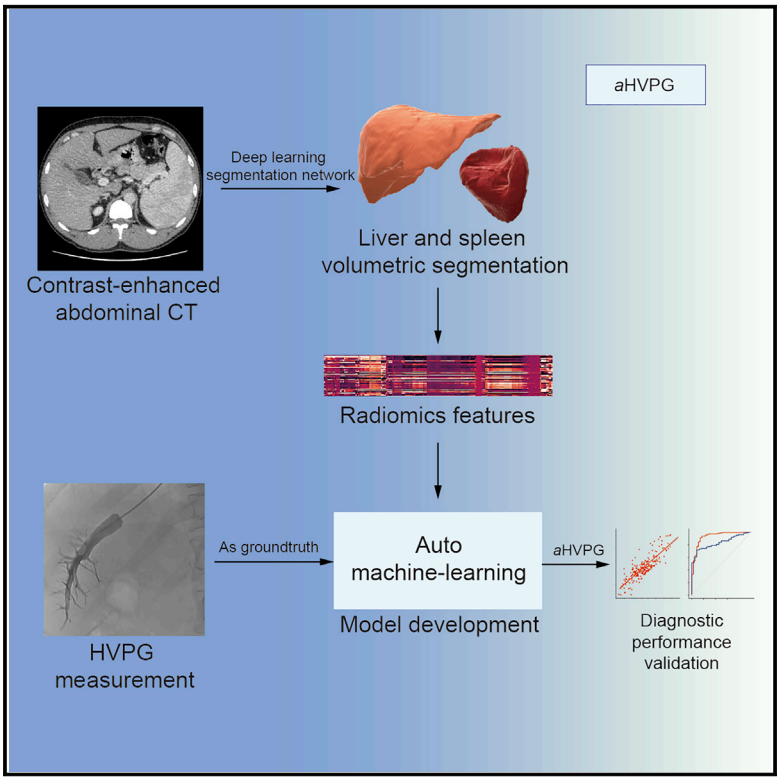
Cell Reports
Medicine

Article

An imaging-based artificial intelligence model for non-invasive grading of hepatic venous pressure gradient in cirrhotic portal hypertension

Graphical abstract



Highlights

- *a*HVPG is an automated HVPG quantitative estimation model based on CT
- *a*HVPG has the potential to assess HVPG and outperforms other non-invasive tools
- Non-invasive tools may help PHT monitoring when invasive HVPG is not available

Authors

Qian Yu, Yifei Huang, Xiaoguo Li, ..., Özgün Ömer Asiller, Shenghong Ju, Xiaolong Qi

Correspondence

jsh@seu.edu.cn (S.J.), qixiaolong@vip.163.com (X.Q.)

In brief

Yu et al. present an automated hepatic venous pressure gradient (HVPG) quantitative estimation model, which can improve diagnostic performance compared with conventional non-invasive tools. The proposed model may expand the clinical role of radiological methods from clinically significant portal hypertension to multiple HVPG categories.

Yu et al., 2022, Cell Reports Medicine *3*, 100563
March 15, 2022
https://doi.org/10.1016/j.xcrm.2022.100563

CellPress

Cell Reports Medicine

CellPress
OPEN ACCESS

Article

An imaging-based artificial intelligence model for non-invasive grading of hepatic venous pressure gradient in cirrhotic portal hypertension

Qian Yu,[1,16] Yifei Huang,[2,16] Xiaoguo Li,[2,16] Michael Pavlides,[3,16] Dengxiang Liu,[4] Hongwu Luo,[5] Huiguo Ding,[6] Weimin An,[7] Fuquan Liu,[8] Changzeng Zuo,[4] Chunqiang Lu,[1] Tianyu Tang,[1] Yuancheng Wang,[1] Shan Huang,[1] Chuan Liu,[2] Tianlei Zheng,[2] Ning Kang,[2] Changchun Liu,[7] Jitao Wang,[4] Seray Akçalar,[9] Emrecan Çelebioğlu,[9] Evren Üstüner,[9] Sadık Bilgiç,[9] Qu Fang,[10] Chi-Cheng Fu,[10] Ruiping Zhang,[11] Chengyan Wang,[12] Jingwei Wei,[13,14] Jie Tian,[13,14] Necati Örmeci,[15] Zeynep Ellik,[15] Özgün Ömer Asiller,[15] Shenghong Ju,[1,17,*] and Xiaolong Qi[2,*]

[1]Department of Radiology, Zhongda Hospital, School of Medicine, Southeast University, Nanjing, China
[2]CHESS Center, Institute of Portal Hypertension, First Hospital of Lanzhou University, Lanzhou, China
[3]Radcliffe Department of Medicine, Oxford Centre for Magnetic Resonance Research, John Radcliffe Hospital, University of Oxford, Oxford, UK
[4]CHESS Working Party, Xingtai People's Hospital, Xingtai, China
[5]Department of General Surgery, Third Xiangya Hospital of Central South University, Changsha, China
[6]Department of Gastroenterology and Hepatology, Beijing You'an Hospital, Capital Medical University, Beijing, China
[7]Department of Radiology, Fifth Medical Center of PLA General Hospital, Beijing, China
[8]Department of Interventional Therapy, Beijing Shijitan Hospital, Capital Medical University, Beijing, China
[9]Department of Radiology, Ankara University School of Medicine, Ankara, Turkey
[10]Shanghai Aitrox Technology Corporation, Shanghai, China
[11]Department of Radiology, Shanxi Bethune Hospital, Third Hospital of Shanxi Medical University, Shanxi, China
[12]Human Phenome Institute, Fudan University, Shanghai, China
[13]Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, China
[14]Beijing Key Laboratory of Molecular Imaging, Beijing, China
[15]Department of Gastroenterology, Ankara University School of Medicine, Ankara, Turkey
[16]These authors contributed equally
[17]Lead contact
*Correspondence: jsh@seu.edu.cn (S.J.), qixiaolong@vip.163.com (X.Q.)
https://doi.org/10.1016/j.xcrm.2022.100563

SUMMARY

The hepatic venous pressure gradient (HVPG) is the gold standard for cirrhotic portal hypertension (PHT), but it is invasive and specialized. Alternative non-invasive techniques are needed to assess the hepatic venous pressure gradient (HVPG). Here, we develop an auto-machine-learning CT radiomics HVPG quantitative model (*a*HVPG), and then we validate the model in internal and external test datasets by the area under the receiver operating characteristic curves (AUCs) for HVPG stages ($\geq$10, $\geq$12, $\geq$16, and $\geq$20 mm Hg) and compare the model with imaging- and serum-based tools. The final *a*HVPG model achieves AUCs over 0.80 and outperforms other non-invasive tools for assessing HVPG. The model shows performance improvement in identifying the severity of PHT, which may help non-invasive HVPG primary prophylaxis when transjugular HVPG measurements are not available.

