Article

Discovery Medicine 2026; 38(211): 2514–2523
https://doi.org/10.24976/Discov.Med.202638211.198

A Simple Conventional Ultrasound-Based Risk Score for Predicting Progression to Cirrhosis in Patients With Compensated Early-Stage Liver Fibrosis

Yu Chen[1], Yu Cao[2], Jiqing Xuan[1,*]

[1]Department of Ultrasound Medicine, The Affiliated Hospital of Southwest Medical University, 646099 Luzhou, Sichuan, China
[2]Department of Rehabilitation, Gulin County People's Hospital, 646500 Luzhou, Sichuan, China
*Correspondence: xuanjiqing2026@126.com (Jiqing Xuan)
Submitted: 9 April 2026 Revised: 24 June 2026 Accepted: 30 June 2026 Published: 24 August 2026

Background: Patients with compensated early-stage liver fibrosis (F1–F2) are at risk of progressing to cirrhosis, yet routine surveillance with liver biopsy or transient elastography is limited by invasiveness, cost, or accessibility. Simple, non-invasive tools are therefore needed for risk stratification. This study aims to develop a simple risk scoring system based on conventional ultrasound features and evaluate its ability to predict progression to cirrhosis within 3 years in patients with compensated early-stage liver fibrosis (F1–F2).

Methods: This retrospective study included 502 patients with biopsy-proven compensated early-stage liver fibrosis. Four ultrasound features—liver surface, liver parenchymal echogenicity, liver margin morphology, and splenic volume (calculated as length × thickness × width × 0.52)—were semiquantitatively scored (0–2 points each) at baseline. The total score (0–8) was used to classify patients into low-risk (≤4) and high-risk (>4) groups. The primary endpoint was progression to compensated cirrhosis within 3 years. A Cox proportional hazards model was used to assess the association between the risk score and cirrhosis progression. Model discrimination was evaluated using the C-index and time-dependent ROC curves. Progression-free survival curves were plotted using the Kaplan–Meier method and compared by the log-rank test.

Results: Among 502 patients, 374 (74.5%) were classified as low-risk (Risk Score ≤4) and 128 (25.5%) as high-risk (Risk Score >4). The 3-year cumulative cirrhosis progression rates were 8.6% and 35.2%, respectively (log-rank $p < 0.0001$). The risk score was an independent predictor (hazard ratio (HR) = 1.43, 95% confidence interval (CI) 1.29–1.59, $p < 0.001$), with a C-index of 0.716 and a 36-month time-dependent area under the receiver operating characteristic curve (AUC) of 0.734. Bootstrap internal validation yielded an optimism-corrected C-index of 0.716. The model showed good calibration (Brier score = 0.118). Decision curve analysis demonstrated positive net benefit of the score across threshold probabilities of 0.05–0.35. The score showed higher C-index (0.716 vs. 0.650, 0.631, and 0.660) and 36-month AUC (0.734 vs. 0.679, 0.672, and 0.675) compared to FIB-4, APRI, and LSM. Adding the score to liver stiffness measurement (LSM) significantly improved reclassification (net reclassification improvement (NRI) = 0.184, $p = 0.007$; integrated discrimination improvement (IDI) = 0.048, $p = 0.025$). Albumin was also an independent protective factor (HR = 0.93, 95% CI 0.87–0.99, $p = 0.024$) in the multivariable sensitivity model.

Conclusion: This simple risk score (0–8) based on conventional ultrasound features (liver surface, parenchymal echogenicity, liver margin, and splenic volume) is an independent predictor of progression to cirrhosis within 3 years in patients with compensated early-stage liver fibrosis. This simple scoring system may provide a convenient and cost-effective non-invasive tool for risk stratification in clinical practice; however, these findings require validation in prospective, multicenter cohorts.

Keywords: liver fibrosis; cirrhosis; conventional ultrasound; risk score; prediction model; non-invasive diagnosis

Introduction

Chronic liver disease is a major global public health problem with a continuously increasing burden. According to the Global Burden of Disease Study 2021, liver cancer remains one of the leading causes of cancer-related mortality worldwide, with substantial geographical disparities in its prevalence and mortality [1]. Additionally, the global prevalence of non-alcoholic fatty liver disease (NAFLD) has reached approximately 30% and continues to rise, representing a growing health threat across all regions [2]. According to the latest World Health Organization report, viral hepatitis caused approximately 1.3 million deaths in 2022, ranking as the second leading infectious cause of death globally [3]. Liver fibrosis is the common pathological process of chronic liver disease progression. In patients with compensated early-stage fibrosis (F1–F2), timely intervention may allow some degree of regression; however, without effective management, a substantial proportion of patients progress to cirrhosis over 5 to 10 years [4]. Once cirrhosis enters the decompensated stage, the 5-year survival

rate decreases significantly, and the risk of hepatocellular carcinoma increases markedly [5,6]. Therefore, early identification of patients with compensated liver fibrosis who are at high risk of progression is important for optimizing follow-up strategies, enabling timely intervention, and improving prognosis.

Although liver biopsy is the gold standard for assessing the stage of liver fibrosis, it has limitations including invasiveness, sampling error, poor repeatability, and low patient acceptance, making it difficult to use for routine follow-up [7,8]. In recent years, non-invasive assessment techniques for liver fibrosis have developed rapidly. Among them, liver stiffness measurement (LSM) by transient elastography has good diagnostic accuracy and is widely used for fibrosis staging and prognostic assessment [9,10]. However, LSM equipment is expensive, requires operator expertise, and is not widely available in primary healthcare institutions in China; moreover, its success rate can be affected by obesity, narrow intercostal spaces, and other factors [11]. Thus, exploring simple, cost-effective, and easily disseminable non-invasive methods remains clinically important.

Conventional ultrasound is a routine imaging examination for patients with liver disease, offering advantages such as non-invasiveness, convenience, low cost, and high accessibility. Previous studies have shown that ultrasound signs, including nodular liver surface, heterogeneous parenchymal echogenicity, blunt liver margin, and splenomegaly, are closely associated with the severity of liver fibrosis and portal hypertension [12,13]. Berzigotti [14] demonstrated that liver surface morphology is significantly correlated with fibrosis severity, with a sensitivity exceeding 85% for predicting cirrhosis (F4). The splenic volume, a composite measure of spleen size, is positively associated with the grade of esophageal varices and portal pressure and has been validated as an effective parameter for predicting liver disease progression [15,16,17]. Nevertheless, most previous studies have focused on the diagnostic value of individual ultrasound signs, and few have integrated these signs into a quantitative scoring system to predict disease progression.

