Diagn Interv Radiol 2025; DOI: 10.4274/dir.2025.253655

ABDOMINAL IMAGING

ORIGINAL ARTICLE



# Multi-organ non-contrast computed tomography radiomics model to predict hepatic encephalopathy in patients with cirrhosis and hepatorenal failure

Jin-ming Cao[1,2]
Ming-ya Zhang[3]
Xue-mei Ding[2]
Hai-ying Zhou[4]
Xiao-ming Zhang[1,4*]
Tian-wu Chen[5*]

[1]The First Clinical College of Jinan University, Department of Radiology, Guangzhou, China

[2]The Second Clinical Medical School of North Sichuan Medical College/Nanchong Central Hospital, Department of Radiology, Nanchong, China

[3]Nanjing University, Department of Computer Science, Xianlin Campus, Nanjing, China

[4]Medical Imaging Key Laboratory of Sichuan Province, Department of Radiology, Affiliated Hospital of North Sichuan Medical College, Nanchong, China

[5]The Second Affiliated Hospital of Chongqing Medical University, Department of Radiology, Chongqing, China

***Contributed equally to this work.**

**Corresponding author:** Tian-wu Chen

**E-mail:** tianwuchen_nsmc@163.com

Received 17 September 2025; revision requested 21 October 2025; last revision received 27 November 2025; accepted 05 December 2025.

Epub: 11.12.2025

Publication date:

DOI: 10.4274/dir.2025.253655

**PURPOSE**

To develop and validate a model by incorporating abdominal multi-organ non-contrast computed tomography (CT) radiomics and clinical features to predict the feasibility of hepatic encephalopathy (HE) occurrence in patients with cirrhosis and hepatorenal failure.

**METHODS**

In total, 351 consecutive patients with cirrhosis and hepatorenal failure undergoing non-contrast abdominal CT scans at Centers 1 and 2 were enrolled. Patients from Center 1 were randomly allocated to training (n = 191) and internal test (n = 81) groups, and those from Center 2 were assigned to the external test group (n = 79). The nnU-Net framework was used for automated three-dimensional (3D) segmentation of abdominal organs—the liver, spleen, portal and splenic vein, inferior vena cava, esophagogastric junction, stomach, liver-adjacent small bowel, and colon. Segmented multi-organ radiomics features were extracted using 3D Slicer, with R software used for feature processing and model construction. Model performance in predicting HE occurrence was evaluated using receiver operating characteristic (ROC) analysis in the training, internal test, and external test cohorts. Decision curve analysis (DCA) was used to evaluate clinical utility. The SHapley Additive exPlanations (SHAP) tool was used to provide a basis for model interpretability analysis.

**RESULTS**

In total, 351 patients (mean age, 61.3 ± 10.7 years; 231 men) were enrolled in this study. Esophageal variceal bleeding, peritonitis, and ascites were independent clinical predictors of HE. Twenty discriminative radiomics features, selected from the abovementioned multi-organs through intraclass correlation coefficient and least absolute shrinkage and selection operator analysis, were used to construct the radiomics model. The integrated model, incorporating both radiomics and clinical features, obtained higher areas under the ROC curve than the radiomics and clinical models in the training (0.87 vs. 0.83 vs. 0.68), internal test (0.85 vs. 0.81 vs. 0.66), and external test (0.83 vs. 0.78 vs. 0.72) cohorts, as evidenced by favorable integrated discrimination improvement values ($P < 0.05$ for all). The integrated model demonstrated superior clinical utility in DCA. Moreover, SHAP feature contribution analysis revealed that the top five features in terms of contribution were all extracted from the digestive tract.

**CONCLUSION**

The integrated model can effectively predict HE occurrence in patients with cirrhosis and hepatorenal failure.

**CLINICAL SIGNIFICANCE**

This novel model, developed by integrating abdominal multi-organ non-contrast CT radiomics and clinical features, demonstrates robust performance in predicting the occurrence of cirrhosis-related HE in patients with cirrhosis and hepatorenal failure. It thus provides a valuable tool for clinical decision-making, facilitating the prevention of this complication.

**KEYWORDS**

Liver cirrhosis, hepatic encephalopathy, computed tomography, radiomics, hepatorenal function

You may cite this article as: Cao JM, Zhang MY, Ding XM, Zhou HY, Zhang XM, Chen TW. Multi-organ non-contrast computed tomography radiomics model to predict hepatic encephalopathy in patients with cirrhosis and hepatorenal failure. *Diagn Interv Radiol.* 11 December 2025 DOI: 10.4274/dir.2025.253655 [Epub Ahead of Print].

Cirrhosis is highly prevalent globally and is a primary cause of morbidity and mortality in patients with chronic liver disease, accounting for 2.4% of global deaths in 2019.[1-3] With disease progression, cirrhosis often leads to hepatic decompensation, which in turn gives rise to severe complications such as hepatic encephalopathy (HE), esophageal variceal bleeding (EVB), and hepatorenal failure.[1-5] As a major complication of cirrhosis, HE affects more than 45% of patients with cirrhosis, causing substantial cognitive, psychomotor, and psychiatric impairments.[2] Its symptoms progress from subclinical or subtle psychomotor changes to confusion, somnolence, coma, or even death.[6,7] Notably, HE onset marks a critical turning point in cirrhosis progression, signifying the transition from the compensated to the decompensated stage. Consequently, patients with cirrhosis and HE have a considerably shorter median survival time, underscoring the urgency of early intervention.[6,8,9]

Radiomics-based machine learning involves extracting vast amounts of high-dimensional mineable data from radiological images, applying advanced characterization algorithms, and deriving imaging features.[10] This technology has been increasingly used in digital medical image analysis, supporting tasks such as automatic organ segmentation on computed tomography (CT) images.[11] By analyzing large datasets, machine learning equips computers with capabilities for pattern recognition, decision-making, and unbiased comprehensive image feature analysis, thereby facilitating holistic assessment of disease factors and potentially improving clinical outcome prediction accuracy.[2,12,13] Radiomics-based prediction models for HE risk assessment, treatment efficacy evaluation, and prognosis analysis are well established.[14-16] However, prior studies have been limited to single-center data, single-organ analysis, and reliance on contrast-enhanced CT.

**Main points**

- Hepatic encephalopathy (HE) is a severe complication in patients with cirrhosis; early clinical intervention can reduce the occurrence of this complication.
- Predictive models based on contrast-enhanced computed tomography (CT) radiomic features lack clinical utility for patients with hepatorenal failure.
- The integrated model, combining abdominal multi-organ non-contrast CT radiomics and clinical features, effectively predicts HE occurrence in patients with cirrhosis and hepatorenal failure.

Notably, patients with cirrhosis and hepatorenal failure are often unable to undergo contrast-enhanced CT, as underlying hepatic and renal impairment compromises the metabolism of contrast agents. Such impairment not only hinders contrast agent clearance but also further deteriorates hepatic and renal function. This critical limitation thus renders prior prediction models clinically unfeasible. Accordingly, this study aimed to develop and validate a multi-center model on the basis of radiomics features extracted from multiple organs on non-contrast abdominal CT scans to comprehensively predict HE occurrence in patients with cirrhosis and hepatorenal failure, thereby facilitating early clinical intervention.

