) Density histograms of the whole liver and spleen signal intensity. The blue represents the liver region, and the orange represents the spleen. MRI, magnetic resonance imaging; 3D, three-dimensional.

the three corrected signal intensities were given by the three subsequent equations (Eqs. [2–4]) (30). After our correction process, the resulting outputs of the corrected signal intensities (CSI), denoted as CSI_1, CSI_2, CSI_3, retained the form of the image tensors.

$$LSC = \frac{Mean(SI_{liver})}{Mean(SI_{spleen})} \quad [1]$$

$$CSI_1 = SI_{liver} - Mean(SI_{spleen}) \cdot J \quad [2]$$

$$CSI_2 = \frac{1}{SD(SI_{spleen})} \cdot (SI_{liver} - Mean(SI_{spleen}) \cdot J) \quad [3]$$

$$CSI_3 = \frac{1}{Mean(SI_{spleen})} \cdot (SI_{liver} - Mean(SI_{spleen}) \cdot J) \quad [4]$$

For convenience, we referred to CSI_1, CSI_2, and CSI_3 as mean-corrected, SD-corrected, and mean-mean-corrected SI values, respectively.

Feature extraction

Histogram feature extraction was performed using the Pyradiomics 3.0 package (Python 3.7.9, Python Software Foundation, Wilmington, DE, USA) (31). Before and after the three different methods of correction were applied to images, a total of 19 histogram features were calculated, including 10th percentile, 90th percentile, energy, entropy, interquartile range, kurtosis, maximum, mean, mean absolute deviation, median, minimum, range, robust mean absolute deviation, root mean squared, skewness, SD, total energy, uniformity, and variance. *Table 2* provides a concise

Table 2 Description of histogram features

Histogram feature	Description
10th percentile	Only 10% of the data points less than this SI value
90th percentile	Only 10% of the data points more than this SI value
Energy	Measure of the squared magnitude of SI values within all data points
Entropy	Measure of the inherent randomness in the SI values within all data points
Interquartile range	Measure of the spread of the distribution of SI values, defined as the difference between the 75th and 25th percentile
Kurtosis	Measure of the "peakedness" of the distribution of SI values within all data points
Maximum	The maximum SI value within all data points
Mean	The average SI value within all data points
Mean absolute deviation	Mean distance of all SI values from the mean value of the image array
Median	The SI value below 50% of all data points
Minimum	The minimum SI value within all data points
Range	Measure of the difference between the highest and lowest SI values
Robust mean absolute deviation	The average separation of all SI values from the average value calculated on the subset of the image array with SI values in between, or equivalent to the 10th and 90th percentile
Root mean squared	The square root of the average of all the squares of SI values
Skewness	Measure of the asymmetry of the distribution of SI values around the average value
Standard deviation	Standard deviation measure for the amount of variation
Total energy	The value of the energy feature scaled by the volume of the voxel in cubic millimeters
Uniformity	Measure of the homogeneity in the SI values within all data points
Variance	The squared distances of each SI value from the mean

SI, signal intensity.

description of each extracted feature (32).

Statistical analysis

All statistical analyses were performed with R software version 4.0.2 (The R Foundation for Statistical Computing) and SPSS version 26.0 software (IBM Corp., Armonk, NY, USA). Four commonly used machine learning classifiers, including support vector machine (SVM), k-nearest neighbor (KNN; with k=5), random forest (RF; with 100 randomized trees), and Gaussian naïve Bayes (GNB), were applied with stratified fivefold cross-validation. Using cross-validation in lieu of a separate test set could ensure robust model evaluation in the context of our dataset's size, which was justified within the context of our study. Consequently, direct comparisons between classifiers were not conducted. Discrimination ability was qualified using receiver operating characteristic (ROC) curve analysis and the area under the curve (AUC). Features were min–max normalized to 0 and 1. The performance of classifiers were evaluated according to whether the histogram features could effectively distinguish individuals classified as Child-Pugh A from those classified as Child-Pugh B or Child-Pugh C. Features computed from original images and corrected images were also compared to assess the efficacy of correction. The approach with the best classification performance was selected as the final correction method, and the corresponding histogram features were used to analyze the development of clinical outcomes.

The characteristics of participants across various groups were assessed with the Student's *t*-test for continuous variables and the chi-squared test for categorical variables. Data following a normal distribution were evaluated with the one-way analysis of variance test, while nonnormally
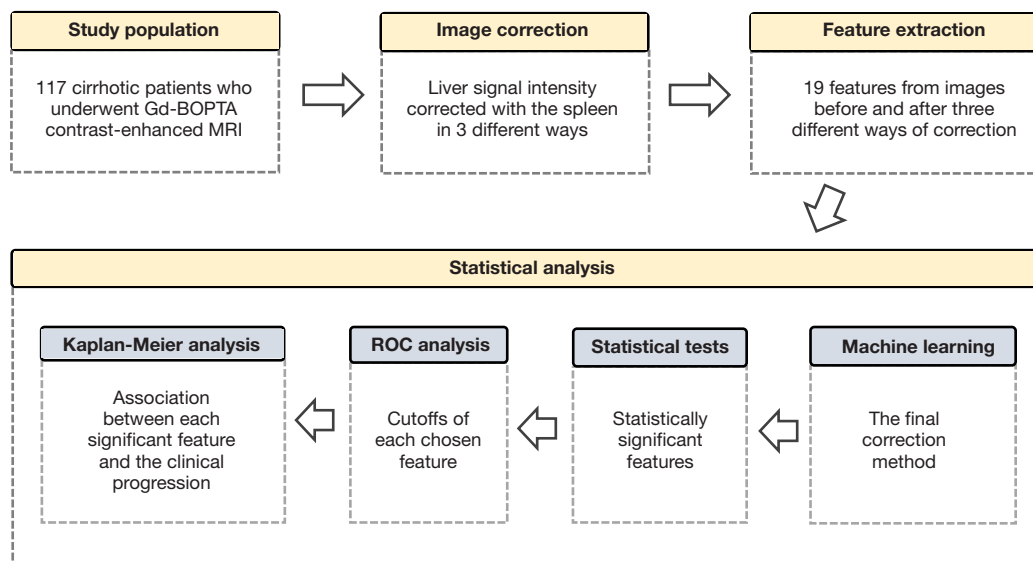
6078