INTRODUCTION

Portal hypertension (PHT), the most prominent non-neoplastic complication of liver cirrhosis, contributes to severe morbidity and mortality.[1] Hepatic venous pressure gradient (HVPG) measurement is the gold standard for the diagnosis of PHT and an important predictor of cirrhotic complications.[2,3] HVPG $\geq$10 mm Hg (clinically significant PHT, CSPH) is the most significant cutoff value in primary prophylaxis, indicating an increased risk of decompensated events in patients with compensated cirrhosis, while HVPG $\geq$12 mm Hg is a high-risk factor for developing variceal bleeding and HVPG $\geq$16 mm Hg suggests an increased risk of death. A higher HVPG $\geq$20 mm Hg (high-risk PHT, high-risk portal hypertension [HRPH]), is associated with bleeding control failure, rebleeding, and mortality.[2–5]

Therefore, continuous HVPG monitoring plays a significant role in the primary prophylaxis and therapeutic management of patients with PHT. However, transjugular HVPG measurement is invasive, expensive, and has professional barriers, which limits its clinical use in PHT management.[3,6] Thus, alternate non-invasive techniques are desperately needed to better assess and monitor HVPG.

Radiological assessments of the liver and spleen, combined with radiomics and deep learning (DL) technology, have

exhibited importance in CSPH diagnosis, according to our previous studies.[7,8] These studies used only two-dimensional (2D) computed tomography (CT) and magnetic resonance (MR) imaging information of the liver and spleen targeting on CSPH, while studies on liver fibrosis have demonstrated that the radiomics information for fibrosis staging from the whole liver are more representative than part of the liver.[9,10] Thus, we hypothesized that the CT information from the whole liver and spleen could be mined to quantify HVPG in patients with PHT using DL and radiomics methods.

To develop an automated HVPG quantitative estimation framework, we tried to develop (1) a DL segmentation network of the liver and spleen on contrast-enhanced CT in PHT patients and (2) an auto-machine learning (AutoML) CT radiomics HVPG quantitative model (called *a*HVPG) for HVPG quantitation and multistage assessment with HVPG $\geq$10, $\geq$12, $\geq$16, and $\geq$20 mm Hg.

RESULTS

Study design and participant characteristics

From 2016 to 2018, 429 consecutive patients were enrolled in CHESS1701 and CHESS1802 (ClinicalTrials.gov NCT03138915 and NCT03766880). Transjugular HVPG measurement was performed in all of the enrolled patients. A flowchart of patient enrollment is shown in Figure 1. Finally, 372 patients with CT data contributed HVPG data to our study. To develop *a*HVPG, 224 (60%) patients were included in the training dataset, and 148 (40%) patients were included in the internal testing dataset.

According to the sample size calculation (Table S1), we had enough patients in the training and testing datasets to validate the model in each HVPG stage. Baseline characteristics of training and internal test cohorts are summarized in Table 1, and the baseline characteristics of external test cohort are summarized in Table S2.

A DL model for liver and spleen segmentation in CT

To automate the analysis of *a*HVPG and reduce the selection bias due to the handcrafted volume, we developed a DL network to segment the 3D liver and spleen volume in portal-venous phase CT from PHT patients. Based on two 3D fully convolutional networks (FCNs, based on the V-Net architecture),[11] the organ segmentation network was divided into two stages (Figure 2A). In stage one, the input images were downsampled and fed into the 3D FCN subnetwork (based on the V-Net architecture, Figure 2A) to obtain the low-resolution segmentation map. In stage two, the low-resolution feature map was upsampled to the original resolution, concatenated with the inputs, and fed into the higher resolution 3D FCNs to obtain the final segmentation results.

Manual segmentation of the liver and spleen in the portal phase by radiologists was treated as the ground truth. The main vessels around the porta hepatis and splenic hilum were excluded. The gallbladder was not delineated. The segmentation results were evaluated using the Dice metric (DM), the Jaccard coefficient, and positive predictive values (PPVs) in the test dataset.

The organ segmentation network accurately outlined the volumes of the liver and spleen in two independent test cohorts (Figure 3A), with average DMs of 0.973 (SD 0.015) and 0.978 (0.009), Jaccard coefficients of 0.948 (0.028) and 0.959 (0.019), and PPVs of 0.961 (0.024) and 0.960 (0.018) for liver volumetric segmentation; and the corresponding values for the spleen were average DMs of 0.974 (0.014) and 0.983 (0.015), Jaccard coefficients of 0.950 (0.026) and 0.966 (0.028), and PPVs of 0.962 (0.021) and 0.975 (0.021).

Development and overall diagnostic performance of *a*HVPG

The workflow of aHVPG development is presented in Figure 2B. Transjugular HVPG measurements were used as the ground truth and the model output the quantitative results.

Portal-venous phase CT images were used for radiomics analysis for their better performance in CSPH diagnosis in previous studies.[7] CT images and masks obtained from the DL network were collated (S.H. and Y.W., board-certified radiologists) and sent to radiomic feature extraction in Pyradiomics.[12] Three feature groups were computed from the normalized and standardized CT images: 14 shape features, 252 first-order features, and 952 textural features. In total, 2,436 features (1,218 features each for the liver and spleen) were extracted from patients.

AutoML method was used for the *a*HVPG development. The tree-based pipeline optimization tool (TPOT)[13] was applied to train a supervised regression model, which could automatically optimize ML pipelines by using genetic programming (including feature preprocessing, feature selection, model selection, and hyperparameter tuning).[13] We used the following parameters to develop the model: 300 generations, a population size of 50, and 10-fold cross-validation. TPOT output the best-performing model and the quantitative results. We selected the best regression model with a Spearman's rho of 0.832 (95% confidence interval [CI] 0.772–0.877, $p < 0.001$, r^2 0.735] on the training dataset (Figure 3B). The top 10 features in the final model included 3 spleen textural features, 6 liver textural features, and 1 liver first-order feature. The importance of the top 10 features added up to 30.2%. The best model pipeline and selected feature importance are shown in Figure S1.

In the internal test dataset, the *a*HVPG results showed a correlation with the ground truth (Spearman's rho: 0.616, 95% CI 0.504–0.711, $p < 0.001$, r^2 0.407; Figure 3B), outperforming the newly developed tools in Qi et al. (2019) (Spearman's rho = 0.605) and Simbrunner et al. (2020) (Spearman's rho = 0.443).[14,15] The diagnostic performance in each HVPG stage is shown in Figure 3C. In the test dataset, the area (AUC) under the receiver operating characteristic curve (ROC) for CSPH diagnosis (0.833, 95% CI 0.76–0.90) was the highest among all of the HVPG stages, followed by the AUC for HRPH (0.814, 95% CI 0.74–0.88). The AUC, sensitivity, specificity, PPVs, negative predictive values (NPVs), and positivity HVPG missed/all positivity cases, and F2 score are summarized in Table 2. The details of the AUCs of the test set in centers C–F, with small samples ranging from 0.71 to 1.00 for HVPG stratification for their different HVPG distributions, are shown in Table S3 and Figure S2.