## Methods

### Patients

The study protocol was approved by the Institutional Review Board of Nanchong Central Hospital [2024 (149)], with informed consent waived due to the study's retrospective nature. It was conducted in accordance with the Declaration of Helsinki and relevant ethical regulations.

Between February 2021 and December 2023, we consecutively collected the data of 411 patients with cirrhosis at two medical centers (Centers 1 and 2) in accordance with the following inclusion criteria: (1) cirrhosis was confirmed with laboratory examinations, radiological manifestations, and clinical findings according to the European Association for the Study of the Liver (EASL) and American Association for the Study of Liver Diseases clinical practice guidelines;[17,18] (2) all patients underwent biochemical workup and non-contrast abdominal CT scans within 1 week after initial hospitalization, and the quality of the CT data was satisfactory for image analysis according to the 5-point scale image-quality scoring system;[19] (3) patients did not experience HE before or during the initial hospitalization; and (4) the presence of hepatorenal failure. Hepatorenal failure was defined as concurrent abnormalities in liver and renal function parameters, which were confirmed via laboratory testing. Specifically, it included severe liver failure (total bilirubin ≥ 171 μmol/L and international normalized ratio ≥ 1.5) and renal insufficiency [estimated glomerular filtration rate (eGFR) < 45 mL/min/1.73 $m^2$], in which an eGFR < 30 mL/min/1.73 $m^2$ was defined as severe renal failure.[20] Other markers, including alanine transaminase, albumin, creatinine, and urine protein, were also used to confirm the presence of hepatorenal failure.

The exclusion criteria were as follows: (1) patients with intrahepatic or splenic space-occupying lesions (n = 31); (2) patients with a history of abdominal surgical procedures (n = 11); (3) patients complicated with other nervous system diseases (n = 5); (4) patients with incomplete clinical data (n = 3); and (5) patients lost to follow-up (n = 10). Based on the inclusion and exclusion criteria, 351 participants were enrolled in the study.

Patient follow-up was conducted in accordance with the following protocol. Monthly telephone assessment after discharge using cognitive assessment scales was conducted during the 2-year follow-up period, focusing on HE-related symptoms such as impaired calculation ability, memory deficits, or abnormal limb signs.[4,17] Patients with suspected HE or onset of related symptoms were advised to be admitted for definitive diagnosis, confirmed in accordance with EASL guidelines based on consistent neurological symptoms or test results. Follow-up ceased upon HE confirmation or after 2 years if undiagnosed, with final follow-up in May 2025.

Of these participants, 191 and 81 patients from Center 1 were randomly assigned to training and internal test groups, respectively, and 79 patients from Center 2 were allocated to the external test group. In both the training and test cohorts, patients with or without HE were divided into HE or non-HE subgroups, respectively, according to the follow-up results. Patient recruitment is illustrated in Figure 1.

### Computed tomography techniques

All the patients underwent non-contrast abdominal CT using a second-generation dual-source scanner (SOMATOM Definition Flash; Siemens Healthineers, Erlangen, Germany) in Center 1 and a 120-slice multi-detector spiral CT scanner (uCT710; United Imaging Healthcare Co., Ltd., Shanghai, China) in Center 2. The CT scanning parameters were a peak voltage of 140 kV, a tube current of 240 mA, detector collimation of 128 × 0.125 mm, a pitch of 1.75, slice thickness of 1.0 mm, and a matrix of 512 × 512 mm. All patients were required to fast for 6–8 hours prior to CT scanning and drink 800–1,000 mL of water 20 minutes before the procedure to distend the gastrointestinal tract. Each patient lay supine on the scanning bed and maintained breath-hold during the non-contrast abdominal CT acquisition. The

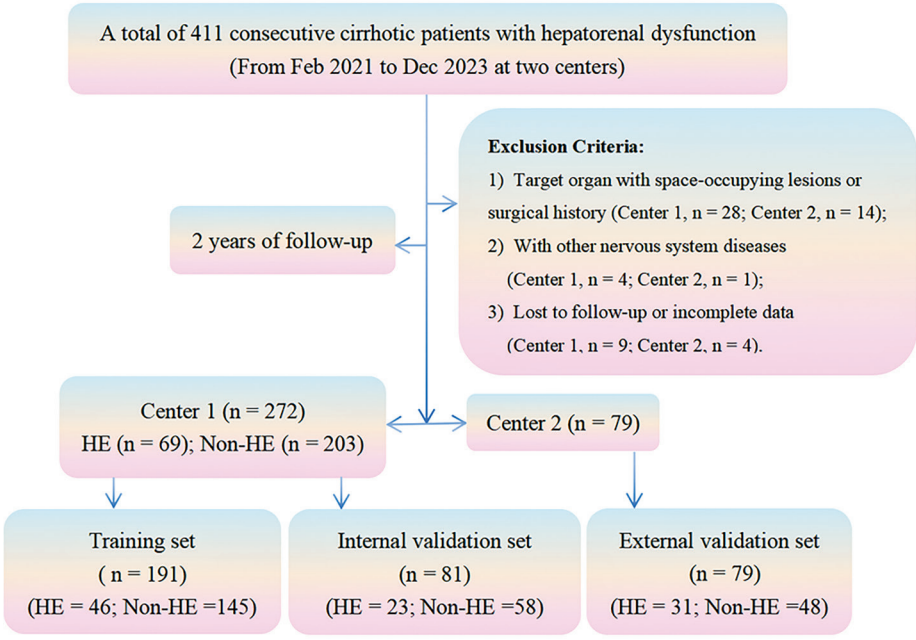


**Figure 1.** Patient selection flowchart. HE, hepatic encephalopathy.

scanning coverage extended from the dome of the diaphragm to the pubic symphysis. Subsequently, the CT data were transferred to the picture archiving and communication system (PACS), and all image data were saved in Digital Imaging and Communications in Medicine (DICOM) format for subsequent image analysis.

## Segmentation of multi-organ on computed tomography imaging

The CT data from both centers were extracted from the PACS system following image quality assessment. These DICOM files were then converted to NIfTI format (.nii.gz), and images were adjusted to standard abdominal window settings [width: 400 Hounsfield unit (HU); level: 50 HU]s to facilitate subsequent segmentation. Preprocessed images were homogenized via resampling with an isotropic resolution of 1×1×1 $mm^3$ and via Z-score normalization to reduce systematic dataset variations, ensuring reliability for subsequent structural segmentation, feature extraction, and model construction. After completing CT image quality assessment and matching the images to clinical records, three-dimensional (3D) Slicer was used for additional preprocessing, including format validation by checking NIfTI compatibility for nnU-Net recognition, and visualization checks via multiplanar reconstruction to exclude poorly covered images. The preprocessed data were fed into the nnU-Net framework, an open-source deep learning tool developed by the German Cancer Research Center and implemented in Python, for automated segmentation of abdominal anatomical structures in non-contrast CT images using "TotalSegmentator".[21] The liver, spleen, portal and splenic vein, inferior vena cava, esophagogastric junction, stomach, liver-adjacent small bowel, and colon were selected as the target structures for subsequent radiomics feature extraction (Figure 2a-f). The selection of these nine organs was hypothesis driven, based on their clear links to HE pathogenesis in our target population. The liver is key for ammonia metabolism; the spleen and portal/splenic veins reflect portal hypertension; the stomach, esophagogastric junction, duodenum, and liver-adjacent bowel/colon relate to the gut–brain axis; and the inferior vena cava reflects systemic venous return changes in patients with cirrhosis and hepatorenal failure.[2,6,7]