Given this background, the present study aimed to develop a simple risk scoring system using four easily obtainable conventional ultrasound features (liver surface, parenchymal echogenicity, liver margin, and splenic volume) and to evaluate its value in predicting progression to cirrhosis within 3 years, specifically in patients with biopsy-proven compensated early-stage liver fibrosis (F1–F2). Unlike previous cross-sectional diagnostic studies that mainly focused on cirrhosis detection, our study provides longitudinal prognostic prediction using a simplified baseline ultrasound score in this specific population. This approach extends the application of conventional ultrasound from a diagnostic to a prognostic tool, with potential advantages in resource-limited settings.

Methods

Study Design and Population

This was a single-center retrospective cohort study. Patients with chronic liver disease who attended the Hepatology Center of The Affiliated Hospital of Southwest Medical University between February 2020 and February 2023 were enrolled. Inclusion criteria were as follows:

1. Biopsy-proven compensated early-stage liver fibrosis, with pathological stage F1 or F2 (according to the METAVIR staging system);
2. Age ≥18 years;
3. Baseline liver function Child-Pugh class A, with no decompensating events such as ascites, esophagogastric variceal bleeding, or hepatic encephalopathy;
4. Complete baseline conventional ultrasound images and clinical data;
5. Follow-up duration of at least 3 years, with ultrasound and liver function tests repeated every 6–12 months.

Exclusion criteria were as follows:

1. Concomitant liver disease with overlapping etiologies (e.g., HBV combined with autoimmune hepatitis, HCV combined with drug-induced liver injury);
2. Concurrent hepatocellular carcinoma or other malignancies;
3. Radiological or pathological evidence of cirrhosis at baseline;
4. Initiation of new antiviral or anti-fibrotic therapy, or a change in the existing treatment regimen during the follow-up period (patients on stable antiviral therapy prior to enrollment were not excluded);
5. Incomplete clinical or imaging data.

Patients with a single etiology other than HBV, HCV, or NAFLD (including cryptogenic liver disease, inherited metabolic liver disease, vascular liver disease, and other rare etiologies) were categorized as “Other” and not excluded, provided that comprehensive evaluations ruled out the major etiologies.

The study was approved by the Ethics Committee of The Affiliated Hospital of Southwest Medical University (approval number: KY2026121) and complied with the ethical principles of the Declaration of Helsinki. Written informed consent was waived by the Ethics Committee because the study posed minimal risk, used anonymized retrospective data obtained from routine clinical practice.

Baseline Conventional Ultrasound Examination and Risk Score Construction

All patients underwent abdominal conventional ultrasound within one week after liver biopsy. Examinations were performed independently by two sonographers with more than 10 years of experience, who were blinded to the pathological results. Ultrasound systems used were Philips iU22 or GE Logiq E9 with convex array probes at 3.5–5.0 MHz. To assess inter-observer agreement, a randomly se-

Table 1. Semiquantitative scoring criteria for conventional ultrasound features.

Ultrasound feature	0 points	1 point	2 points
Liver surface	Smooth	Irregular/serrated	Nodular/undulating
Liver parenchyma	Homogeneous	Mildly heterogeneous/slightly enhanced	Markedly heterogeneous/granular
Liver margin	Sharp	Slightly blunt	Blunt/irregular
Splenic volume (cm^3)	<300	300–480	>480

lected subset of 20 patients was independently reviewed by both sonographers. Weighted kappa statistics were calculated for the categorical ultrasound features (liver surface, parenchymal echogenicity, liver margin), and the intraclass correlation coefficient (ICC) was calculated for the continuous variable (splenic volume). The weighted kappa values were 0.82 (liver surface), 0.79 (parenchymal echogenicity), and 0.84 (liver margin). The ICC for splenic volume was 0.91, indicating good to excellent agreement.

Based on previous literature and clinical practice, four ultrasound features closely associated with liver fibrosis and portal hypertension [18,19] were collected and semiquantitatively scored (0–2 points) according to severity (Table 1).

Splenic volume was calculated using the ellipsoid formula: volume (cm^3) = length × thickness × width × 0.52. Length is the maximum longitudinal diameter, thickness is the maximum thickness perpendicular to the length at the hilum, and width is the maximum transverse diameter on a cross-sectional image. Watanabe et al. [20] reported that splenic volume was significantly correlated with the severity of esophageal varices and liver functional reserve in 110 cirrhotic patients; patients with esophageal varices had significantly larger splenic volumes than those without ($p < 0.0001$), and those with red signs (high-risk stigmata) had even larger volumes ($p = 0.0029$). Fateh et al. [21] also reported a mean splenic volume of 174.4 ± 52.4 mL (range approximately 122–227 mL) in 300 healthy adults. Based on these literature findings, we categorized splenic volume into three grades: <300 cm^3 (0 points, normal), 300–480 cm^3 (1 point, moderate enlargement), and >480 cm^3 (2 points, marked enlargement), where 480 cm^3 was the optimal cutoff for predicting esophageal varices as reported by Katwal et al. [22].

The simple risk score was the sum of the above four items, ranging from 0 to 8 points. The optimal cut-off value for the risk score was determined using the surv_cutpoint function from the survminer package in R software (version 4.5.2; R Foundation for Statistical Computing, Vienna, Austria). This function identifies the value that maximizes the difference in survival between the two groups based on the log-rank test. The analysis was performed with a minimum proportion of 0.1 in each group, and the resulting optimal cut-off was a score of 4 (low-risk: ≤4, high-risk: >4).