Diagnostic performance comparison with conventional models

In each HVPG stage, we compared the diagnostic power of *a*HVPG with the conventional imaging-based and serum-based

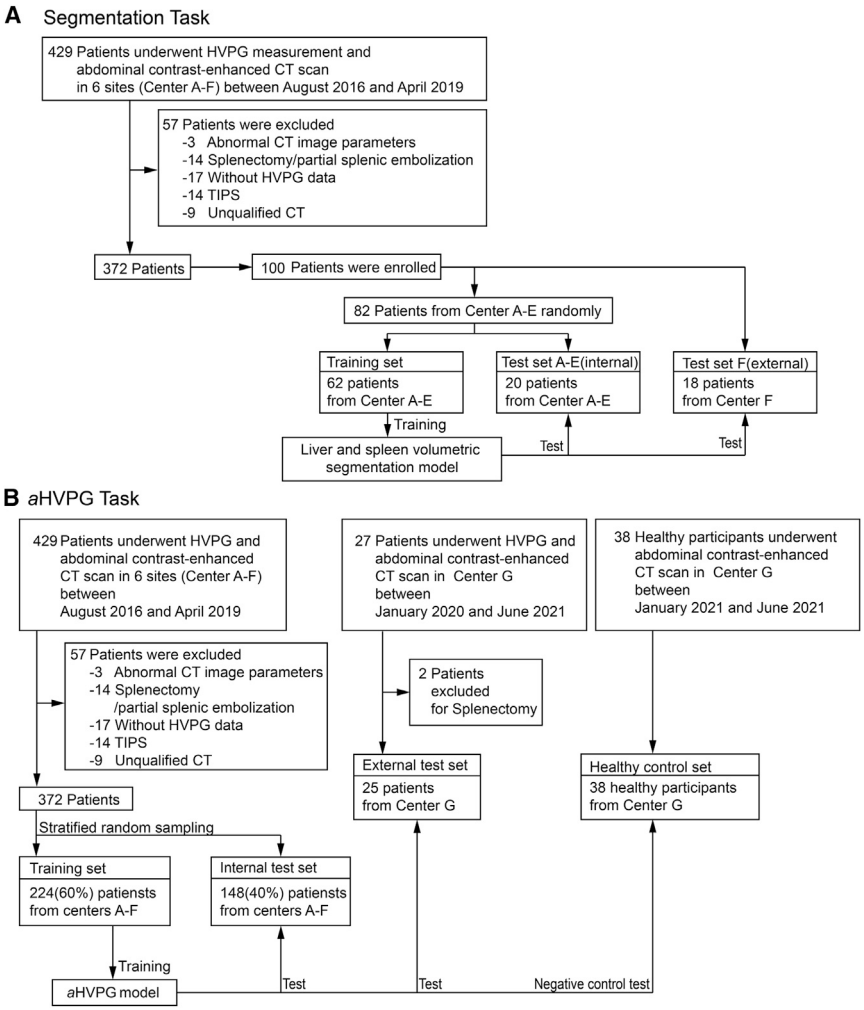


Figure 1. Flowchart of study enrollment

(A) The segmentation task included 100 patients from center A to center F; 62 of 100 patients were enrolled in training the segmentation network, and 38 patients were used for the test.

(B) The *a*HVPG (radiomics) task included 372 patients from center A to center F for model training and internal testing and 27 from center G for external testing. Segmentation and radiomics tasks are independent. Another 38 healthy participants were enrolled in the healthy control dataset. HVPG, hepatic venous pressure gradient.

models, including liver stiffness, CT-based portal pressure score ($HVPG_{CT}$ score),[16] CSPH risk score,[17] King's score,[18] Lok score,[19] aspartate transaminase (AST) to platelet ratio index (APRI),[20] Fibrosis-4 (FIB-4) Index,[21] and AST/alanine transaminase (ALT) ratio (AAR).[22] The calculation methods are shown in Table S4.

In training and testing, *a*HVPG outperformed the $HVPG_{CT}$ score and serum-based models in each HVPG stage (DeLong test, $p < 0.05$). Figure 4 shows the ROC curves of *a*HVPG as well as the top three AUCs within the $HVPG_{CT}$ score and serum-based models (the number of patients and all of the AUCs are shown in Table S5).

Liver stiffness was measured in 84 of 372 enrolled patients. Due to the limited number of patients with liver stiffness, *a*HVPG showed a moderate diagnostic power similar to liver stiffness (DeLong test, $p > 0.05$; Figure S3; Table S6). *a*HVPG achieved better performance for the diagnosis of HVPG $\geq$16 and $\geq$20 mm Hg than liver stiffness in the test dataset (AUC 0.827 versus 0.727, 0.858 versus 0.563), but without significant difference (DeLong test, $p > 0.05$).

For CSPH, according to Baveno VI criteria (liver stiffness $\geq$20 kPa), 29 of 84 patients were identified as CSPH accurately and 37 of 66 CSPH patients were missed. *a*HVPG identified 58 of 84 CSPH patients accurately and missed 8 of 66. All of the patients with liver stiffness $\geq$20 kPa had platelets $<150 \times 10^9$/L.

We included *a*HVPG, serum albumin (ALB), international normalized ratio (INR), AST, ALT, and platelet count (PLT) into a multiple linear regression using the patients with shared indexes in the training dataset. Although the model was statistically significant (F-statistic 87.45), none of the serum markers were associated with HVPG ($p > 0.05$; Table S7).

External test of *a*HVPG

In the external test dataset, the *a*HVPG results showed a correlation with the ground truth (Spearman's rho 0.751, 95% CI 0.434–0.915, $p < 0.001$; Figure S4), and the overall accuracy of classification was 89%. No CSPH was missed. The F2 score for $\geq$10, $\geq$12, $\geq$16, and $\geq$20 mm Hg was 100%, 91.35%, 96.15%, and 64.52%, respectively, which are consistent with training and internal test sets.

Diagnostic robustness and negative control test

For the robustness assessment of *a*HVPG, three different training and internal testing datasets were randomly constructed from all of the patients at a ratio of 6:4 to retrain the model, which were tested and compared with the original model. After retraining and testing, there was no significant variation in the AUCs after different training datasets were applied in all of the HVPG stages (DeLong test, $p > 0.05$; Figure 5; Table S8).