## Evaluation of segmentation result accuracy

To validate the accuracy of automated multi-organ segmentation, all targeted anatomical structures segmented by nnU-Net underwent qualitative and quantitative assessments. For qualitative evaluation, two radiologists (first and third authors with 13 and 15 years of experience in abdominal imaging, respectively) reviewed the segmentation masks to assess alignment between automated segmentation boundaries and actual anatomical contours, with discrepancies resolved through consensus. For quantitative assessment, 20 randomly selected consecutive cases were semi-automatically segmented by the first author using 3D Slicer. The Dice similarity coefficient (DSC) was calculated to compare automated and semi-automated results, with a DSC ≥ 0.85 for all structures defined as the reliability criterion for subsequent feature extraction.[22]

## Computed tomography radiomics feature extraction

After image segmentation and accuracy assessment, CT radiomics features were extracted from the validated volume of interest (VOI) using the "Pyradiomics" extension in 3D Slicer. A total of 1,107 radiomics features were initially extracted from the nine target organs, categorized into four major types: (1) morphological features (e.g., surface area, volume, compactness), reflecting organ size and shape; (2) first-order statistical features (e.g., mean, median, skewness, kurtosis), reflecting the distribution of CT attenuation values within the VOI; (3) texture features from gray-level co-occurrence matrices (GLCMs), gray-level run-length matrices (GLRLMs), and gray-level size-zone matrices (GLSZMs) (e.g., contrast, correlation, energy, entropy), capturing spatial patterns of pixel intensity variations; and (4) higher-order features derived from wavelet transforms, quantifying texture information across multiple scales.

## Consistency analysis for radiomics feature selection

To ensure the stability and reliability of radiomics features, intraclass correlation coefficient (ICC) analysis was performed on 20 randomly selected cases. For these cases, radiomics features were extracted from both the initial automatic segmentation results via nnU-Net and the semi-automated segmentation results generated by the first author. ICC analysis was then used to compare the two feature sets, with those demonstrating an ICC > 0.8 retained as indicators of good reproducibility for subsequent analyses.[17] This approach enhanced the stability of features, mitigated potential biases from variability in image acquisition and segmentation, and laid a robust foundation for subsequent feature optimization.

## Feature selection and model construction

After ICC analysis, R software (version 4.5.1) was used to implement a rigorous pipeline encompassing feature optimization and selection, model construction, and performance evaluation. The packages included "rms," "glmnet," "caret," "pROC," "dplyr," and "ggplot2."

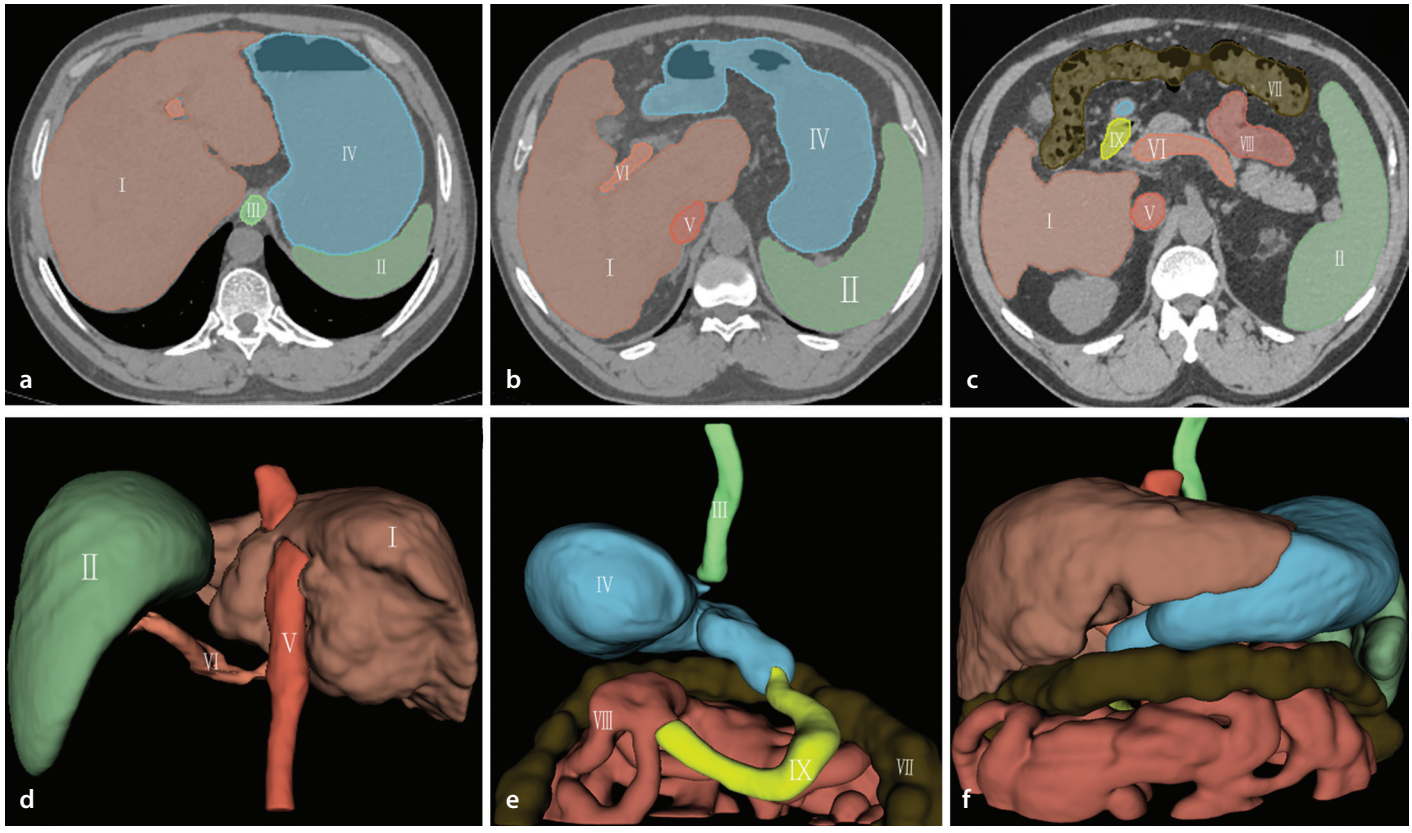


**Figure 2.** Example of automatic segmentation of multiple abdominal target organs. Images **(a-c)** show contour delineations of three two-dimensional cross sections. Images **(d-f)** depict three-dimensional visualizations of the segmented organs. I, liver; II, spleen, III, esophagogastric junction;s IV, stomach; V, inferior vena cava; VI, portal vein and splenic vein; VII, colon; VIII, small bowel; IX, duodenum.