Qiao et al. MRI histogram analysis in patients with cirrhosis

Figure 3 A workflow of the study. Gd-BOPTA, gadobenate dimeglumine; MRI, magnetic resonance imaging; ROC, receiver operating characteristic.

distributed data were examined using the Kruskal-Wallis test. All 19 features were compared separately for images before and after correction via the Kruskal-Wallis test. For each histogram feature with a two-sided P value <0.05 indicating statistical significance, nonparametric ROC curve analysis was carried out individually to evaluate the diagnostic capability and determine the optimal cutoff. The cutoff was defined as the threshold of the equal error rate point on ROC curves, in which the false-positive rate equaled the false-negative rate.

According to the cutoff of a significant feature, patients were classified into two groups to determine the relationship between this feature and clinical progression. The endpoint event was the development of hepatic insufficiency. The adjusted Kaplan-Meier curve model was applied to analyze the cumulative incidence of hepatic insufficiency between the two groups.

A workflow of the study process is shown in *Figure 3*.

Results

Clinical characteristics of patients

Among 117 patients (mean age, 51.05±10.14 years; 33 females) included in the final study, 43 (36.8%) were grouped in the Child-Pugh A, 49 (41.9%) in the Child-Pugh B, and 25 (21.4%) in the Child-Pugh C. The clinical data for all patients are summarized in *Table 1*. After the

follow-up [42.3±19.1 (8–93) months], hepatic insufficiency occurred in 14 patients (12%), resulting in 8 deaths.

Performance of classification models using histogram features

The ROC curves of four classifiers built on 19 histogram features before and after correction are shown in *Figure 4*. For four classifiers built on uncorrected image histogram features, the average AUC values of fivefold cross-validation spanned from 0.83 [95% confidence interval (CI): 0.75–0.91] to 0.87 (95% CI: 0.81–0.94). Four classifiers built on mean-corrected (CSI_1) features obtained the highest AUCs, with a range from 0.89 (95% CI: 0.82–0.95) to 0.92 (95% CI: 0.87–0.97). Histogram features extracted from SD-corrected (CSI_2) and mean-mean–corrected (CSI_3) images achieved AUCs ranging from 0.84 (95% CI: 0.78–0.92) to 0.87 (95% CI: 0.80–0.94) and from 0.82 (95% CI: 0.75–0.90) to 0.87 (95% CI: 0.79–0.93), respectively. Therefore, subsequent analyses employed the features derived from the mean-corrected (CSI_1) method.

For different classifiers and feature groups, the comparisons of accuracy, sensitivity, and specificity are listed in *Table 3*. The classification models built on mean-corrected (CSI_1) features outperformed those of the others, with substantially higher accuracy and sensitivity and comparable specificity.
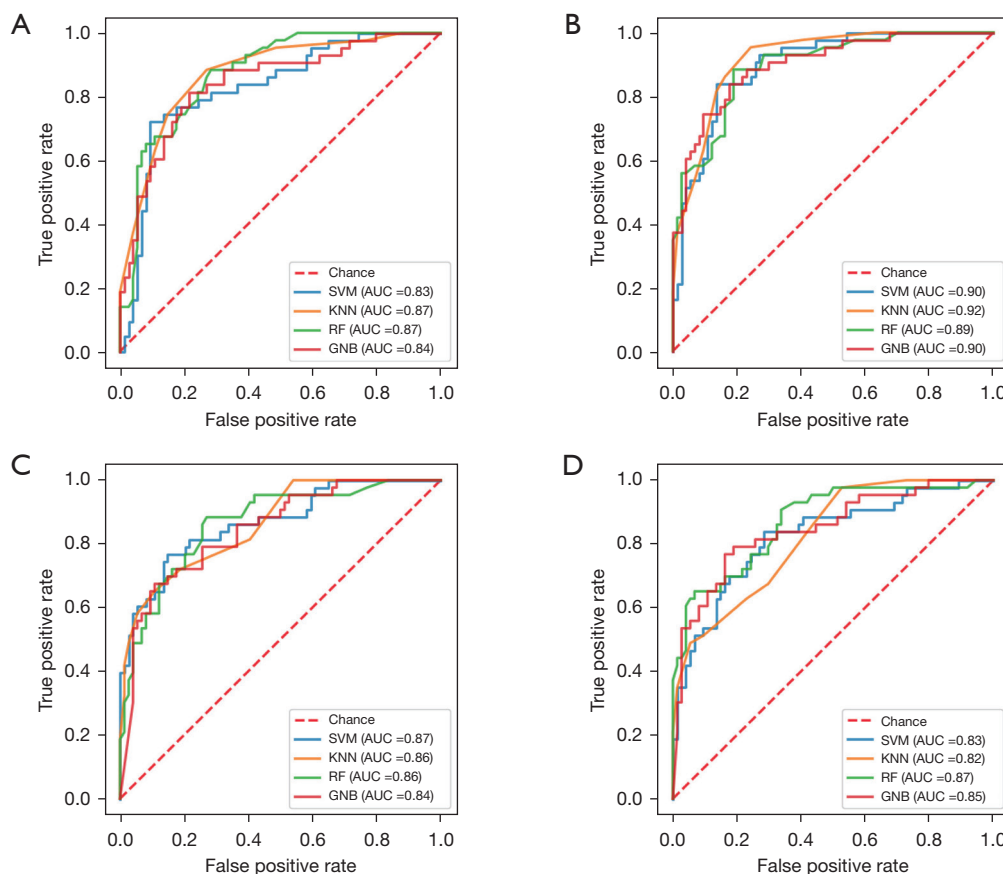
Figure 4 Receiver operating characteristic curve of four classifiers with (A) uncorrected, (B) mean-corrected, (C) standard deviation-corrected, and (D) mean-mean-corrected histogram features. SVM, support vector machine; AUC, area under the curve; KNN, k-nearest neighbor; RF, random forest; GNB, Gaussian naïve Bayes.