Data Collection

All patients were followed up every 6–12 months by outpatient visits or telephone interviews. The follow-up start point was the date of the baseline ultrasound, and the endpoint was the time of first progression to compensated cirrhosis or completion of 3 years of follow-up. Progression to compensated cirrhosis was defined based on in-person assessments during outpatient visits (imaging, elastography, or biopsy). Telephone interviews were used only to ascertain survival status or the occurrence of major decompensating events (e.g., ascites, variceal bleeding). To ensure independence between baseline scoring and endpoint adjudication, baseline ultrasound scoring and endpoint determination were performed by different radiologists who were mutually blinded to each other's assessments. At baseline, two radiologists independently reviewed all ultrasound images to confirm the absence of definite cirrhosis signs; if baseline imaging showed suspicious findings, the biopsy result (F1–F2) overruled the imaging suspicion. For endpoint diagnosis, at least two imaging criteria (nodular liver surface, widened hepatic fissures, or enlarged caudate lobe) were required, not solely liver surface nodularity, to minimize potential circularity.

Progression to compensated cirrhosis was defined as meeting at least one of the following criteria:

1. Imaging criteria: new definite signs of cirrhosis on follow-up abdominal ultrasound, CT, or MRI, including at least two of the following: nodular liver surface, widened hepatic fissures, or enlarged caudate lobe, and LSM by transient elastography ≥13.0 kPa within 3 months of the imaging; confirmed by two radiologists independently;

2. Elastography criteria: LSM by transient elastography ≥13.0 kPa and persistently elevated (two measurements at least 6 months apart);

3. Pathological criteria: biopsy-proven F4 stage during follow-up.

Patients who did not experience the endpoint event during 3 years of follow-up were censored at their last follow-up visit.

The following clinical data were collected from the electronic medical record system: Demographic characteristics: age, sex; Laboratory parameters: platelet count (PLT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), albumin, AST/ALT ratio; Fibrosis assessment: baseline liver biopsy stage (F1/F2), LSM by transient elastography; Comorbidities and lifestyle: diabetes mellitus, history of alcohol consumption (defined as

average daily alcohol intake >30 g for ≥5 years); Etiology: chronic hepatitis B, chronic hepatitis C, NAFLD, and others. The predictive performance of the ultrasound risk score was compared with established non-invasive tests, including FIB-4, APRI, and LSM, using the C-index, time-dependent area under the receiver operating characteristic curve (AUC), net reclassification improvement (NRI), and integrated discrimination improvement (IDI).

In this study, alcoholic liver disease (ALD) and alcohol consumption history were distinguished as follows: ALD was defined as average daily alcohol intake >30 g (males) or >20 g (females) for ≥5 years, with other etiologies excluded, combined with laboratory findings (AST/ALT ratio >1.5, markedly elevated GGT) and imaging features, confirmed by two hepatologists; patients meeting these criteria were excluded. Alcohol consumption history was defined as average daily alcohol intake >30 g for ≥5 years with a clear concurrent diagnosis of HBV, HCV, or NAFLD, where the primary cause of liver injury was determined to be viral hepatitis or NAFLD rather than alcohol; these patients were included, and alcohol history was analyzed as a covariate. All liver biopsy specimens were obtained using 16G Menghini-type needles. Specimens measuring ≥15 mm or containing ≥6 complete portal tracts were considered adequate. Histological staging was performed according to the METAVIR system by two independent pathologists (>10 and >15 years of experience, respectively) who were blinded to all clinical, laboratory, and imaging data. In cases of discordance, a final consensus was reached through discussion, or a third senior pathologist was consulted.

Statistical Analysis

Statistical analyses were performed using R software version 4.3.0. Normality of continuous variables was assessed using the Shapiro-Wilk test. Normally distributed variables are presented as mean ± SD and were compared using one-way ANOVA; non-normally distributed variables are presented as median (interquartile range) and were compared using the Kruskal-Wallis test. Categorical variables are presented as frequencies (percentages) and were compared using the chi-square test or Fisher's exact test. A two-sided $p < 0.05$ was considered statistically significant.

The association between the risk score and cirrhosis progression was assessed using the Cox proportional hazards model, with hazard ratios (HRs) and 95% confidence intervals (CIs) reported. Univariate Cox regression was first performed to screen for variables associated with progression. In the primary model, only the risk score was included. Collinearity among predictor variables was assessed using the variance inflation factor (VIF). In a multivariable sensitivity model, covariates with $p < 0.05$ in univariate analysis were entered together with the risk score to evaluate the independent predictive value of the score. Model discrimination was evaluated using the C-index and the 36-month time-dependent AUC. A C-index >0.6 indicated acceptable discriminative ability. The time-dependent ROC curve was calculated using the Kaplan–Meier method to assess the predictive accuracy of the risk score at a fixed time point (36 months). Based on the risk score, patients were divided into low- and high-risk groups, and cumulative progression-free survival curves were plotted using the Kaplan–Meier method and compared by the log-rank test. Internal validation was performed using the bootstrap method with 1000 resamples to calculate the optimism-corrected C-index. Calibration of the 3-year predicted risk was assessed using a calibration plot and the Brier score. Decision curve analysis (DCA) was performed to evaluate the clinical utility of the risk score across a range of threshold probabilities. All tests were two-sided, and $p < 0.05$ was considered statistically significant.

Results

Study Population and Baseline Characteristics

A total of 502 patients with compensated early-stage liver fibrosis were included and classified into low-risk (n = 374) and high-risk (n = 128) groups according to the risk score. There were significant differences in albumin, platelet count, AST/ALT ratio, LSM, prevalence of diabetes mellitus, history of alcohol consumption, and risk score between the two groups (all $p < 0.05$). The distribution of etiologies also differed significantly ($p < 0.001$), with a higher proportion of HCV and NAFLD in the high-risk group. Age, sex and baseline fibrosis stage did not differ significantly between the groups. Detailed baseline characteristics are shown in Table 2.

Association Between Risk Score and Disease Progression

Univariate analysis showed that albumin, platelet count, AST/ALT ratio, LSM, diabetes, alcohol history, and the risk score were significantly associated with progression (all $p < 0.05$, Table 3).

In the primary model that included only the risk score, each 1-point increase in the score was associated with a 43% increased risk of progression (HR = 1.43, 95% CI 1.29–1.59, $p < 0.001$), with a C-index of 0.716. In the multivariable sensitivity model adjusting for other univariately significant variables, the risk score remained independently predictive (HR = 1.35, 95% CI 1.15–1.59, $p < 0.001$), and albumin was also an independent protective factor (HR = 0.93, 95% CI 0.87–0.99, $p = 0.024$) (Table 4). Variance inflation factor (VIF) values for all variables in the multivariable sensitivity model were below 5, indicating the absence of significant multicollinearity.