The dataset from Center 1 was first split into training and internal test sets at a 7:3 ratio. Quantitative features in the training set were standardized using the Z-score method, with the same transformation parameters applied to the test set to avoid data leakage. Sample weights were calculated for the training set to address class imbalance. For features in the training set, Pearson correlation matrices with Holm–Bonferroni correction were computed to identify feature pairs with $|r| > 0.8$ (adjusted $P < 0.05$). Next, the variance inflation factor (VIF) was used to quantify multicollinearity (with VIF > 5 indicating significant multicollinearity). The specific feature that exhibits the highest VIF was explicitly identified and prioritized for removal to mitigate multicollinearity. Subsequently, a 10-fold cross-validated least absolute shrinkage and selection operator (LASSO) analysis was applied to the preprocessed training set using the binomial family with L1 regularization. The optimal regularization parameter (λ) was determined by minimizing the cross-validation error (λ-min). Feature stability was evaluated through bootstrapping (100 iterations), retaining only variables with Cohen's d > 0.5 selected in >70% of resamples to ensure robustness. A radiomics score (RS) was calculated for each case in the training, internal test, and external test sets using the predict function [predict(model, newdata = feature_data, type = "link")] in R software, and this function output the model's linear predictor. The calculation uses the following formula: $RS = \beta_0 + \beta_1X_1 + ... + \beta_nX_n$, where β represents regression coefficients and X represents radiomics features. Subgroup differences in the RS between patients with and without HE were compared.

Additionally, univariate analysis for clinical features was used to identify clinically important factors between the HE and non-HE groups. For this analysis, categorical factors, including sex, EVB, peritonitis, and ascites, were assessed using the chi-square test, and the continuous factor (age) was evaluated with the independent samples t-test.

Based on insights from previous radiomics and clinical feature analyses, three predictive models were developed using multivariate logistic regression—a clinical model, a radiomics model, and an integrated model—by combining radiomics features with the clinical factors. All three models were trained through 10-fold cross-validation in the training set, with coefficients optimized using maximum likelihood estimation, and predictive performance of the three models was evaluated in both the internal and external test sets. To interpret the predictive contribution of radiomics features, SHapley Additive exPlanations (SHAP) analysis was conducted. Mean absolute SHAP values were computed to quantify the importance of each radiomics feature, whereas SHAP beeswarm plots and feature importance bar plots were employed to visualize the impact of individual features and their contributions across samples.

## Statistical analysis

The performance of the models was evaluated using receiver operating characteristic (ROC) curve analysis and area under the ROC curves (AUCs), with the ROC test used to compare model performance. Based on the predicted classifications of the three models, confusion matrices were derived, and accuracy, sensitivity, and specificity were calculated for all datasets. Additionally, Brier scores were used to assess the calibration of predicted probabilities for each model, and integrated discrimination improvement (IDI) was employed to compare their discriminative ability.[23] DeLong's test was used to evaluate model overfitting across datasets.

Violin plots were used to visually depict the distributional differences in the RS between groups with and without HE. For baseline characteristics comparisons, normally distributed continuous variables were analyzed using the independent t-test; otherwise, the Mann–Whitney U test was used. Categorical variables were analyzed using the chi-square test or Fisher's exact test, as appropriate. A two-sided $P < 0.05$ was considered statistically significant.

## Results

### Baseline characteristics

The development set included 272 patients (173 men, 99 women), aged 61.2 ± 10.5 years (mean ± standard deviation), with an age range of 22–86 years. Of these, 69 (25.4%) had HE and 203 (74.6%) did not. The external test set comprised 79 patients (58 men, 21 women), aged 61.6 ± 11.1 years, with an age range of 34–87 years, including 31 (39.2%) with HE and 48 (60.8%) without. No significant differences were observed between sets ($P > 0.05$ for all). Among the clinical features, ascites, peritonitis, and EVB were identified as independent predictors of HE in the development and test sets ($P < 0.05$ for all) (Table 1).

### Radiomics feature extractions and selections

Across all target organs (the liver, spleen, portal and splenic veins, inferior vena cava, esophagogastric junction, stomach, and liver-adjacent small bowel and colon), the median DSC value between automated and semi-automated segmentations was 0.91 (ranging from 0.87 to 0.95), indicating the accuracy of automated multi-organ segmentation. Based on ICC analysis, 119 initial features with an ICC ≤ 0.8 were excluded, and the remaining 988 features with an ICC > 0.8 were retained for subsequent radiomics feature selection. To address multicollinearity, the feature with the highest VIF was removed, followed by 10-fold cross-validated LASSO logistic regression and correlation analysis. Eventually, a total of 20 discriminative radiomics features were selected (Figure 3a, b).

For these 20 selected features, categories included 3 first-order statistical features, 13 texture features (including 4 from GLCMs, 1 from GLRLMs, 3 from neighboring gray-tone difference matrices, 2 from gray-level dependence matrices, and 3 from GLSZMs), and 4 shape features. Based on the distribution across organs, the features were derived as follows: two from the spleen, one from the liver, one from the stomach, one from the

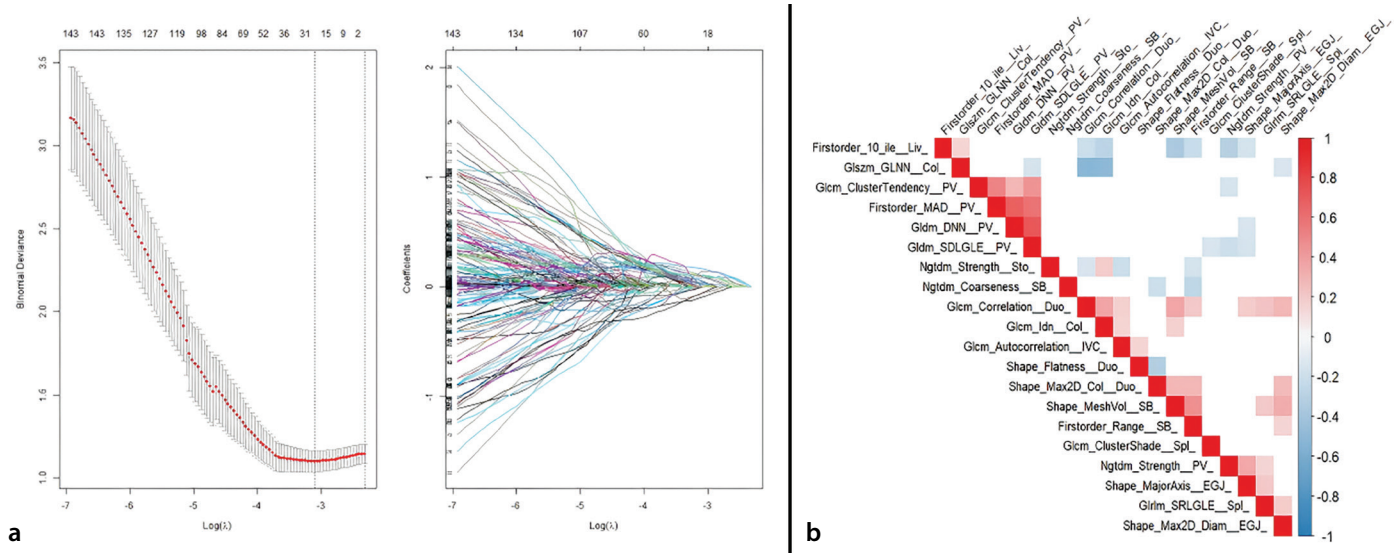


**Figure 3.** Feature selection using least absolute shrinkage and selection operator. Image **(a)** shows 10-fold cross-validation used for parameter tuning to obtain the best lambda, which is subsequently used for feature selection. Image **(b)** shows the correlation of selected features without multicollinearity in a heatmap. Spl, spleen; Liv, liver; Sto, stomach; IVC, inferior vena cava; PV, portal vein; EGJ, esophagogastric junction; SB, small bowel; Duo, duodenum; Col, colon.