To further investigate whether the model may identify a normal CT scan as CSPH, 38 healthy participants were enrolled as the healthy control dataset, and the accuracy of classification as a non-CSPH patient is 84% (2 of 38). The model may not identify a normal CT scan as CSPH patients.

DISCUSSION

In this post hoc study, we proposed a fully automated HVPG quantitative estimation framework based on CT, including a DL

Table 1. Baseline characteristics of the patients in training and internal test dataset

	Median (interquartile ranges), or n/total (%)			
Characteristics	All (N = 372)	Training set (n = 224)	Internal test set (n = 148)	*P*
Age, y	50 (41–57.00)	50 (41–56)	49 (41–57)	0.850
Sex				0.052
Female	118/372 (32)	62/224 (28)	56/148 (38)	
Male	254/372 (68)	162/224 (72)	92/148 (62)	
BMI, kg/m^2	23.39 (21.10–25.60)	23.52 (21.27–25.78)	23.14 (20.72–25.35)	0.301
Child-Pugh class				0.748
A	212/325 (65)	129/194 (66)	83/131 (63)	
B	85/325 (26)	50/194 (26)	35/131 (27)	
C	28/325 (9)	15/194 (8)	13/131 (10)	
Etiology				0.402
Hepatitis B	205/372 (55)	126/224 (56)	79/148 (53)	
Hepatitis C	17/372 (5)	13/224 (6)	4/148 (3)	
Alcoholic liver disease	28/372 (8)	15/224 (7)	13/148 (9)	
Other[a]	122/372 (33)	70/224 (31)	52/148 (35)	
Ascites				0.055
No	219/342 (64)	142/208 (68)	77/134 (57)	
Yes	123/342 (36)	66/208 (32)	57/134 (43)	
HVPG, mm Hg	16.54 (11.95–21.01)	16.54 (11.95–21.08)	16.54 (11.94–21.01)	0.898
HVPG stage, mm Hg				
<10	59/372 (16)	35/224 (16)	24/148 (16)	0.886
≥10	313/372 (84)	189/224 (84)	124/148 (84)	0.886
≥12	278/372 (75)	167/224 (75)	111/148 (75)	>0.99
≥16	209/372 (56)	126/224 (56)	83/148 (56)	>0.99
≥20	115/372 (31)	70/224 (31)	45/148 (30)	0.954
TBIL, μmol/L	18.50 (12.85–25.50)	18.55 (12.62–25.92)	18.50 (13.60–24.60)	>0.99
ALB, g/L	35.41 (32.05–39.00)	35.00 (32.00–38.10)	36.00 (32.38–39.48)	0.297
INR, U/L	1.17 (1.08–1.31)	1.19 (1.09–1.33)	1.15 (1.08–1.28)	0.193
AST, U/L	31.00 (23.00–42.00)	31.00 (22.00–42.00)	32.00 (23.50–42.50)	0.282
ALT, U/L	22.00 (16.00–31.00)	21.00 (15.00–31.75)	23.00 (17.00–31.00)	0.131
Platelets, $10^9/L$	65.00 (46.00–93.00)	64.00 (43.25–91.75)	67.00 (48.50–94.00)	0.433
Liver stiffness, kPa	17.05 (12.38–27.70)	17.10 (12.60–27.25)	16.50 (12.00–27.70)	0.907
$HVPG_{CT}$ score	17.37 (15.74–19.86)	17.37 (15.55–19.41)	17.75 (16.34–20.65)	0.025
AAR	1.41 (1.15–1.73)	1.45 (1.16–1.76)	1.38 (1.10–1.71)	0.211
APRI	1.22 (0.74–1.85)	1.22 (0.71–1.88)	1.22 (0.76–1.84)	0.953
CSPH risk score	6.02 (3.54–8.84)	6.27 (4.12–9.10)	5.41 (3.05–8.35)	0.082
FIB-4	4.95 (3.02–7.56)	5.09 (3.06–7.75)	4.93 (3.02–7.19)	0.493
King's score	28.47 (16.37–46.41)	28.20 (15.86–47.38)	28.83 (18.36–43.54)	0.849
Lok score	1.87 (1.17–2.76)	1.92 (1.27–2.82)	1.68 (1.12–2.57)	0.130
Center				0.350
A	237/372 (64)	147/224 (66)	90/148 (61)	
B	66/372 (18)	41/224 (18)	25/148 (17)	
C	18/372 (5)	12/224 (5)	6/148 (4)	
D	17/372 (5)	8/224 (4)	9/148 (6)	
E	16/372 (4)	6/224 (3)	10/148 (7)	
F	18/372 (5)	10/224 (4)	8/148 (5)	

AAR, AST to ALT ratio; ALB: albumin; ALT, alanine transaminase; APRI, AST to platelet ratio index; AST, aspartate transaminase; BMI, body mass index; CSPH, clinically significant portal hypertension; HVPG, hepatic venous pressure gradient; TBIL, total bilirubin.

[a]Other etiologies included hepatic sinusoidal obstruction syndrome, autoimmune liver disease, primary biliary cirrhosis, non-alcoholic steatohepatitis (NASH), and unknown.

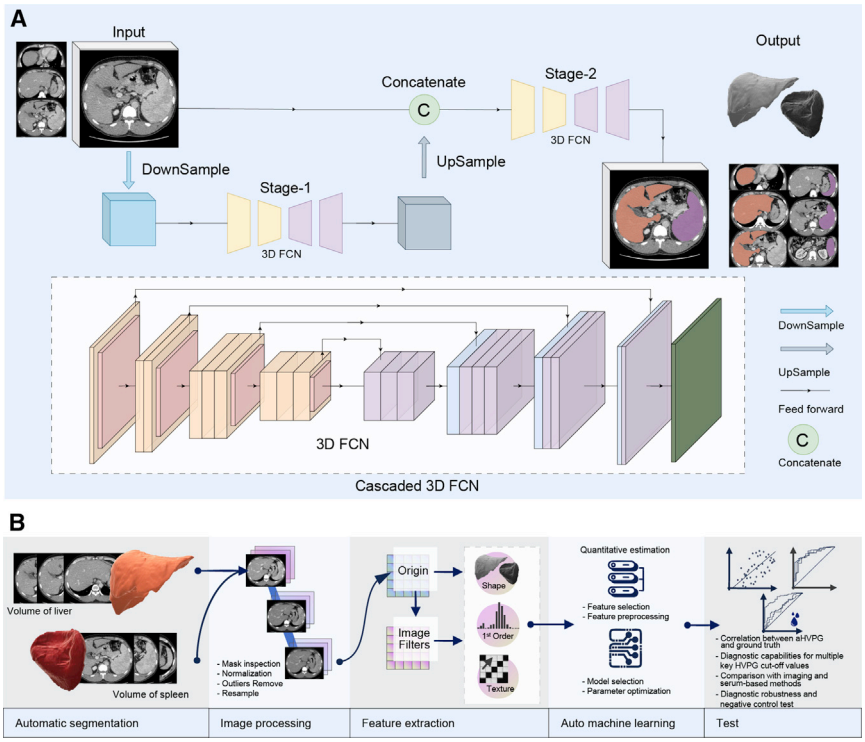


Figure 2. Deep learning segmentation framework and radiomics development workflow

(A) The organ segmentation framework included two 3D fully convolutional networks (3D FCNs). In stage 1, the first 3D FCN generated the low-resolution segmentation map. In stage 2, the low-resolution feature map was upsampled to the original resolution, concatenated with the inputs, and fed into the higher resolution 3D FCNs to obtain the final segmentation results.