In the sensitivity model further adjusted for etiology, the ultrasound risk score remained a robust independent predictor (HR = 1.33, 95% CI 1.13–1.57, $p < 0.001$), com-

Table 2. Baseline characteristics of patients in different risk groups.

Characteristic	High-risk group (n = 128)	Low-risk group (n = 374)	Statistic	*p* value
Age (years), mean ± SD	47.47 ± 10.13	45.51 ± 9.67	$t = 1.96$	0.051
Sex, n (%)			$\chi^2 = 0.50$	0.481
Female	59 (46.09)	159 (42.51)		
Male	69 (53.91)	215 (57.49)		
Baseline fibrosis, n (%)			$\chi^2 = 3.04$	0.081
F1	79 (61.72)	262 (70.05)		
F2	49 (38.28)	112 (29.95)		
Albumin (g/L), mean ± SD	40.54 ± 3.99	43.61 ± 3.27	$t = –7.87$	<0.001
Etiology, n (%)			$\chi^2 = 22.60$	<0.001
HBV	53 (41.41)	213 (56.95)		
HCV	31 (24.22)	35 (9.36)		
NAFLD	33 (25.78)	79 (21.12)		
Other	11 (8.59)	47 (12.57)		
Platelet count ($\times 10^9$/L), mean ± SD	154.27 ± 42.99	203.43 ± 47.67	$t = –10.32$	<0.001
AST/ALT ratio, M (Q_1, Q_3)	1.12 (0.86, 1.36)	0.85 (0.69, 1.03)	$Z = –6.71$	<0.001
LSM (kPa), M (Q_1, Q_3)	12.20 (9.97, 14.62)	6.90 (5.80, 8.30)	$Z = –13.70$	<0.001
Diabetes mellitus, n (%)			$\chi^2 = 32.08$	<0.001
No	87 (67.97)	334 (89.30)		
Yes	41 (32.03)	40 (10.70)		
Alcohol history, n (%)			$\chi^2 = 12.99$	<0.001
No	86 (67.19)	308 (82.35)		
Yes	42 (32.81)	66 (17.65)		
Risk score, M (Q_1, Q_3)	6.00 (5.00, 7.00)	2.00 (1.00, 3.00)	$Z = –17.10$	<0.001

AST, aspartate aminotransferase; ALT, alanine aminotransferase; LSM, liver stiffness measurement.

Table 3. Univariate Cox regression analysis.

Variable	HR	95% CI	*p* value
Age	1.02	0.99–1.04	0.149
Sex	0.92	0.59–1.45	0.731
Baseline fibrosis	1.10	0.68–1.76	0.700
Albumin	0.86	0.81–0.91	<0.001
Platelet count	0.99	0.99–0.99	<0.001
AST/ALT ratio	3.30	1.82–5.99	<0.001
LSM	1.15	1.09–1.22	<0.001
Diabetes	1.94	1.16–3.22	0.011
Alcohol history	1.70	1.05–2.76	0.030
Etiology (Reference: HBV)			
HCV	1.61	0.87–2.98	0.129
NAFLD	1.37	0.80–2.35	0.247
Other	0.59	0.23–1.51	0.276
Risk score	1.43	1.29–1.59	<0.001

parable to the primary multivariable model (HR = 1.35). Albumin also retained its independent protective effect (HR = 0.93, 95% CI 0.87–0.99, $p = 0.025$). None of the other covariates, including platelet count, AST/ALT ratio, LSM, diabetes, or alcohol history, reached statistical significance (all $p > 0.05$). The overall effect of etiology was not significant ($p = 0.498$), with no individual etiology category showing a significant association (HCV: HR = 1.20, $p = 0.579$; NAFLD: HR = 1.13, $p = 0.670$; Other: HR = 0.55, $p = 0.218$). These findings suggest that the predictive value of the ultrasound risk score is independent of etiological differences.

Risk Stratification and Cumulative Progression Rates

Based on the risk score, the two groups showed a significant difference in the 3-year cumulative cirrhosis progression rate: 8.6% in the low-risk group, and 35.2% in the high-risk group (Table 5). The Kaplan–Meier curve showed that progression-free survival was significantly lower in the high-risk group than in the low-risk group (log-rank $p < 0.0001$, Fig. 1).

Discriminative Ability of Models

The Cox regression model containing only the risk score had a C-index of 0.716, and the 36-month time-dependent ROC AUC was 0.734 (Fig. 2A), indicating good discriminative ability for predicting progression to cirrhosis in patients with compensated early-stage liver fibrosis. Bootstrap internal validation yielded an optimism-corrected C-index of 0.716. Calibration was assessed using a calibration plot (Fig. 2B), calibration slope (1.019, optimism-corrected via bootstrap), Brier score (0.118), and the Hosmer-Lemeshow test ($p = 0.322$), all indicating good agreement between predicted and observed 36-month risks. Decision curve analysis demonstrated positive net benefit of the score across threshold probabilities of 0.05–0.35 (Fig.

Table 4. Multivariable Cox regression analysis.

Model	Variable	HR	95% CI	*p* value
Primary model	Risk score	1.43	1.29–1.59	<0.001
Multivariable sensitivity model	Risk score	1.35	1.15–1.59	<0.001
	Albumin	0.93	0.87–0.99	0.024
	Platelet count	1.00	0.99–1.00	0.290
	AST/ALT ratio	1.24	0.66–2.35	0.507
	LSM	0.94	0.86–1.03	0.195
	Diabetes	1.35	0.79–2.32	0.275
	Alcohol history	1.32	0.80–2.17	0.278
Sensitivity model (adjusted for etiology)	Risk score	1.33	1.13–1.57	<0.001
	Albumin	0.93	0.87–0.99	0.025
	Platelet count	1.00	0.99–1.00	0.345
	AST/ALT ratio	1.35	0.70–2.57	0.369
	LSM	0.95	0.86–1.04	0.234
	Diabetes	1.38	0.80–2.38	0.246
	Alcohol history	1.29	0.78–2.14	0.321
	Etiology (overall)			0.498
	HCV (vs. HBV)	1.20	0.63–2.27	0.579
	NAFLD (vs. HBV)	1.13	0.65–1.94	0.670
	Other (vs. HBV)	0.55	0.22–1.42	0.218

Table 5. 3-year cumulative cirrhosis progression rates by risk group.