**Table 1.** Patient characteristics in the development and external test sets

| Characteristics | Description | Development set (n = 272) | | | Test set (n = 79) | | |
|---|---|---|---|---|---|---|---|
| | | HE (n = 69) | Non-HE (n = 203) | *P* value | HE (n = 31) | Non-HE (n = 48) | *P* value |
| Gender, n (%) | Women | 25 (36.2%) | 74 (36.5%) | 1.0 | 8 (25.8%) | 13 (27.1%) | 1.0 |
| | Men | 44 (63.8%) | 129 (63.5%) | | 23 (74.2%) | 35 (72.9%) | |
| Age, years | Mean ± SD | 59.9 ± 11.2 | 61.7 ± 10.3 | 0.23 | 62.35 ± 11.11 | 61.17 ± 11.25 | 0.65 |
| EVB, n (%) | Y | 30 (43.5%) | 58 (28.6%) | 0.03 | 16 (51.6%) | 9 (18.8%) | 0.005 |
| | N | 39 (56.5%) | 145 (71.4%) | | 15 (48.4%) | 39 (81.2%) | |
| Ascites, n (%) | Y | 40 (58.0%) | 66 (32.5%) | <0.001 | 14 (45.2%) | 7 (14.6%) | 0.006 |
| | N | 29 (42.0%) | 137 (67.5%) | | 17 (54.8%) | 41 (85.4%) | |
| Peritonitis, n (%) | Y | 55 (79.7%) | 118 (58.1%) | 0.002 | 22 (71.0%) | 21 (43.8%) | 0.03 |
| | N | 14 (20.3%) | 85 (41.9%) | | 27 (56.2%) | 27 (56.2%) | |

HE, hepatic encephalopathy; SD, standard deviation; EVB, esophageal variceal bleeding; Y, yes; N, no.

inferior vena cava, four from the portal and splenic veins, two from the esophagogastric junction, three from the liver-adjacent small bowel, three from the liver-adjacent colon, and three from the duodenum.

Violin plots revealed distinct distributions of RSs between the HE and non-HE subgroups across the three cohorts. Patients with HE exhibited significantly higher RSs than those without HE ($P < 0.001$ for all), indicating that the RS effectively discriminated between both subgroups (Figure 4a-c).

### Performance of predictive models

Three predictive models were constructed: a radiomics model (incorporating selected radiomics features), a clinical model (incorporating selected clinical features), and an integrated model (combining all selected radiomics and clinical features). The performance and stability of all three models were evaluated using AUCs, accuracy, sensitivity, specificity, and Brier scores. The integrated model demonstrated superior clinical utility through decision curve analysis (DCA), revealing a higher net benefit at a range of 0.2–0.8 clinically relevant threshold probabilities (Figure 5a-d). Additionally, the integrated model showed significantly better predictive accuracy than the radiomics model, as evidenced by the IDI values, with the differences reaching statistical significance ($P < 0.05$) (Table 2). Moreover, SHAP feature contribution analysis revealed that the top five features in terms of contribution were extracted from the digestive tract (stomach, duodenum, liver-adjacent small bowel, and colon), including two texture features and three shape features (Figure 6a, b).

### Overfitting assessment for radiomics and integrated models

Overfitting was assessed by comparing model performance across datasets using DeLong's test (Table 3). For both models, pairwise comparisons between the training and internal and external test sets showed no statistically significant differences in AUCs ($P > 0.05$ for all), verifying that the models had no substantial overfitting risk.

## Discussion

This study developed and tested three predictive models for HE prediction in patients with cirrhosis and hepatorenal failure using multi-organ radiomics from non-contrast abdominal CT and clinical features. The integrated model, by combining multi-organ radiomics and clinical features, showed robust performance across the training, internal test, and external test cohorts, with AUCs of 0.87, 0.85, and 0.83, respectively; the absence of significant overfitting risk further strengthened its potential for clinical utility. Within these models, the RS demonstrated strong discriminative power for HE, with significantly higher values across cohorts. Additionally, three clinical features (ascites, EVB, and peritonitis) were identified as independent HE predictors, and their integration with radiomics features demonstrated superior performance, which strongly reinforced the merit of combining radiomics with clinical insights. This aligns with established mechanisms: ascites indicates decompensated cirrhosis, a critical transition from compensated liver dysfunction to overt failure of hepatic synthetic and metabolic capacity;[24] peritonitis triggers systemic inflammation, disrupting hepatic clearance of neurotoxins;[25] and variceal bleeding increases the ammonia burden by releasing nitrogen-rich blood into the gut, where bacterial metabolism further raises toxic ammonia levels.[26]

Notably, the radiomics features in this study were derived from nine abdominal organs, encompassing three categories: first-order statistical, texture, and shape features. This multi-organ strategy offers a distinct advantage over conventional radiomics studies, which often focus on isolated organs and thus overlook systemic pathological changes in cirrhosis.[2,4] By contrast, our multiple-organ coverage comprehensively captures multi-system dysfunction, giving it a critical advantage for evaluating complex disorders such as HE. Beyond features from well-studied organs (liver, spleen, and portal vein), 11 gastrointestinal radiomics features from the stomach, esophagogastric junction, liver-adjacent small bowel, and colon add unique value. These features not only highlight associations between portal hypertensive gastroenteropathy and decompensated cirrhosis but also provide radiomics-based evidence supporting the gut–brain axis and gut microbiota theories in HE pathogenesis. Specifically, the top five contributing radiomics features identified by SHAP extracted from the stomach, duodenum, liver-adjacent small bowel, and colon indicate that our observations align closely with the previously proposed gut–brain axis theory, corroborating its critical association with HE development; these features are insufficiently represented in radiomics studies limited to hepatic or splenic features.[2,27,28] More importantly, our logistic regression-based model offers superior interpretability, allowing clinicians to clearly quantify how multi-organ radiomics features contribute to HE risk predictions, which is critical for clinical trust and model adoption.