Predictive values of histogram features in hepatic insufficiency

As listed in *Table 4*, five features were significantly different between the groups with or without hepatic insufficiency according to the Kruskal-Wallis test: median (P<0.01), 90th percentile (P<0.01), root mean squared (P<0.01), mean (P<0.01), and 10th percentile (P<0.05). The nonparametric ROC analyses demonstrated the association between each of those five features before and after mean correction and the development of hepatic insufficiency (*Figure 5*). In the comparison of these histogram features extracted from original and mean-corrected images, the diagnostic ability of five features improved after correction. Specifically, the AUCs for median, 90th percentile, root mean squared, mean , and 10th percentile were 0.723 (95% CI: 0.653–0.793), 0.722 (95% CI: 0.652–0.792), 0.722 (95% CI: 0.652–0.792), 0.721 (95% CI: 0.651–0.791) and 0.674 (95%

CI: 0.600–0.748), respectively, for the corrected images; meanwhile, for the original images, the AUCs were 0.583 (95% CI: 0.504–0.662), 0.59 (95% CI: 0.511–0.669), 0.58 (95% CI: 0.501–0.659), 0.58 (95% CI: 0.501–0.659), and 0.527 (95% CI: 0.447–0.607), respectively (all P values <0.05).

To clarify the association of mean-corrected histogram features and clinical outcomes of hepatic insufficiency, the cutoff of each significant feature at the equal error point was calculated, and the total cohort was then categorized into two groups according to the cutoff. The adjusted Kaplan-Meier curves of how histogram features were correlated with the incidence of hepatic insufficiency are shown in *Figure 6*. Patients with a feature level less than the cutoff, as compared to those with a level above the cutoff, showed a statistically shorter progression-free survival and higher incidences of hepatic insufficiency for significant features

Table 3 Classification performance of four classifiers for features before and after correction

Correction	Classifier	Accuracy	Sensitivity	Specificity	AUC (95% CI)
Uncorrected	SVM	0.83	0.72	0.89	0.83 (0.75–0.91)
	KNN	0.79	0.63	0.89	0.87 (0.81–0.94)
	RF	0.79	0.67	0.86	0.87 (0.80–0.93)
	GNB	0.79	0.72	0.82	0.84 (0.77–0.92)
Mean corrected	SVM	0.81	0.70	0.88	0.90 (0.84–0.95)
	KNN	0.80	0.63	0.91	0.92 (0.87–0.97)
	RF	0.79	0.72	0.84	0.89 (0.82–0.95)
	GNB	0.84	0.74	0.89	0.90 (0.84–0.96)
SD corrected	SVM	0.82	0.58	0.96	0.87 (0.80–0.94)
	KNN	0.81	0.58	0.95	0.86 (0.79–0.93)
	RF	0.80	0.67	0.88	0.86 (0.79–0.93)
	GNB	0.81	0.67	0.89	0.84 (0.78–0.92)
Mean-mean corrected	SVM	0.74	0.53	0.86	0.83 (0.75–0.91)
	KNN	0.76	0.51	0.91	0.82 (0.75–0.90)
	RF	0.79	0.65	0.86	0.87 (0.79–0.93)
	GNB	0.79	0.60	0.89	0.85 (0.77–0.92)

AUC, area under the curve; CI, confidence interval; SVM, support vector machine; KNN, k-nearest neighbor; RF, random forest; GNB, Gaussian naïve Bayes; SD, standard deviation.