Risk group	Number of patients	Number of events	3-year progression rate (%)
Low-risk	374	32	8.6
High-risk	128	45	35.2

2C). The ultrasound risk score showed numerically higher C-index (0.716 vs. 0.650, 0.631, and 0.660, respectively) and 36-month AUC values (0.734 vs. 0.679, 0.672, and 0.675) compared to FIB-4, APRI, and LSM in this cohort (Fig. 2D). Adding the score to LSM significantly improved risk reclassification (NRI = 0.184, $p = 0.007$; IDI = 0.048, $p = 0.025$).

Discussion

Based on a retrospective cohort of 502 patients with compensated early-stage liver fibrosis, we developed and validated a simple risk scoring system (0–8 points) that includes four conventional ultrasound features: liver surface, parenchymal echogenicity, liver margin, and splenic volume. The main findings are as follows: (1) The risk score was an independent predictor of progression to cirrhosis within 3 years (HR = 1.43, 95% CI 1.29–1.59, $p < 0.001$), with good model discrimination (C-index = 0.716, 36-month AUC = 0.734). (2) The score effectively stratified patients into low- and high-risk groups, with 3-year cumulative progression rates of 8.6% and 35.2%, respectively (log-rank $p < 0.001$). (3) In the sensitivity analysis, albumin also showed an independent protective effect (HR = 0.93, $p = 0.024$), suggesting that liver synthetic function is independently associated with progression risk. Compared to FIB-4, APRI, and LSM, our simple ultrasound score demonstrated numerically higher discriminative ability in this cohort, although these differences require confirmation in larger external validation studies. Its main advantages lie in its reliance on widely available conventional ultrasound, its low cost, and the immediate availability of results, making it particularly suitable for primary care and resource-limited settings where LSM or complex blood test-derived indices are not feasible.

Most previous studies have focused on the cross-sectional diagnostic value of single ultrasound signs for staging liver fibrosis, with relatively few longitudinal studies predicting disease progression. Colli et al. [23] reported that the sensitivity of liver surface nodularity for predicting cirrhosis was approximately 88%, but they did not provide long-term follow-up data. The present study is the first to integrate four conventional ultrasound features into a semi-quantitative score and validate its ability to predict progression to cirrhosis over 3 years, filling a gap in translating imaging findings into prognostic assessment.

The splenic volume, as a composite measure of spleen size, has been shown to be positively correlated with the hepatic venous pressure gradient and the severity of esophageal varices [16]. By incorporating splenic volume into the scoring system, our model can sensitively capture early changes in portal hypertension. Notably, platelet count, LSM, and other indicators were significantly associ-

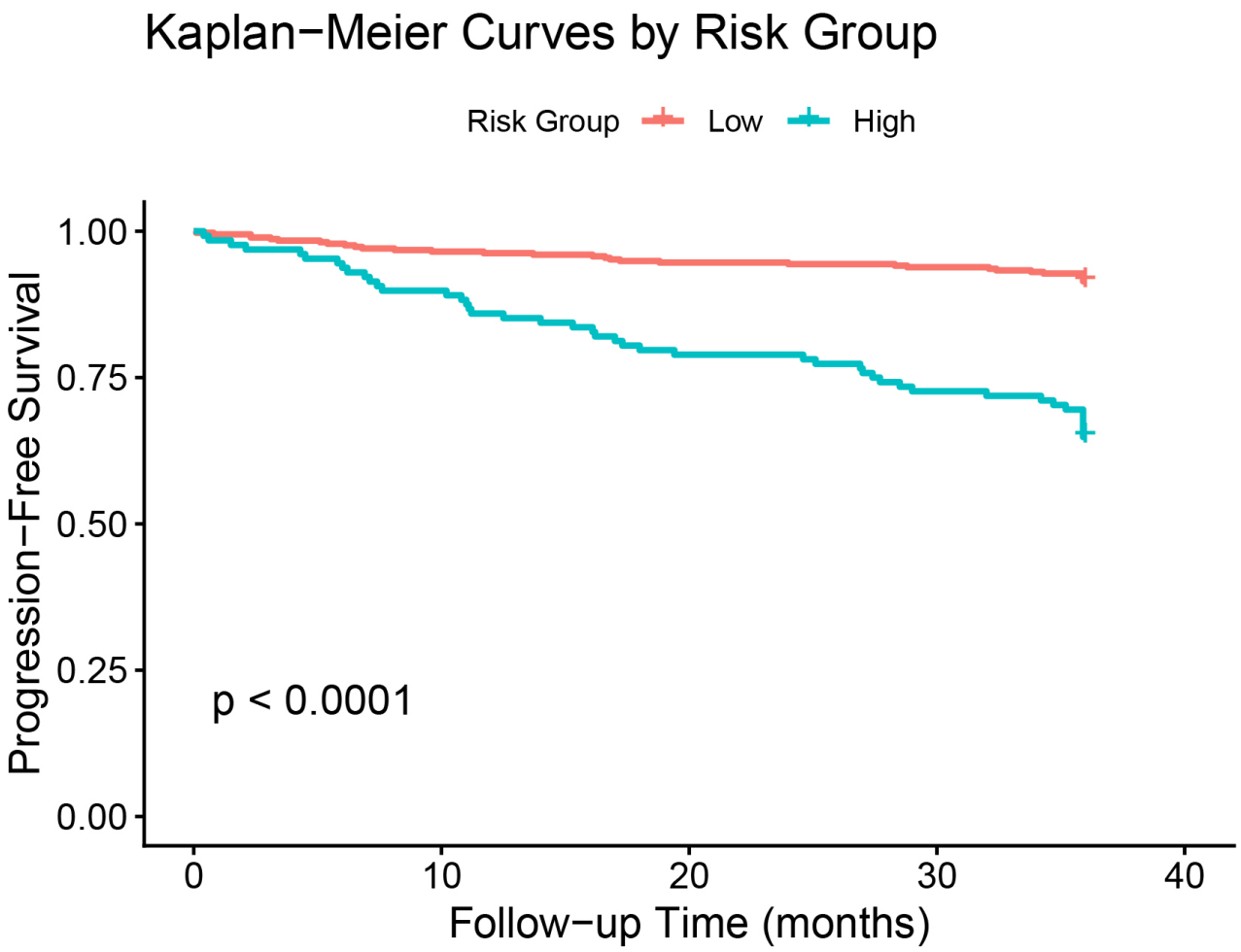

Fig. 1. Progression-free survival curves for patients in different risk groups.