(B) Radiomics analysis workflow of *a*HVPG. CT images and masks obtained from the deep learning network were pre-processed, and 1,218 features each for the liver and spleen were extracted. The tree-based pipeline optimization tool was applied to train a supervised regression model, and the hepatic venous pressure gradient was used as the ground truth. Finally, the model output the quantitative results and would be validated in internal and external test cohorts.

organ volumetric segmentation model and *a*HVPG, which is an AutoML radiomics model. The segmentation model exhibited excellent performance in liver and spleen segmentation. The proposed model showed diagnostic power better than the traditional imaging- and serum-based models for assessing HVPG stages.

Compared with our previous studies on CSPH diagnosis,[7,8,15] we can rationalize the radiological workflow in HVPG assessment using DL and Auto-ML methods. This study realized an automated model for HVPG estimation and multistage assessment, which expanded the clinical role of radiological methods from CSPH to multi-HVPG severity, while previous studies need manual intervention and only focused on CSPH.

The performance of *a*HVPG for assessing HVPG stages may be useful to improve risk stratification and clinical decision making in patients with cirrhosis and PHT. In the test dataset, the *a*HVPG results showed a moderate correlation (Spearman's rho = 0.616) with the ground truth, outperforming the newly developed tools reported in Qi et al. (2019) and Simbrunner et al. (2020).[14,15] Also, we found the diagnostic performance of *a*HVPG considerable for detecting different severities of PHT. This may help clinicians preliminarily assess the risk of complications (e.g., esophageal varices bleeding) or identify patients at high risk needing proactive treatment measures (e.g., transjugular intrahepatic portosystemic shunt [TIPS]) when an invasive procedure is unavailable. Finally, no significant variation was observed after re-training and testing, suggesting the robustness and reproducibility of our model.

Almost existing non-invasive imaging methods, including our study, tend to seek external reflections of profound structural alteration in the liver and spleen. As chronic liver disease progresses, damaged liver parenchymal and non-parenchymal cells cause structural changes to the liver tissue, including fibrosis, regenerative nodules, and destruction of vascular structures, leading to significant liver structure changes, intrahepatic vascular resistance increase, and then PHT. At the same time, another organ of the portal system, the spleen, is also involved in the vicious cycle of PHT. This is based not only on passive congestion of the spleen but also on structural changes resulting from angiogenesis, fibrogenesis, and hyperactivation of the splenic lymphatic compartment.[1,23] Therefore, significant structural changes in the liver and spleen may be revealed by imaging methods, and conventional CT methods should not be limited to morphological changes only, but deep into areas beyond the reach of the human eye.

Radiomics provides us with the means to extract high-throughput CT data and uncover possible tissue cell-level alterations hidden in pixels.[24] In our study, the main features included in *a*HVPG were second-order (9 of the top 10 features selected). These features were generated from the interrelationship between neighboring voxels and were insensitive to the absolute gray value. Furthermore, these features buried information about the coarseness of the texture and the spatial heterogeneity of the liver and spleen, which may be an imaging reflection of changes in the structure, and the pathogenetic mechanisms hidden within the image textures may be related to liver and spleen stiffness. Liver stiffness has been considered an alternative diagnostic method for CSPH in Baveno VI[2,25] and showed diagnostic efficacy for severe PHT in our study. The spleen stiffness measurement has also been validated in ruling out high-risk varices in PHT alone or when combined with Baveno VI criteria.[23,26] However, liver stiffness performed worse than *a*HVPG in diagnosing higher HVPG in our study, which may be attributed to extrahepatic factors in cirrhosis progression.[1]

The $HVPG_{CT}$ score, which is calculated by the liver and spleen volume and ascites, was the third most accurate model in the test