Table 4 Results of Kruskal-Wallis test for the mean-corrected features

Feature	H-statistic	P value
Median	7.27	0.007
90th percentile	7.22	0.007
Root mean squared	7.22	0.007
Mean	7.18	0.007
10th percentile	4.44	0.035
Interquartile range	3.48	0.062
Robust mean absolute deviation	3.26	0.071
Mean absolute deviation	2.36	0.124
Kurtosis	2.21	0.137
Maximum	1.87	0.171
Minimum	1.63	0.202
Standard deviation	1.59	0.208
Variance	1.59	0.208
Skewness	1.30	0.253

Table 4 (continued)

Table 4 (continued)

Feature	H-statistic	P value
Uniformity	1.27	0.260
Energy	1.03	0.310
Total energy	1.03	0.310
Entropy	0.650	0.420
Range	0.376	0.540

of median (cutoff =26.001; 21.28% versus 5.71%; P=0.02), 90th percentile (cutoff =86.263; 20.41% versus 5.88%; P<0.01), root mean squared (cutoff =1,028.477; 19.15% versus 7.14%, P=0.049), and mean (cutoff =27.484; 19.15% versus 7.14%; P=0.049). Patients with a 10th percentile less than −39.811 also showed higher cumulative incidences of hepatic insufficiency than did those with a 10th percentile greater than the cutoff (0.18% versus 7.46%; P=0.22); however, the difference in survival rates between the cohorts did not reach statistical significance.
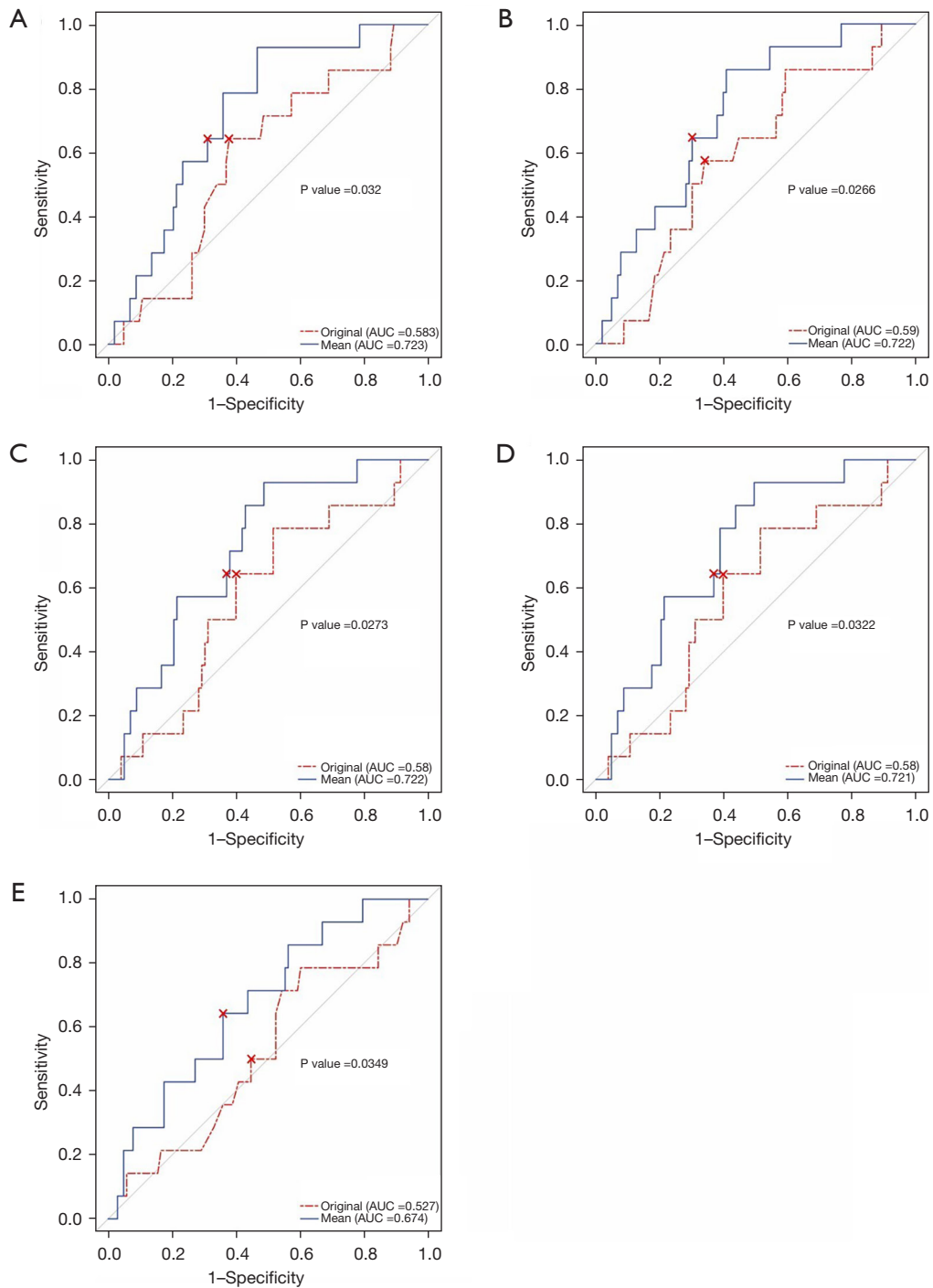
Figure 5 Receiver operating characteristic curves of significant features before (original) and after mean correction (mean) for differentiating patients with hepatic insufficiency. (A) Median, (B) 90th percentile, (C) root mean squared, (D) mean, and (E) 10th percentile. Cutoffs are marked in red crosses. AUC, area under the curve.
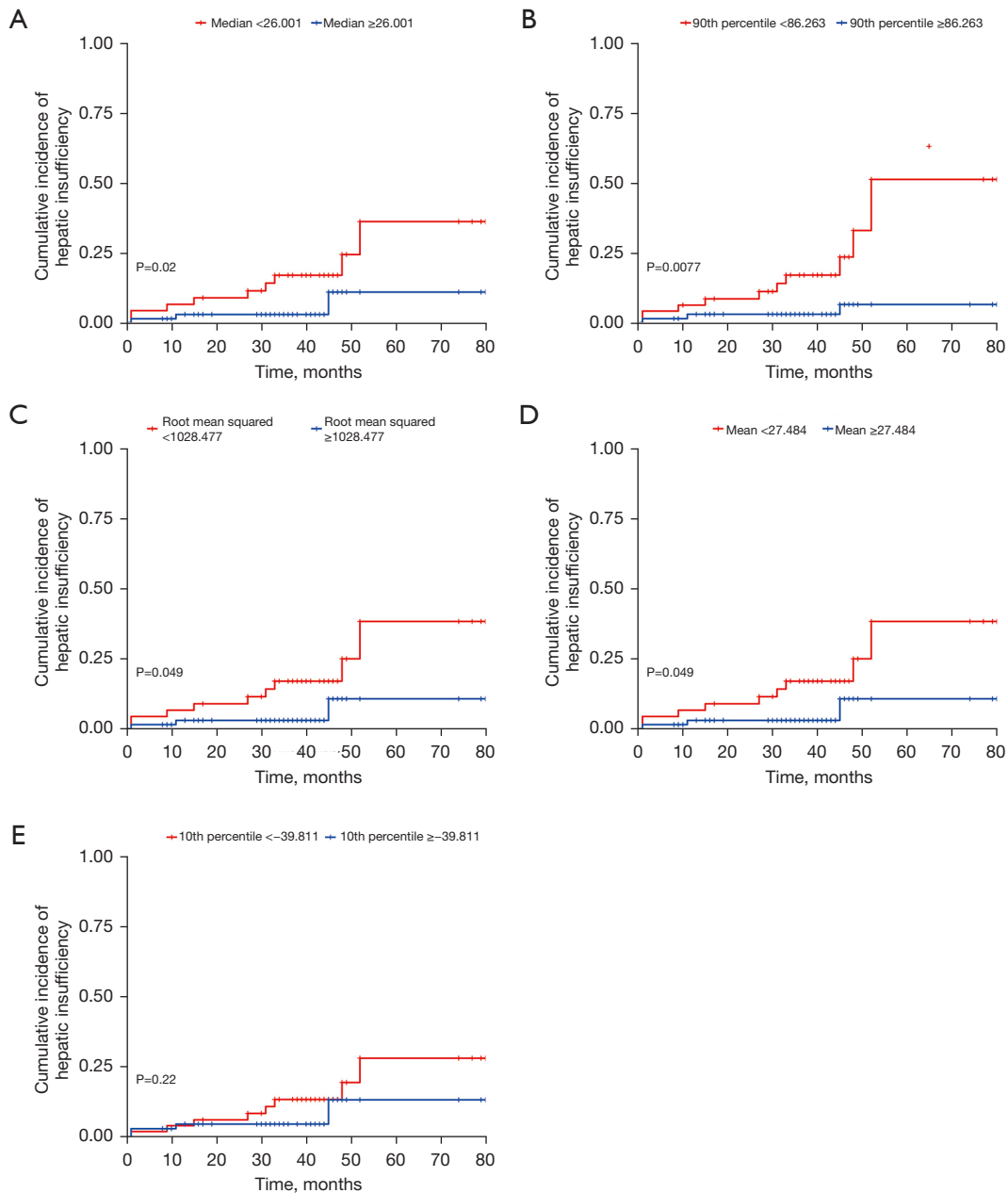
Quant Imaging Med Surg 2024;14(8):6072-6086 | https://dx.doi.org/10.21037/qims-24-109