ated with progression in univariate analysis but lost significance in the multivariate multivariable sensitivity model. This finding suggests that the information on fibrosis/portal hypertension reflected by thrombocytopenia and elevated LSM is largely encompassed by the ultrasound risk score, further supporting the intrinsic value of the ultrasound score as a comprehensive non-invasive indicator.

Albumin retained independent predictive value in the multivariable sensitivity model (HR = 0.93, $p = 0.024$), consistent with previous studies. Albumin is synthesized by hepatocytes, and a decline in its level directly reflects impaired liver functional reserve; it has been established as an independent predictor of cirrhosis progression and mortality [24]. However, albumin measurement requires blood sampling, whereas the ultrasound score is completely non-invasive and available in real time. Combining the two may further improve predictive performance, a hypothesis worthy of future investigation.

Etiological heterogeneity was observed in our cohort, with significant differences in the distribution of etiologies between the low- and high-risk groups. Given that fibrosis progression rates may vary by etiology, we performed a sensitivity analysis adjusting for etiology to exclude potential confounding. Notably, after adjusting for etiology, the ultrasound risk score maintained its independent predictive value (HR = 1.334, 95% CI 1.134–1.570, $p < 0.001$), while the overall effect of etiology itself was not significant ($p = 0.498$). This suggests that the prognostic information captured by the ultrasound score—reflecting fibrotic changes and portal hypertension—is largely consistent across etiologies. Nevertheless, due to the relatively limited sample sizes within each etiological subgroup (particularly for HCV and NAFLD in the high-risk group), these findings should be interpreted as hypothesis-generating. Future prospective studies with larger, homogeneous etiological cohorts (e.g., HBV-only or NAFLD-only) are warranted to validate the score's generalizability and to determine whether etiology-specific cutoff values are needed.

The risk scoring system has several clinical advantages, particularly for resource-limited settings where blood tests or transient elastography are not readily available: (1) Ultrasound-only, no blood draw: Unlike FIB-4, APRI, or serum albumin, which require venipuncture and laboratory infrastructure, our score uses only ultrasound features. This

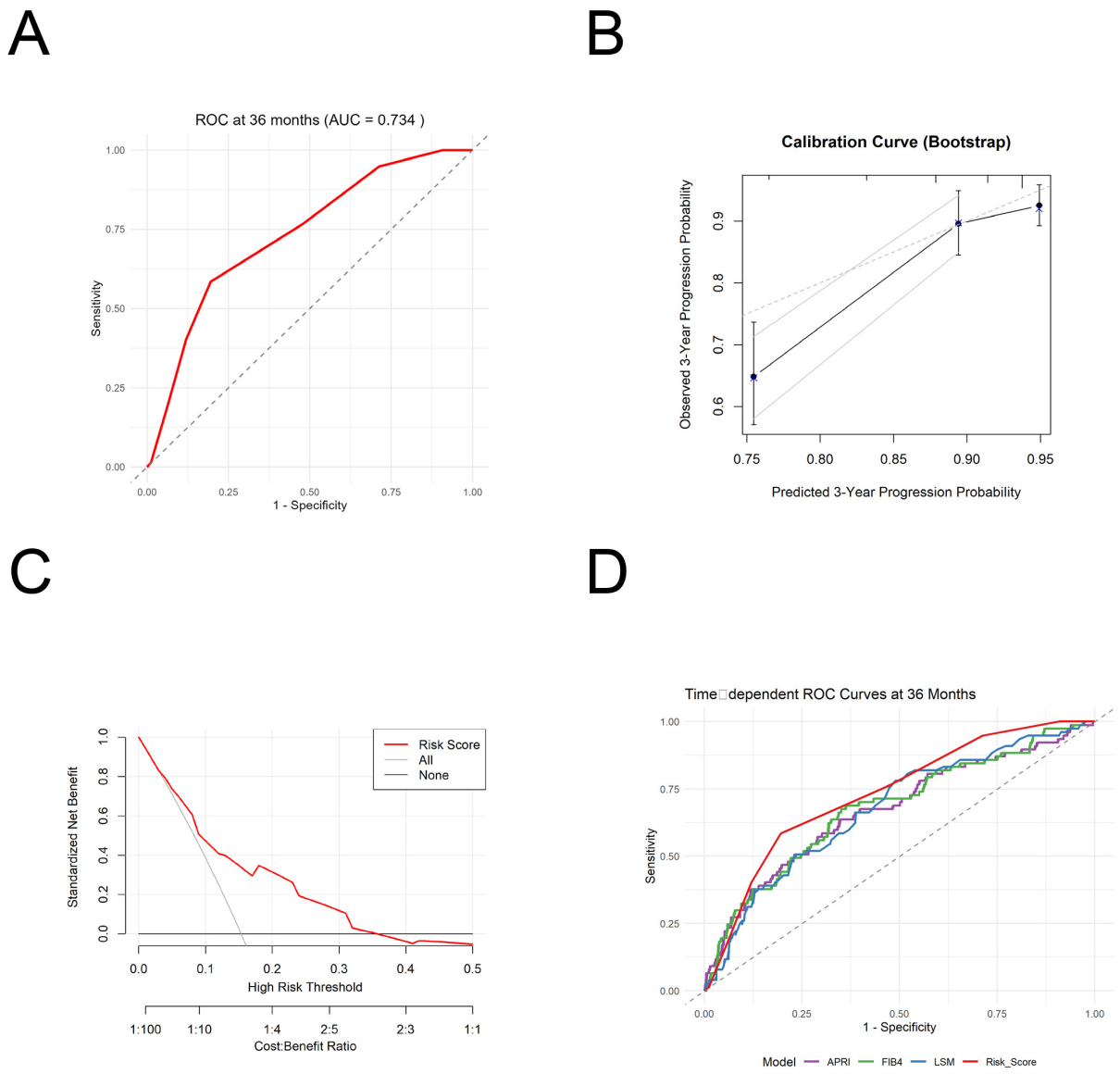

Fig. 2. 36-month predictive performance of the ultrasound risk score model. (A) 36-month time-dependent ROC curve. (B) Bootstrap calibration curve. (C) Decision curve analysis. (D) Performance comparison with FIB-4, APRI and LSM.