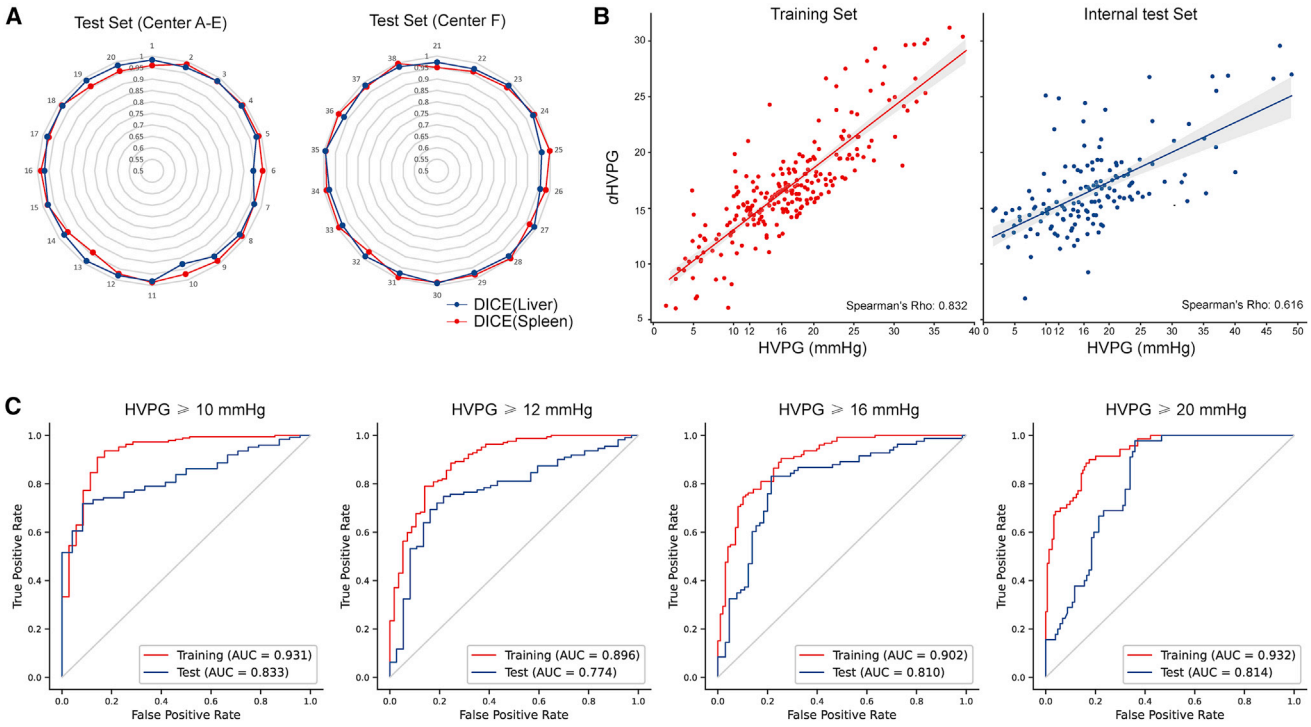


Figure 3. Segmentation accuracy and diagnostic performance of the deep learning network and *a*HVPG

(A) Dice metric, Jaccard coefficient of the deep learning segmentation network for the liver and spleen in the internal test dataset (centers A–E) and external test dataset (center F).

(B) Correlation between *a*HVPG and invasive HVPG. Scatterplot shows agreement between the *a*HVPG and the invasive HVPG in training and internal test datasets.

(C) Receiver operating characteristic curves of the *a*HVPG for assessing hepatic venous pressure gradient stages, including ≥10, ≥12, ≥16, and ≥20 mm Hg in training (red line) and internal test sets (blue line).

AUC, area under the curve.

dataset, while the shape features showed less importance in *a*HVPG. These results indicated that features involving the liver and spleen volume may be less important in assessing HVPG in imaging studies,[16,27,28] which is consistent with a previous study.[7]

Furthermore, serum biomarkers showed poor performance in HVPG assessment, while they were originally designed for liver fibrosis or cirrhosis[18,20,21,29] and are consistent with our previous studies.[7,8] Simbrunner et al. (2020) proposed the enhanced liver fibrosis (ELF) score in CSPH and HRPH assessment, with AUCs of 0.833 and 0.677, respectively, which presented a strong CSPH diagnostic ability.[14] Un-like CT data, liver stiffness and the ELF score have not been included in the PHT management

Table 2. Diagnostic accuracy of *a*HVPG for each HVPG stage

HVPG	Group	AUC[a]	95% CI	Cutoff value	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Missed (%)[b]	F2 score (%)
≥10 mm Hg	training	0.93	0.88–0.98	12.7	95.24	77.14	95.74	75	4.76	95.34
	internal test	0.83	0.76–0.90		95.16	20.83	86.13	45.45	4.84	93.21
≥12 mm Hg	training	0.90	0.85–0.94	13.7	94.61	63.16	88.27	80	5.39	93.27
	internal test	0.77	0.68–0.85		87.39	32.43	79.51	46.15	12.61	85.69
≥16 mm Hg	training	0.90	0.86–0.94	14.7	95.24	59.18	75	90.62	4.76	90.36
	internal test	0.81	0.73–0.88		89.16	50.77	69.81	78.57	10.84	84.47
≥20 mm Hg	training	0.93	0.90–0.96	16	95.71	62.99	54.03	97	4.29	82.92
	internal test	0.81	0.74–0.88		88.89	66.02	53.33	93.15	11.11	78.43

AUC, area under the receiver operating characteristic curve; CI, confidence interval; HVPG, hepatic venous pressure gradient; NPV, negative predictive value; PPV, positive predictive value.

[a]$p < 0.001$.

[b]Positivity HVPG missed/all positivity cases.

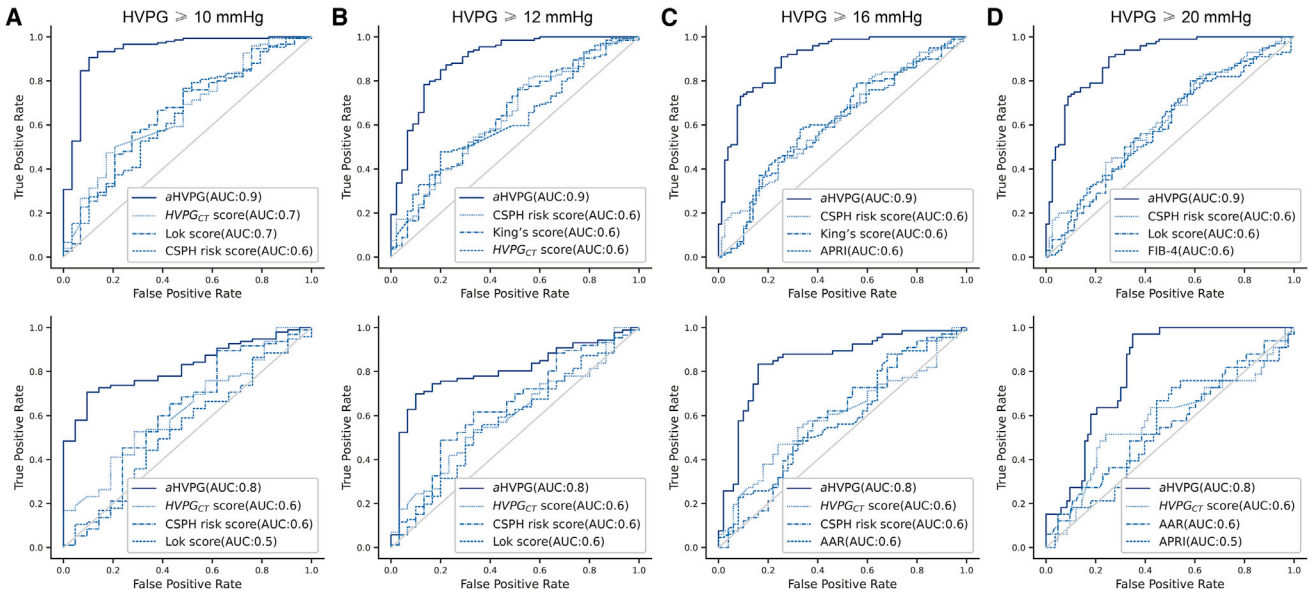


Figure 4. Receiver operating characteristic curves of *a*HVPG and the top 3 conventional non-invasive tools

(A–D) Top: Receiver operating characteristic curves (ROCs) of *a*HVPG and the top three conventional non-invasive tools in the training dataset for assessing hepatic venous pressure gradient (HVPG) stages including ≥10, ≥12, ≥16, and ≥20 mm Hg. Bottom: ROCs in the internal test dataset for assessing HVPG stages.