Figure 6 Adjusted Kaplan-Meier curve showing the cumulative incidences of hepatic insufficiency in (A) 47 patients with median <26.001 compared with 70 patients with median ≥26.001, (B) 49 patients with 90th percentile <86.263 compared with 68 patients with 90th percentile ≥86.263, (C) 47 patients with root mean squared <1,028.477 compared with 70 patients with root mean squared ≥1,028.477, (D) 47 patients with mean <27.484 compared with 70 patients with mean ≥27.484, and (E) 50 patients with 10th percentile <−39.811 compared with 67 patients with 10th percentile ≥−39.811.

Discussion

In this study, we employed a radiomics approach that used whole-liver histogram analysis of Gd-BOPTA-enhanced MRI to evaluate liver function and predict progression in patients with cirrhosis. Radiomics, a rapidly evolving field in medical imaging, involves the extraction and analysis of a

large number of quantitative features from medical images to uncover hidden information and patterns that may not be discernible to the naked eye (33). Histogram features, including median, 90th percentile, root mean squared, mean, and 10th percentile, were significantly different between groups with and without hepatic insufficiency. The ROC curve demonstrated the diagnostic ability of these five features was improved after correction of SI via the subtraction of the mean value of the spleen. Median, 90th percentile, root mean squared, and mean were four significant factors for predicting liver insufficiency. To verify the performance of each feature, the simplest method was employed: each of these features was used to independently to predict the prognosis of cirrhosis. The median, without being skewed by outliers, was found capable of providing a robust measure. Meanwhile, the root mean squared was more sensitive to outlier regions to the mean and median, thereby potential providing information on fibrotic development. Finally, the 90th percentile was able to capture the higher-intensity values indicative of contrast enhancement or tissue heterogeneity, which are often associated with pathological changes. In liver cirrhosis, these features might reveal the presence of fibrotic nodules or inflammatory activities, which frequently exhibit heightened contrast compared to healthy tissue (34).

Histogram analysis involves the analysis of the SI distribution of ROIs. Kim *et al.* stated that histograms would be more suitable for evaluating liver function and achieve more consistent quantification of SI distribution compared to the degree of enhancement (35). Studies on histogram analysis have examined T1 maps (36) and hepatocyte-specific contrast-enhanced MRI (25,35) with limited ROIs drawn on only one representative slice for predicting liver function and fibrosis; however, few studies have attempted to predict liver prognosis in patients with cirrhosis. In this study, the histogram analysis was performed on the whole liver region to better depict the uneven SI distribution of the liver.

Our study focused not on a direct comparison of the performance of four widely used machine learning classifiers but rather on the validation and efficacy of a proposed calibration method. We evaluated the classification effectiveness by comparing various feature-correction approaches. Specifically, we distinguished patients diagnosed with Child-Pugh A cirrhosis from those with more advanced Child-Pugh B or C cirrhosis, as illustrated in *Figure 4* and detailed in *Table 3*. The model built on mean-corrected features outperformed others, with much higher AUCs spanning from 0.89 (95% CI: 0.82–0.95) to 0.92 (95% CI: 0.87–0.97). Its accuracy and sensitivity surpassed those of others on each classifier while its specificity was also not substantially inferior. The performance of the diagnostic ability of five features improved after correction as compared to the original images. Therefore, the mean-corrected method was employed since the mean of the spleen region for calibration is more meaningful for quantitatively evaluating the liver function. This result is similar to the study of Park *et al.* (37), in which SI normalization was performed based on the spleen, and a higher predictive accuracy was achieved for liver fibrosis staging. Nonetheless, the study employed a singular calibration approach. Our evaluation infers that the intrinsic relative nature of MRI SI values may render normalization by division with the spleen's SD ($SD[SI_{spleen}]$) suboptimal for discernible efficacy. In our study, the median, 90th percentile, root mean squared, and mean from the whole-liver histogram analysis were clearly shown to be four significant factors in predicting hepatic insufficiency. For each feature, patients with feature levels less than the cutoffs showed higher incidences of hepatic insufficiency, providing crucial prognostic information for patients with cirrhosis.

Furthermore, notwithstanding the absence of patients with severe hepatic iron overload within our study cohort, the potential impact of mild iron overload on T1-weighted MRI SI warrants consideration. To quantify liver iron concentration (LIC), three MRI methods have been validated: SI ratio and R2- and R2*-based relaxometry (38). Including iron overload as a quantifiable metric is likely to bolster the precision of prognostic models.

In the future, our approach holds the potential to serve as a supplementary method alongside standard clinical measures such as the Child-Pugh score or indocyanine green retention test at 15 minutes (ICG-15) for the management of diseases in patients with cirrhosis. By integrating automatic segmentation technology (39), our method may enable the automated transformation of patient MRI data into corrected features. Through using cutoff standards derived from the analysis of a large sample set, it would be possible to assess the risk of disease progression in individual patients. The adoption of this methodology could lead to the generation of automated recommendations that support clinical decisions, providing a more robust prognostic tool that complements traditional, more subjective assessments.

Limitations

The limitations to this study were as follows. Only 117 samples from a single hospital were collected in this retrospective study. Given the imbalanced distribution of samples among various Child-Pugh scores, the classification of patients was examined by grouping those diagnosed as Child-Pugh A separately from those diagnosed as Child-Pugh B or Child-Pugh C. Our experimental results still need further validation on multicenter external datasets. Moreover, while using cross-validation in lieu of a separate test set aimed at discerning the most effective correction method for optimal classification performance was deemed appropriate, this involved some limitations. Furthermore, the potential impact of mild iron overload on T1-weighted MR SI warrants consideration. In addition, the manual delineation of ROIs was laborious, and automatic segmentation algorithms that can relieve doctors from time-consuming labeling should be investigated.