Critically, we employed non-contrast CT, a deliberate choice tailored to patients with cirrhosis and hepatorenal failure. Contrast agents are known to exacerbate renal dysfunction in this population,[20] yet many existing HE prediction models rely on contrast-enhanced CT,[2] restricting their applicability. Our use of non-contrast CT avoids this risk, leverages routine clinical data, and reduces costs, facilitating its integration into

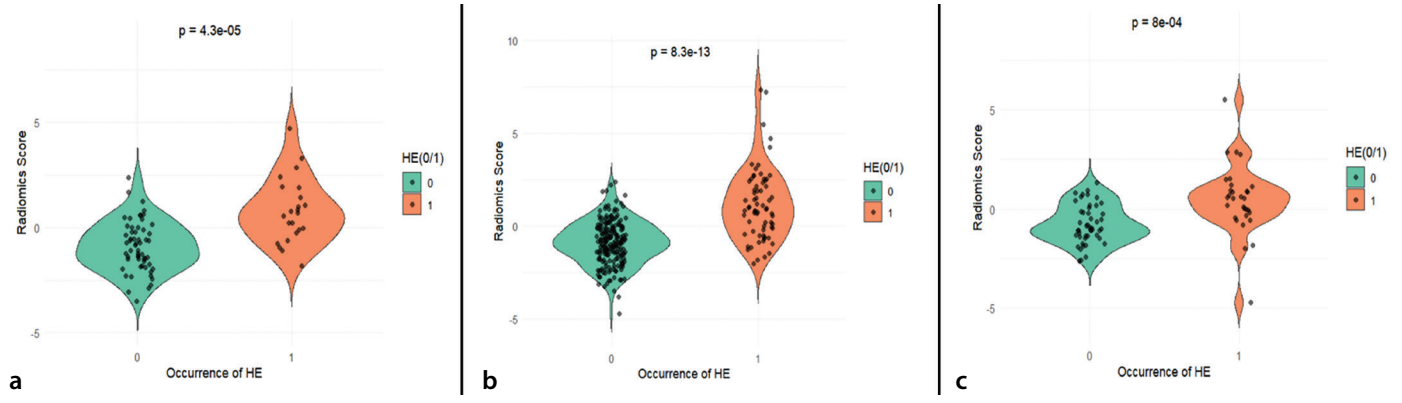


**Figure 4.** Violin plots show the distribution of radiomics scores. Training set **(a)**, internal test set **(b)**, and external test set **(c)**. HE, hepatic encephalopathy; 1, with hepatic encephalopathy; 0, without hepatic encephalopathy.

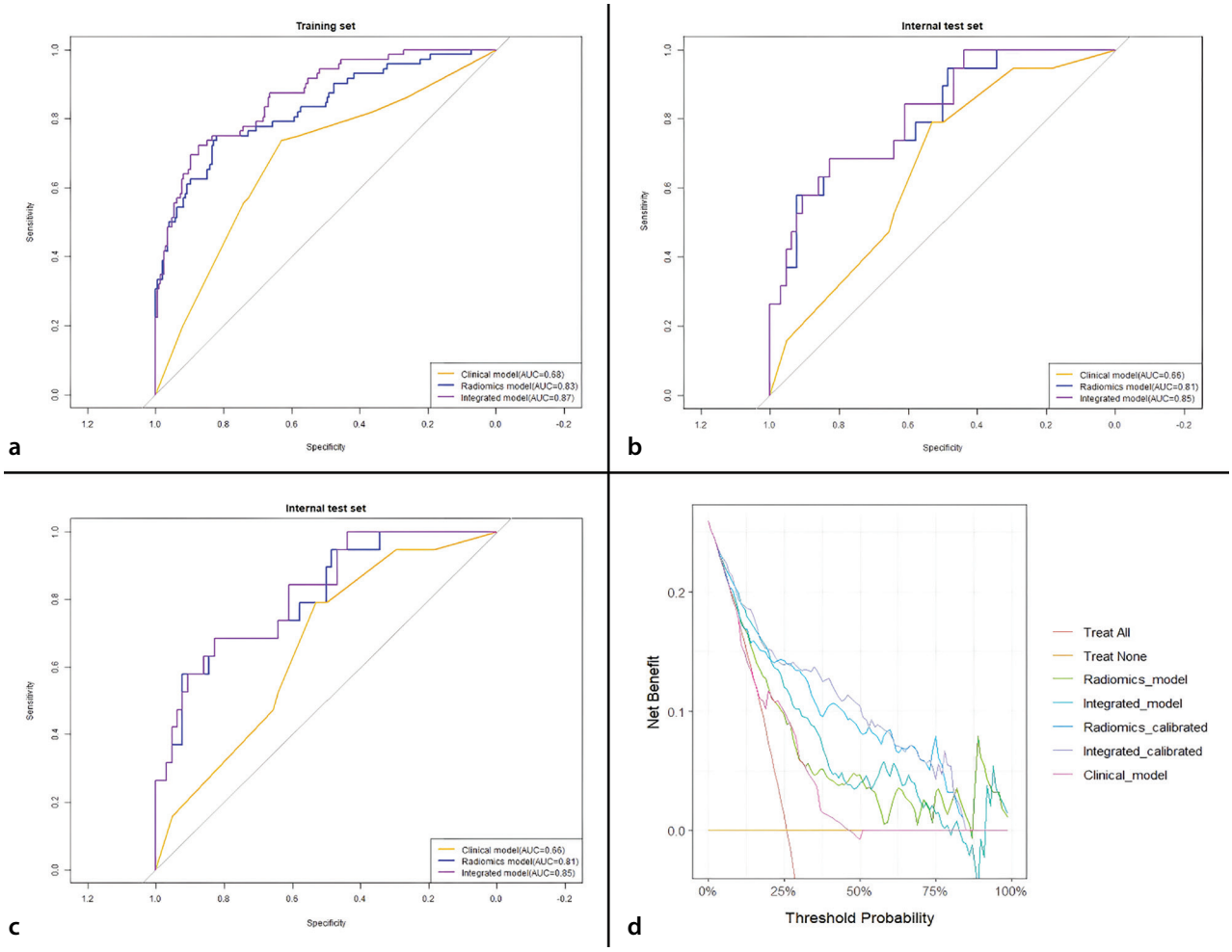


**Figure 5.** Receiver operating characteristic curves comparing the integrated clinical and radiomics models in the training set **(a)**, internal test set **(b)**, and external test set **(c)**. Decision curve analysis depicts net benefit curves to compare models across threshold probabilities, highlighting the integrated model's advantage **(d)**.

**Table 2.** Prediction performance of models in each dataset

| Model categories | Metrics | Training set | Internal test set | External test set |
|---|---|---|---|---|
| Clinical model | AUC (95% CI) | 0.68 (0.61–0.76) | 0.66 (0.54–0.79) | 0.72 (0.61–0.84) |
| | Acc | 0.66 | 0.59 | 0.72 |
| | Sen | 0.74 | 0.79 | 0.61 |
| | Spe | 0.63 | 0.53 | 0.79 |
| | Brier score | 0.1749 | 0.1648 | 0.1979 |
| Radiomics model | AUC (95% CI) | 0.83 (0.77–0.89) | 0.81 (0.71–0.92) | 0.78 (0.66–0.89) |
| | Acc | 0.85 | 0.75 | 0.72 |
| | Sen | 0.62 | 0.74 | 0.87 |
| | Spe | 0.92 | 0.76 | 0.63 |
| | Brier score | 0.1284 | 0.1498 | 0.1815 |
| Integrated model | AUC (95% CI) | 0.87 (0.82–0.91) | 0.85 (0.76–0.94) | 0.83 (0.74–0.93) |
| | Acc | 0.84 | 0.74 | 0.79 |
| | Sen | 0.64 | 0.91 | 0.77 |
| | Spe | 0.91 | 0.67 | 0.79 |
| | Brier score | 0.1212 | 0.1397 | 0.1569 |
| IDI (*P* value) | IM vs. CM | < 0.001 | < 0.001 | < 0.001 |
| | IM vs. RM | < 0.001 | 0.043 | < 0.001 |

AUC, area under the receiver operating characteristic curve; CI, confidence interval; IDI, integrated discrimination improvement; Acc, accuracy; Sen, sensitivity; Spe, specificity; IM, integrated model; CM, clinical model; RM, radiomics model.