AAR, AST to ALT ratio; APRI, AST to platelet ratio index; CSPH, clinically significant portal hypertension.

workflow in most hospitals, but it is possible to provide a non-invasive HVPG staging method by combining CT scan, liver or spleen stiffness, and the ELF score.

The AutoML method is also an exciting tool for bioinformatics problems. Balancing complexity and interpretability is a timeless topic in clinical model building, and AutoML methods may provide a possible solution to designing a more complex pipeline, yielding satisfactory outcomes not inferior to those made by humans, as well as more interpretable results compared with some DL methods.[8,13]

In conclusion, we developed an automated and non-invasive HVPG quantitative estimation method to evaluate HVPG stages based on CT. Owing to the convenience of CT examinations, *a*HVPG, as a non-invasive method, may help non-invasive HVPG primary prophylaxis when transjugular HVPG measurements are not available.

Limitations of the study

For the multicenter situation, we used a mixed-samples strategy to make the model conform to real-world distributions. A trained model that uses images with multiple parameters or from different scanners may improve its generalizability for manufacturers.

Selection bias is the most significant limitation of this study. Because of the concerns of patients with mild symptoms about the procedure and the price of HVPG measurement, the limited sample size with a lower HVPG level (i.e., HVPG <5 mm Hg) caused an imbalanced HVPG distribution. According to the sample size calculation, we had enough patients in the training and testing datasets to validate the model performance, especially in distinguishing CSPH. The negative control test also demonstrated enough power to identify healthy people or non-portal hypertensive patients by aHVPG, but such imbalanced HVPG distribution has resulted in overestimates. Given the lack of contrast between the patients with HVPG <5 mm Hg and >5 mm Hg, we could not obtain a threshold value for the diagnosis of PHT. Patients with non-portal hypertensive cirrhosis need to be included in the future to update the model. In addition, the follow-up data of patients in this study were not collected. Finally, ionizing radiation from CT requires attention, although CT can provide a rapid examination process and information added about the whole abdomen.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Study design and participants
  - CT examinations
  - Transjugular HVPG measurement
  - Liver stiffness and clinical features
  - Deep learning network for liver and spleen segmentation
  - Performance analysis of deep learning network

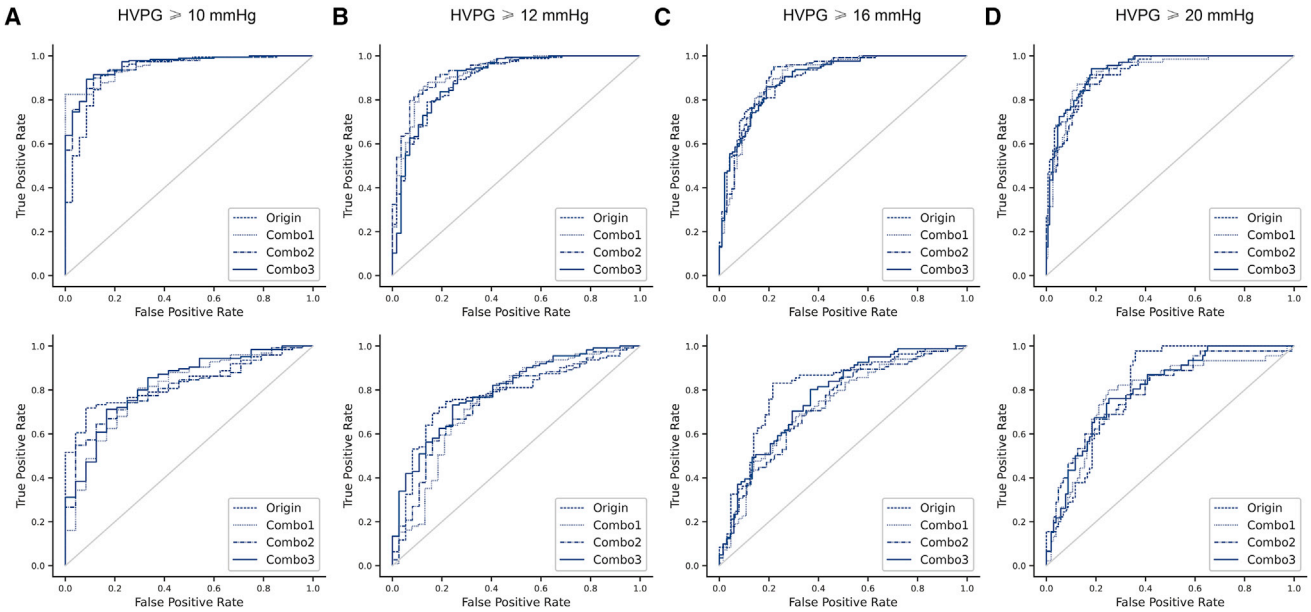


Figure 5. Receiver operating characteristic curves of the robustness test

(A–D) Top: Receiver operating characteristic curves (ROCs) of the original model and 3 times-retrained models for assessing HVPG stages, including ≥10, ≥12, ≥16, and ≥20 mm Hg in the training dataset (DeLong test, $p > 0.05$). Bottom: ROCs of the robustness test in the test dataset for assessing HVPG stages (DeLong test, $p > 0.05$).

- ○ Diagnostic accuracy assessment of *a*HVPG
- ● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrm.2022.100563.

ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China (NSFC, grant nos. 81830053 and 61821002).

AUTHOR CONTRIBUTIONS

S.J., X.Q., Q.Y., and M.P. contributed to the conception and design of the study. Y.H. and X.L. led the patient recruitment. Y.H., X.L., D.L., H.L., H.D., W.A., F.L., C.Z., C. Liu, T.Z., N.K., C.L., J.W., S.A., E.Ç., E.Ü., S.B., N.Ö., Z.E., and Ö.Ö.A. contributed to the acquisition of the data. Q.Y., C. Lu, T.T., Y.W., and S.H. contributed to the analysis or interpretation of the data. Development of the method was performed by Q.Y. and supervised by Q.F., C.-C.F., R.Z., C.W., J.W., and J.T. Drafting of the manuscript was done by Q.Y. and revised by S.J., X.Q., C.Lu, Y.H., X.L., M.P., Q.F., C.-C.F., R.Z., C.W., J.W., and J.T. All of the authors contributed to the critical review and final approval of the manuscript. The corresponding author attests that all of the listed authors met the authorship criteria and that no others meeting the criteria have been omitted.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 10, 2021
Revised: December 19, 2021
Accepted: February 17, 2022
Published: March 15, 2022