Conclusions

This study provides evidence that Gd-BOPTA-enhanced hepatobiliary phase MRI histogram analysis holds potential for evaluating hepatic function and prognosing the progression in patients with cirrhosis. Our technique for correcting liver SIs relative to the spleen has enhanced the diagnostic accuracy of key imaging features. This approach shows considerable promise as a noninvasive technique for the clinical evaluation and management of cirrhotic conditions.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-24-109/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-24-109/coif). X.Q. reports that this research was supported in part by the Shandong Provincial Natural Science Foundation (Nos. ZR2021MF057 and ZR2020MH285) and the Natural Science Foundation of China (No. U1806202). The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and was approved by the ethics committee of Shandong Provincial Hospital Affiliated to Shandong First Medical University (SWYX: No. 2021-399). Due to the retrospective nature of this study, the requirement for informed consent was waived by the ethics committee.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Ginès P, Krag A, Abraldes JG, Solà E, Fabrellas N, Kamath PS. Liver cirrhosis. Lancet 2021;398:1359-76.
2. Fleming KM, Aithal GP, Card TR, West J. The rate of decompensation and clinical progression of disease in people with cirrhosis: a cohort study. Aliment Pharmacol Ther 2010;32:1343-50.
3. D'Amico G, Morabito A, D'Amico M, Pasta L, Malizia G, Rebora P, Valsecchi MG. Clinical states of cirrhosis and competing risks. J Hepatol 2018;68:563-76.
4. D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. J Hepatol 2006;44:217-31.
5. Ellis EL, Mann DA. Clinical evidence for the regression of liver fibrosis. J Hepatol 2012;56:1171-80.
6. Ramachandran P, Iredale JP, Fallowfield JA. Resolution of liver fibrosis: basic mechanisms and clinical relevance. Semin Liver Dis 2015;35:119-31.
7. Zipprich A, Kuss O, Rogowski S, Kleber G, Lotterer E,

Seufferlein T, Fleig WE, Dollinger MM. Incorporating indocyanin green clearance into the Model for End Stage Liver Disease (MELD-ICG) improves prognostic accuracy in intermediate to advanced cirrhosis. Gut 2010;59:963-8.

8. Garcia-Tsao G, Lim JK; Members of Veterans Affairs Hepatitis C Resource Center Program. Management and treatment of patients with cirrhosis and portal hypertension: recommendations from the Department of Veterans Affairs Hepatitis C Resource Center Program and the National Hepatitis C Program. Am J Gastroenterol 2009;104:1802-29.

9. Vos JJ, Wietasch JK, Absalom AR, Hendriks HG, Scheeren TW. Green light for liver function monitoring using indocyanine green? An overview of current clinical applications. Anaesthesia 2014;69:1364-76.

10. Harrison SA, Oliver D, Arnold HL, Gogia S, Neuschwander-Tetri BA. Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. Gut 2008;57:1441-7.

11. Cholongitas E, Papatheodoridis GV, Vangeli M, Terreni N, Patch D, Burroughs AK. Systematic review: The model for end-stage liver disease--should it replace Child-Pugh's classification for assessing prognosis in cirrhosis? Aliment Pharmacol Ther 2005;22:1079-89.

12. Botta F, Giannini E, Romagnoli P, Fasoli A, Malfatti F, Chiarbonello B, Testa E, Risso D, Colla G, Testa R. MELD scoring system is useful for predicting prognosis in patients with liver cirrhosis and is correlated with residual liver function: a European study. Gut 2003;52:134-9.

13. Lee SS, Park SH, Kim HJ, Kim SY, Kim MY, Kim DY, Suh DJ, Kim KM, Bae MH, Lee JY, Lee SG, Yu ES. Non-invasive assessment of hepatic steatosis: prospective comparison of the accuracy of imaging examinations. J Hepatol 2010;52:579-85.

14. Mukherjee R, Tewary S, Routray A. Diagnostic and Prognostic Utility of Non-Invasive Multimodal Imaging in Chronic Wound Monitoring: a Systematic Review. J Med Syst 2017;41:46.

15. Yu QJ, Luo YC, Zuo ZW, Xie C, Yang TY, Wang T, Cheng L. Utility of gadoxetate disodium-enhanced magnetic resonance imaging in evaluating liver failure risk after major hepatic resection. Quant Imaging Med Surg 2024;14:3731-43.

16. Poetter-Lang S, Bastati N, Messner A, Kristic A, Herold A, Hodge JC, Ba-Ssalamah A. Quantification of liver function using gadoxetic acid-enhanced MRI. Abdom Radiol (NY) 2020;45:3532-44.

17. Yamada A, Hara T, Li F, Fujinaga Y, Ueda K, Kadoya M,

Doi K. Quantitative evaluation of liver function with use of gadoxetate disodium-enhanced MR imaging. Radiology 2011;260:727-33.

18. Zheng W, Guo W, Xiong M, Chen X, Gao L, Song Y, Cao D. Clinic-radiological features and radiomics signatures based on Gd-BOPTA-enhanced MRI for predicting advanced liver fibrosis. Eur Radiol 2023;33:633-44.

19. Sandrasegaran K, Cui E, Elkady R, Gasparis P, Borthakur G, Tann M, Liangpunsakul S. Can functional parameters from hepatobiliary phase of gadoxetate MRI predict clinical outcomes in patients with cirrhosis? Eur Radiol 2018;28:4215-24.

20. Bastati N, Beer L, Mandorfer M, Poetter-Lang S, Tamandl D, Bican Y, Elmer MC, Einspieler H, Semmler G, Simbrunner B, Weber M, Hodge JC, Vernuccio F, Sirlin C, Reiberger T, Ba-Ssalamah A. Does the Functional Liver Imaging Score Derived from Gadoxetic Acid-enhanced MRI Predict Outcomes in Chronic Liver Disease? Radiology 2020;294:98-107.