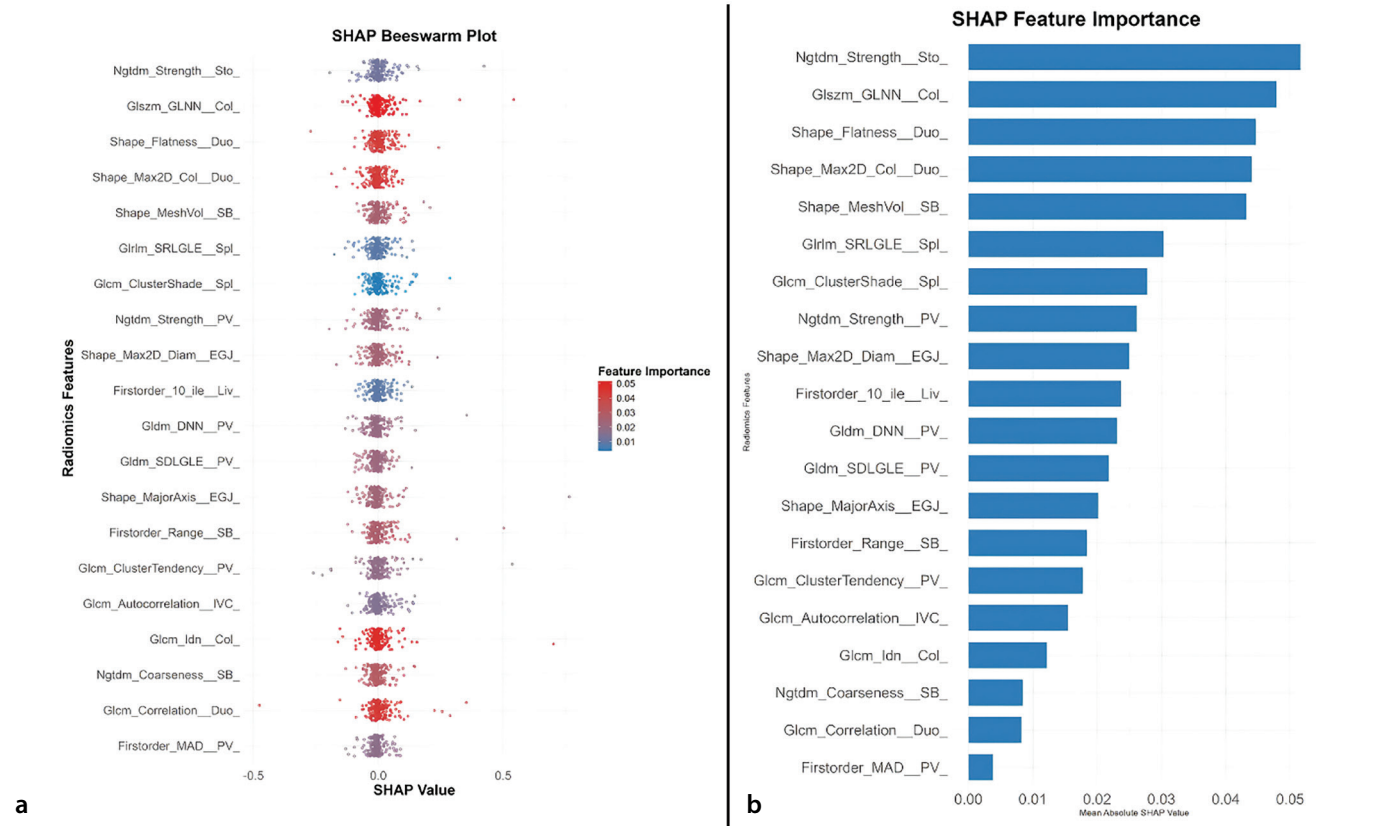


**Figure 6.** SHapley Additive exPlanations (SHAP) beeswarm plot with radiomics features on the y-axis (ordered by descending importance), SHAP values on the x-axis (denoting prediction impact), color-coded importance (red = high, blue = low), and dots representing each feature's per-sample contribution **(a)**. SHAP feature importance bar plot ranking radiomics features by their mean absolute SHAP values in descending order to illustrate their impact on model predictions **(b)**. Spl, spleen; Liv, liver; Sto, stomach; IVC, inferior vena cava; PV, portal vein; EGJ, esophagogastric junction; SB, small bowel; Duo, duodenum; Col, colon.

**Table 3.** Overfitting assessment across datasets using DeLong's test

| Model | Comparison group | AUC (TR-set) | AUC (TE-set) | *P* value | Overfitting (Y or N) |
|---|---|---|---|---|---|
| Clinical model | T-set vs. IT-set | 0.68 | 0.66 | 0.80 | N |
| | T-set vs. ET-set | 0.68 | 0.72 | 0.60 | N |
| Radiomics model | T-set vs. IT-set | 0.83 | 0.81 | 0.81 | N |
| | T-set vs. ET-set | 0.83 | 0.78 | 0.41 | N |
| Integrated model | T-set vs. IT-set | 0.87 | 0.85 | 0.74 | N |
| | T-set vs. ET-set | 0.87 | 0.83 | 0.56 | N |

TR-set, training set; IT-set, internal test set; ET-set, external test set; TE-set, test set; AUC, area under receiver operating characteristic curve, Y, yes; N, no.

practice, particularly in resource-limited settings. Notably, a prior study reported a weighted random forest model with an AUC of 0.82 for HE prediction in general patients with cirrhosis.[29] By contrast, our integrated model achieves comparable or superior performance, with an AUC ranging from 0.83 to 0.87, specifically for patients with cirrhosis and hepatorenal failure in which contrast-enhanced CT is contraindicated. This underscores our model's advantage in both filling a clinical gap and maintaining robust accuracy. These methodological strengths are further supported by our DCA results. Specifically, the integrated calibrated model delivered superior net benefit at a 0.2–0.8 threshold, which matches the clinical intervention time-point for HE prevention in patients with cirrhosis. Compared with the clinical and radiomics model, this advantage addresses the unmet need for early HE risk stratification, helping clinicians distinguish between low-risk and medium-risk patients to balance the benefits with overtreatment.