makes it suitable for primary care clinics or field settings where phlebotomy or lab equipment is unavailable. (2) Low cost and wide availability: Conventional ultrasound is substantially less expensive than transient elastography (approximately one-third to one-fifth the cost) and is available in most primary healthcare institutions across China, even in rural areas. (3) Immediate result at point of care: The total score (0–8) can be calculated within minutes during the ultrasound examination, without waiting for blood test results or specialized software. (4) Clear stratification: The progression rates show a distinct gradient, which helps to formulate individualized follow-up plans. For example, low-risk patients may extend follow-up intervals to 2 years, whereas high-risk patients should undergo ultrasound and liver function tests every 6 months.

We emphasize that this tool is not intended to replace LSM or FIB-4 in well-equipped centers, but rather to fill a gap in settings where those modalities are not feasible. The present results should be interpreted as strictly hypothesis-generating given the single-center, retrospective design, and the absence of external validation. At this stage, the scoring system should not be used to guide clinical decision-making. External validation in independent, preferably prospective, cohorts is mandatory before this scoring system can be considered for routine clinical utility. We acknowledge that the equal-weight assignment and the variable selection based on clinical knowledge rather than data-driven methods (e.g., LASSO) deviate from standard prediction model development. This approach was an intentional choice to prioritize clinical practicality and gen-

eralizability, making the tool suitable for resource-limited settings. Therefore, this study should be considered an exploratory validation of a clinically pre-defined tool, not a de novo model development.

This study has several limitations. First, the single-center retrospective design may introduce selection bias, and the requirement for complete 3-year follow-up may have systematically excluded patients with more rapid disease progression or those lost to follow-up, potentially biasing the cohort toward more stable cases. Additionally, the optimal cut-off value (score >4) was determined and the resulting group comparison was performed within the same dataset, which may introduce optimism bias due to data-adaptive selection; therefore, these cut-off-defined results should be considered exploratory and require independent external validation. Second, despite the use of semi-quantitative criteria, ultrasound scoring remains subject to inter-observer variability, particularly in grading liver surface and parenchymal echogenicity; standardized training and pictorial guidelines could mitigate this in future studies. Third, there is a potential risk of incorporation bias, as both the baseline ultrasound score and components of the composite endpoint (imaging criteria) rely on ultrasound findings, which may lead to some degree of overestimation of the score's predictive performance. Consequently, our model predicts progression as defined by the composite clinical endpoint (imaging-elastography-pathology), not solely histological progression to F4. Fourth, although the cohort included three main etiologies (HBV, HCV, and NAFLD), other etiologies were underrepresented, and the sample size within each etiological subgroup was limited. Nevertheless, a sensitivity analysis adjusting for etiology confirmed the score's robustness (HR = 1.334, $p < 0.001$); however, validation in larger homogeneous etiological cohorts remains warranted. Fifth, the inclusion of patients on stable pre-existing therapy alongside untreated patients creates a mixed cohort that may introduce treatment-related confounding, and the score's performance in patients initiating new therapy requires separate validation. Sixth, the composite endpoint (imaging, elastography, or histology) may introduce heterogeneity despite stringent criteria for each modality, and the 3-year fixed follow-up limits assessment of longer-term predictive value. Seventh, the splenic volume thresholds were derived from cirrhotic populations and require validation specifically in F1–F2 patients. Eighth, the relatively high baseline LSM values in the high-risk group (median 12.20 kPa), despite biopsy-proven F1–F2, highlight the potential discordance between histology and elastography; factors such as sampling error, active inflammation, or steatosis may contribute to this discrepancy. This suggests that some patients may have already had subclinical structural changes at baseline, making them more likely to reach the composite endpoint, which could partly amplify the observed predictive performance. Finally, although our ultrasound features (e.g., nodular surface, splenomegaly) are more typical of advanced fibrosis, all patients had biopsy-proven F1–F2; this apparent discordance underscores that ultrasound signs reflect macroscale structural and hemodynamic changes that may not fully align with histological staging, and further studies are needed to clarify this relationship.

Conclusion

In conclusion, this retrospective single-center study demonstrates that a simple ultrasound-based risk score is associated with 3-year cirrhosis progression in patients with compensated early-stage fibrosis. However, due to the exploratory nature of the analysis and the lack of external validation, these findings are hypothesis-generating only. The score's potential as a screening tool requires confirmation in prospective, multicenter studies before any clinical application can be recommended.

Availability of Data and Materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author Contributions

YCh and JQX designed the research study. YCh performed the research. YCa analyzed the data. YCh and JQX drafted the article. All authors contributed to important editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

The study was approved by the Ethics Committee of The Affiliated Hospital of Southwest Medical University (approval number: KY2026121), and all procedures followed the principles of the Declaration of Helsinki. Written informed consent was waived by the Ethics Committee because the study posed minimal risk, used anonymized retrospective data obtained from routine clinical practice.

Acknowledgment

Not applicable.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Jiang Z, Zeng G, Dai H, Bian Y, Wang L, Cao W, et al. Global, regional and national burden of liver cancer 1990-2021: a systematic analysis of the global burden of disease study 2021. BMC Public Health. 2025; 25: 931. https://doi.org/10.1186/s12889-025-22026-6

[2] Younossi ZM, Golabi P, Paik JM, Henry A, Van Dongen C, Henry L. The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): a systematic review. Hepatology (Baltimore, Md.). 2023; 77: 1335–1347. https://doi.org/10.1097/HEP.0000000000000004

[3] World Health Organization. Global hepatitis report 2024: action for access in low-and middle-income countries. World Health Organization: Geneva. 2024.