21. Kurata C, Saito K, Shirota N, Araki Y, Sugimoto K, Tajima Y, Yunaiyama D. The feasibility of superparamagnetic iron oxide-enhanced magnetic resonance imaging for assessing liver lesions in patients with contraindications for iodine CT contrast media or gadolinium-based MR contrast media: a retrospective case-control study. Quant Imaging Med Surg 2022;12:4612-21.

22. Chen Y, Liu Z, Mo Y, Li B, Zhou Q, Peng S, Li S, Kuang M. Prediction of Post-hepatectomy Liver Failure in Patients With Hepatocellular Carcinoma Based on Radiomics Using Gd-EOB-DTPA-Enhanced MRI: The Liver Failure Model. Front Oncol 2021;11:605296.

23. Zhao X, Huang M, Zhu Q, Wang T, Liu Q. The relationship between liver function and liver parenchymal contrast enhancement on Gd-BOPTA-enhanced MR imaging in the hepatocyte phase. Magn Reson Imaging 2015;33:768-73.

24. Zhou IY, Catalano OA, Caravan P. Advances in functional and molecular MRI technologies in chronic liver diseases. J Hepatol 2020;73:1241-54.

25. Asayama Y, Nishie A, Ishigami K, Ushijima Y, Takayama Y, Okamoto D, Fujita N, Yoshizumi T, Hida T, Honda H. Histogram analysis of noncancerous liver parenchyma on gadoxetic acid-enhanced MRI: predictive value for liver function and pathology. Abdom Radiol (NY) 2016;41:1751-7.

26. Just N. Improving tumour heterogeneity MRI assessment with histograms. Br J Cancer 2014;111:2205-13.

27. de Graaf W, van Lienden KP, van Gulik TM, Bennink

RJ. (99m)Tc-mebrofenin hepatobiliary scintigraphy with SPECT for the assessment of hepatic function and liver functional volume before partial hepatectomy. J Nucl Med 2010;51:229-36.

28. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, Gerig G. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 2006;31:1116-28.

29. Motosugi U, Ichikawa T, Sou H, Sano K, Tominaga L, Kitamura T, Araki T. Liver parenchymal enhancement of hepatocyte-phase images in Gd-EOB-DTPA-enhanced MR imaging: which biological markers of the liver function affect the enhancement? J Magn Reson Imaging 2009;30:1042-6.

30. Feier D, Balassy C, Bastati N, Stift J, Badea R, Ba-Ssalamah A. Liver fibrosis: histopathologic and biochemical influences on diagnostic efficacy of hepatobiliary contrast-enhanced MR imaging in staging. Radiology 2013;269:460-8.

31. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin JC, Pieper S, Aerts HJWL. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Res 2017;77:e104-7.

32. Boca Petresc B, Caraiani C, Popa L, Lebovici A, Feier DS, Bodale C, Buruian MM. The Utility of ADC First-Order Histogram Features for the Prediction of Metachronous Metastases in Rectal Cancer: A Preliminary Study. Biology (Basel) 2022;11:452.

33. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, Sanduleanu S, Larue RTHM, Even AJG, Jochems A, van Wijk Y, Woodruff H, van Soest J, Lustberg T, Roelofs E, van Elmpt W, Dekker A, Mottaghy FM, Wildberger JE, Walsh S. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749-62.

34. Hanna RF, Aguirre DA, Kased N, Emery SC, Peterson MR, Sirlin CB. Cirrhosis-associated hepatocellular nodules: correlation of histopathologic and MR imaging features. Radiographics 2008;28:747-69.

35. Kim H, Park SH, Kim EK, Kim MJ, Park YN, Park HJ, Choi JY. Histogram analysis of gadoxetic acid-enhanced MRI for quantitative hepatic fibrosis measurement. PLoS One 2014;9:e114224.

36. Xu X, Zhu H, Li R, Lin H, Grimm R, Fu C, Yan F. Whole-liver histogram and texture analysis on T1 maps improves the risk stratification of advanced fibrosis in NAFLD. Eur Radiol 2021;31:1748-59.

37. Park HJ, Lee SS, Park B, Yun J, Sung YS, Shim WH, Shin YM, Kim SY, Lee SJ, Lee MG. Radiomics Analysis of Gadoxetic Acid-enhanced MRI for Staging Liver Fibrosis. Radiology 2019;290:380-7.

38. Reeder SB, Yokoo T, França M, Hernando D, Alberich-Bayarri Á, Alústiza JM, Gandon Y, Henninger B, Hillenbrand C, Jhaveri K, Karçaaltıncaba M, Kühn JP, Mojtahed A, Serai SD, Ward R, Wood JC, Yamamura J, Martí-Bonmatí L. Quantification of Liver Iron Overload with MRI: Review and Guidelines from the ESGAR and SAR. Radiology 2023;307:e221856.

39. Martí-Aguado D, Jiménez-Pastor A, Alberich-Bayarri Á, Rodríguez-Ortega A, Alfaro-Cervello C, Mestre-Alagarda C, Bauza M, Gallén-Peris A, Valero-Pérez E, Ballester MP, Gimeno-Torres M, Pérez-Girbés A, Benlloch S, Pérez-Rojas J, Puglia V, Ferrández A, Aguilera V, Escudero-García D, Serra MA, Martí-Bonmatí L. Automated Whole-Liver MRI Segmentation to Assess Steatosis and Iron Quantification in Chronic Liver Disease. Radiology 2022;302:345-54.