This study has several limitations. First, cirrhosis was not stratified by etiology; however, our integrated model had broad potential applications despite HE resulting from any causes. We will design cohort studies stratified by different etiologies to enhance the model's generalizability in future research. Second, although automatic segmentation via nnU-Net achieved high DSC values, its reliability may be decreased in cases of severe anatomical distortion (e.g., massive ascites or organ displacement). Follow-up studies could integrate manual correction for complex cases or explore more advanced segmentation algorithms to improve feature stability. Third, air in the gastrointestinal lumen may still introduce residual bias. Subsequent studies could use water filling or artificial intelligence-based air segmentation to further reduce this interference, ensuring more accurate extraction of radiomics features. Fourth, differences in CT scanners may introduce systematic biases in radiomics feature

extraction despite standardization through Z-score normalization. Future studies should include device-specific calibration or multi-center harmonization protocols to mitigate this issue. In conclusion, extending the follow-up period or evaluating longitudinal changes in radiomics features may provide new insights into HE progression; we plan to investigate this in subsequent studies.

In conclusion, our combined model, integrating multi-organ radiomics features on abdominal non-contrast CT and important clinical factors, demonstrated robust performance in the prediction of HE secondary to cirrhosis in patients with hepatorenal failure. This model could provide valuable evidence for clinical decision-making regarding potential HE in this population.

### Footnotes

### Conflict of interest disclosure

The authors declared that there is no conflict of interest.

## References


1. Ginès P, Krag A, Abraldes JG, Solà E, Fabrellas N, Kamath PS. Liver cirrhosis. *Lancet*. 2021;398(10308):1359-1376. [Crossref]
2. Cao JM, Yang JQ, Ming ZQ, et al. A radiomics model of liver CT to predict risk of hepatic encephalopathy secondary to hepatitis B related cirrhosis. *Eur J Radiol*. 2020;130:109201. [Crossref]
3. Huang DQ, Terrault NA, Tacke F, et al. Global epidemiology of cirrhosis - aetiology, trends and predictions. *Nat Rev Gastroenterol Hepatol*. 2023;20(6):388-398. [Crossref]
4. Tan BG, Tang Z, Ou J, et al. An improved model based on quantitative features of right liver lobe, maximum varices, and portal vein system measured on magnetic resonance imaging to predict oesophagogastric variceal haemorrhage secondary to hepatitis B-related cirrhosis. *Quant Imaging Med Surg*. 2023;13(12):7741-7752. [Crossref]
5. Tapper EB, Ufere NN, Huang DQ, Loomba R. Review article: current and emerging therapies for the management of cirrhosis and its complications. *Aliment Pharmacol Ther*. 2022;55(9):1099-1115. [Crossref]
6. Louissaint J, Deutsch-Link S, Tapper EB. Changing epidemiology of cirrhosis and hepatic encephalopathy. *Clin Gastroenterol Hepatol*. 2022;20(8S):S1-S8. [Crossref]
7. Häussinger D, Dhiman RK, Felipo V, et al. Hepatic encephalopathy. *Nat Rev Dis Primers*. 2022;8(1):43. [Crossref]
8. 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(1):217-231. [Crossref]
9. Tapper EB, Aberasturi D, Zhao Z, Hsu CY, Parikh ND. Outcomes after hepatic encephalopathy in population-based cohorts of patients with cirrhosis. *Aliment Pharmacol Ther*. 2020;51(12):1397-1405. [Crossref]
10. Mirestean CC, Pagute O, Buzea C, Iancu RI, Iancu DT. Radiomic machine learning and texture analysis - new horizons for head and neck oncology. *Maedica (Bucur)*. 2019;14(2):126-130. [Crossref]
11. Mai DVC, Drami I, Pring ET, et al. A systematic review of automated segmentation of 3D computed-tomography scans for volumetric body composition analysis. *J Cachexia Sarcopenia Muscle*. 2023;14(5):1973-1986. [Crossref]
12. Skrede OJ, De Raedt S, Kleppe A, et al. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *Lancet*. 2020;395(10221):350-360. [Crossref]
13. Rakaee M, Tafavvoghi M, Ricciuti B, et al. Deep learning model for predicting immunotherapy response in advanced non-small cell lung cancer. *JAMA Oncol*. 2025;11(2):109-118. [Crossref]
14. Bo Z, Song J, He Q, et al. Application of artificial intelligence radiomics in the diagnosis, treatment, and prognosis of hepatocellular carcinoma. *Comput Biol Med*. 2024;173:108337. [Crossref]
15. Avanzo M, Wei L, Stancanello J, et al. Machine and deep learning methods for radiomics. *Med Phys*. 2020;47(5):e185-e202. [Crossref]
16. Chen B, Zhang R, Gan Y, Yang L, Li W. Development and clinical application of radiomics in lung cancer. *Radiat Oncol*. 2017;12(1):154. [Crossref]
17. European Association for the Study of the Liver. EASL Clinical Practice Guidelines on the management of hepatic encephalopathy. *J Hepatol*. 2022;77(3):807-824. [Crossref]
18. Karvellas CJ, Bajaj JS, Kamath PS, et al. AASLD Practice guidance on acute-on-chronic liver failure and the management of critically ill patients with cirrhosis. *Hepatology*. 2024;79(6):1463-1502. [Crossref]
19. Chen TW, Yang ZG, Dong ZH, et al. Whole tumour first-pass perfusion using a low-dose method with 64-section multidetector row computed tomography in oesophageal squamous cell carcinoma. *Eur J Radiol*. 2011;80(2):284-91. [Crossref]
20. Lu S, Robyak K, Zhu Y. The CKD-EPI 2021 equation and other creatinine-based race-independent eGFR equations in chronic kidney disease diagnosis and staging. J Appl Lab Med. 2023;8(5):952-961. [Crossref]
21. Wasserthal J, Breit HC, Meyer MT, et al. TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. *Radiol Artif Intell*. 2023;5(5):e230024. [Crossref]
22. Häntze H, Xu L, Mertens CJ, et al. Segmenting whole-body MRI and CT for multiorgan anatomic structure delineation. *Radiol Artif Intell*. 2025;7(6):e240777. [Crossref]
23. Huang S, Liang X, Lou K, et al. Comparing habitat, radiomics, and fusion models for predicting micropapillary/solid components in stage I lung adenocarcinoma. *Acad Radiol*. 2025;32(10):6307-6319. [Crossref]
24. Gracia-Sancho J, Marrone G, Fernández-Iglesias A. Hepatic microcirculation and mechanisms of portal hypertension. *Nat Rev Gastroenterol Hepatol*. 2019;16(4):221-234. [Crossref]
25. Lima LCD, Miranda AS, Ferreira RN, Rachid MA, Simões E Silva AC. Hepatic encephalopathy: lessons from preclinical studies. *World J Hepatol*. 2019;11(2):173-185. [Crossref]
26. Jindal A, Jagdish RK. Sarcopenia: ammonia metabolism and hepatic encephalopathy. *Clin Mol Hepatol*. 2019;25(3):270-279. [Crossref]
27. Agirman G, Yu KB, Hsiao EY. Signaling inflammation across the gut-brain axis. *Science*. 2021;374(6571):1087-1092. [Crossref]
28. Fried DE, Watson RE, Robson SC, Gulbransen BD. Ammonia modifies enteric neuromuscular transmission through glial γ-aminobutyric acid signaling. *Am J Physiol Gastrointest Liver Physiol*. 2017;313(6):G570-G580. [Crossref]
29. Yang H, Li X, Cao H, et al. Using machine learning methods to predict hepatic encephalopathy in cirrhotic patients with unbalanced data. *Comput Methods Programs Biomed*. 2021;211:106420. [Crossref]