[4] Rockey DC. Hepatic fibrosis and cirrhosis. Yamada's Textbook of Gastroenterology. 2022: 2000–23.

[5] D'Amico G, Morabito A, D'Amico M, Pasta L, Malizia G, Rebora P, et al. New concepts on the clinical course and stratification of compensated and decompensated cirrhosis. Hepatology International. 2018; 12: 34–43. https://doi.org/10.1007/s12072-017-9808-z

[6] Asrani SK, Kamath PS. Natural history of cirrhosis. Current Gastroenterology Reports. 2013; 15: 308. https://doi.org/10.1007/s11894-012-0308-y

[7] Bedossa P, Carrat F. Liver biopsy: the best, not the gold standard. Journal of Hepatology. 2009; 50: 1–3. https://doi.org/10.1016/j.jhep.2008.10.014

[8] Regev A, Berho M, Jeffers LJ, Milikowski C, Molina EG, Pyrsopoulos NT, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. The American Journal of Gastroenterology. 2002; 97: 2614–2618. https://doi.org/10.1111/j.1572-0241.2002.06038.x

[9] European Association for the Study of the Liver. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis - 2021 update. Journal of Hepatology. 2021; 75: 659–689. https://doi.org/10.1016/j.jhep.2021.05.025

[10] de Franchis R, Bosch J, Garcia-Tsao G, Reiberger T, Ripoll C, Baveno VII Faculty. Baveno VII - Renewing consensus in portal hypertension. Journal of Hepatology. 2022; 76: 959–974. https://doi.org/10.1016/j.jhep.2021.12.022

[11] van Doorn DJ, Holleboom AG, Takkenberg RB, Verheij J, Lantinga MA. Can liver stiffness measurement accurately predict disease progression and clinical outcome in patients with metabolic dysfunction-associated steatotic liver disease and bridging fibrosis or cirrhosis? Hepatobiliary Surgery and Nutrition. 2023; 12: 912–915. https://doi.org/10.21037/hbsn-23-445

[12] Ferraioli G, Barr RG, Berzigotti A, Sporea I, Wong VWS, Reiberger T, et al. WFUMB Guideline/Guidance on Liver Multiparametric Ultrasound: Part 1. Update to 2018 Guidelines on Liver Ultrasound Elastography. Ultrasound in Medicine & Biology. 2024; 50: 1071–1087. https://doi.org/10.1016/j.ultrasmedbio.2024.03.013

[13] Bauer DJM, SilvestriI AD, Mare R, Maiocchi L, Raimondi A, Semmler G, et al. Two-dimensional shear wave elastography (ElastQ) accurately rules out liver fibrosis and rules in advanced chronic liver disease across liver disease etiologies: a prospective multicenter study. Ultrasonography (Seoul, Korea). 2023; 42: 544–554. https://doi.org/10.14366/usg.23069

[14] Berzigotti A. Non-invasive evaluation of portal hypertension using ultrasound elastography. Journal of Hepatology. 2017; 67: 399–411. https://doi.org/10.1016/j.jhep.2017.02.003

[15] Min YW, Bae SY, Gwak GY, Paik YH, Choi MS, Lee JH, et al. A clinical predictor of varices and portal hypertensive gastropathy in patients with chronic liver disease. Clinical and Molecular Hepatology. 2012; 18: 178–184. https://doi.org/10.3350/cmh.2012.18.2.178

[16] Berzigotti A, Seijo S, Arena U, Abraldes JG, Vizzutti F, García-Pagán JC, et al. Elastography, spleen size, and platelet count identify portal hypertension in patients with compensated cirrhosis. Gastroenterology. 2013; 144: 102–111.e1. https://doi.org/10.1053/j.gastro.2012.10.001

[17] Hirooka M, Tanaka T, Koizumi Y, Yano R, Sunago K, Watanabe T, et al. Measurement of multiple spleen lengths is not necessary for non-invasive prediction of high-risk esophagogastric varices. Hepatology Research: the Official Journal of the Japan Society of Hepatology. 2022; 52: 187–198. https://doi.org/10.1111/hepr.13716

[18] Moini M, Onofrio F, Hansen BE, Adeyi O, Khalili K, Patel K. Combination of FIB-4 with ultrasound surface nodularity or elastography as predictors of histologic advanced liver fibrosis in chronic liver disease. Scientific Reports. 2021; 11: 19275. https://doi.org/10.1038/s41598-021-98776-1

[19] Han SK, Kim MY, Kang SH, Baik SK. Application of ultrasound for the diagnosis of cirrhosis/portal hypertension. Journal of Medical Ultrasonics (2001). 2022; 49: 321–331. https://doi.org/10.1007/s10396-022-01191-w

[20] Watanabe S, Hosomi N, Kitade Y, Kurokohchi K, Arima K, Kawabata H, et al. Assessment of the presence and severity of esophagogastric varices by splenic index in patients with liver cirrhosis. Journal of Computer Assisted Tomography. 2000; 24: 788–794. https://doi.org/10.1097/00004728-200009000-00022

[21] Fateh SM, Mohammed NA, Mahmood KA, Hasan AH, Tahir SH, Kakamad FH, et al. Sonographic measurement of splenic size and its correlation with body parameters. Medicine International. 2023; 3: 7. https://doi.org/10.3892/mi.2023.67

[22] Katwal S, Ansari MA, Suwal S, Rayamajhi S, Ghimire P, Ghimire A. Measurement of portal vein indices and splenic index by ultrasound and their association with gastroesophageal varices in cirrhosis of liver. Annals of Medicine and Surgery (2012). 2023; 85: 5926–5931. https://doi.org/10.1097/MS9.0000000000001483

[23] Colli A, Fraquelli M, Andreoletti M, Marino B, Zuccoli E, Conte D. Severe liver fibrosis or cirrhosis: accuracy of US for detection–analysis of 300 cases. Radiology. 2003; 227: 89–94. https://doi.org/10.1148/radiol.2272020193

[24] Bernardi M, Angeli P, Claria J, Moreau R, Gines P, Jalan R, et al. Albumin in decompensated cirrhosis: new concepts and perspectives. Gut. 2020; 69: 1127–1138. https://doi.org/10.1136/gutjnl-2019-318843