REFERENCES

1. Gracia-Sancho, J., Marrone, G., and Fernández-Iglesias, A. (2019). Hepatic microcirculation and mechanisms of portal hypertension. Nat. Rev. Gastroenterol. Hepatol. *16*, 221–234.
2. Franchis, R. de (2015). Expanding consensus in portal hypertension: report of the Baveno VI Consensus Workshop: stratifying risk and individualizing care for portal hypertension. J. Hepatol. *63*, 743–752.
3. Qi, X., Berzigotti, A., Cardenas, A., and Sarin, S.K. (2018). Emerging noninvasive approaches for diagnosis and monitoring of portal hypertension. Lancet Gastroenterol. Hepatol. *3*, 708–719.
4. D'Amico, G., Garcia-Pagan, J.C., Luca, A., and Bosch, J. (2006). Hepatic vein pressure gradient reduction and prevention of variceal bleeding in cirrhosis: a systematic review. Gastroenterology *131*, 1611–1624.
5. Veldhuijzen van Zanten, D., Buganza, E., and Abraldes, J.G. (2021). The role of hepatic venous pressure gradient in the management of cirrhosis. Clin. Liver Dis. *25*, 327–343.
6. Garcia-Tsao, G., Abraldes, J.G., Berzigotti, A., and Bosch, J. (2017). Portal hypertensive bleeding in cirrhosis: risk stratification, diagnosis, and management: 2016 practice guidance by the American Association for the study of liver diseases. Hepatology *65*, 310–335.
7. Liu, F., Ning, Z., Liu, Y., Liu, D., Tian, J., Luo, H., An, W., Huang, Y., Zou, J., Liu, C., et al. (2018). Development and validation of a radiomics signature for clinically significant portal hypertension in cirrhosis (CHESS1701): a prospective multicenter study. EBioMedicine *36*, 151–158.
8. Liu, Y., Ning, Z., Örmeci, N., An, W., Yu, Q., Han, K., Huang, Y., Liu, D., Liu, F., Li, Z., et al. (2020). Deep convolutional neural network-aided detection of portal hypertension in patients with cirrhosis. Clin. Gastroenterol. Hepatol. *18*, 2998–3007.e5.
9. Yasaka, K., Akai, H., Kunimatsu, A., Abe, O., and Kiryu, S. (2017). Liver fibrosis: deep convolutional neural network for staging by using gadoxetic acid–enhanced hepatobiliary phase MR images. Radiology *287*, 146–155.
10. Wang, K., Lu, X., Zhou, H., Gao, Y., Zheng, J., Tong, M., Wu, C., Liu, C., Huang, L., Jiang, T., et al. (2019). Deep learning Radiomics of shear

wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study. Gut *68*, 729–741.

11. Milletari, F., Navab, N., and Ahmadi, S. (2016). V-net: fully convolutional neural networks for volumetric medical image segmentation. In 2016 Fourth International Conference on 3D Vision (3DV) (IEEE), pp. 565–571.
12. van Griethuysen, J.J.M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R.G.H., Fillion-Robin, J.-C., Pieper, S., and Aerts, H.J.W.L. (2017). Computational radiomics system to decode the radiographic phenotype. Cancer Res. *77*, e104–e107.
13. Le, T.T., Fu, W., and Moore, J.H. (2020). Scaling tree-based automated machine learning to biomedical big data with a feature set selector. Bioinformatics *36*, 250–256.
14. Simbrunner, B., Marculescu, R., Scheiner, B., Schwabl, P., Bucsics, T., Stadlmann, A., Bauer, D.J.M., Paternostro, R., Eigenbauer, E., Pinter, M., et al. (2020). Non-invasive detection of portal hypertension by enhanced liver fibrosis score in patients with different aetiologies of advanced chronic liver disease. Liver Int. *40*, 1713–1724.
15. Qi, X., An, W., Liu, F., Qi, R., Wang, L., Liu, Y., Liu, C., Xiang, Y., Hui, J., Liu, Z., et al. (2019). Virtual hepatic venous pressure gradient with CT angiography (CHESS 1601): a prospective multicenter study for the noninvasive diagnosis of portal hypertension. Radiology *290*, 370–377.
16. Iranmanesh, P., Vazquez, O., Terraz, S., Majno, P., Spahr, L., Poncet, A., Morel, P., Mentha, G., and Toso, C. (2014). Accurate computed tomography-based portal pressure assessment in patients with hepatocellular carcinoma. J. Hepatol. *60*, 969–974.
17. Berzigotti, A., Gilabert, R., Abraldes, J.G., Nicolau, C., Bru, C., Bosch, J., and García-Pagan, J.C. (2008). Noninvasive prediction of clinically significant portal hypertension and esophageal varices in patients with compensated liver cirrhosis. Am. Coll. Gastroenterol. *103*, 1159–1167.
18. Cross, T.J.S., Rizzi, P., Berry, P.A., Bruce, M., Portmann, B., and Harrison, P.M. (2009). King's Score: an accurate marker of cirrhosis in chronic hepatitis C. Eur. J. Gastroenterol. Hepatol. *21*, 730–738.
19. Lok, A.S.F., Ghany, M.G., Goodman, Z.D., Wright, E.C., Everson, G.T., Sterling, R.K., Everhart, J.E., Lindsay, K.L., Bonkovsky, H.L., Bisceglie, A.M.D., et al. (2005). Predicting cirrhosis in patients with hepatitis C based on standard laboratory tests: results of the HALT-C cohort. Hepatology *42*, 282–292.
20. Wai, C.-T., Greenson, J.K., Fontana, R.J., Kalbfleisch, J.D., Marrero, J.A., Conjeevaram, H.S., and Lok, A.S.-F. (2003). A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology *38*, 518–526.
21. Vallet-Pichard, A., Mallet, V., Nalpas, B., Verkarre, V., Nalpas, A., Dhalluin-Venier, V., Fontaine, H., and Pol, S. (2007). FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. Hepatology *46*, 32–36.
22. Bonacini, M., Hadi, G., Govindarajan, S., and Lindsay, K.L. (1997). Utility of a discriminant score for diagnosing advanced fibrosis or cirrhosis in patients with chronic hepatitis C virus infection. Am. J. Gastroenterol. *92*, 1302–1304.
23. Colecchia, A., Montrone, L., Scaioli, E., Bacchi–Reggiani, M.L., Colli, A., Casazza, G., Schiumerini, R., Turco, L., Di Biase, A.R., Mazzella, G., et al. (2012). Measurement of spleen stiffness to evaluate portal hypertension and the presence of esophageal varices in patients with HCV-related cirrhosis. Gastroenterology *143*, 646–654.
24. Gillies, R.J., Kinahan, P.E., and Hricak, H. (2015). Radiomics: images are more than pictures, they are data. Radiology *278*, 563–577.
25. Berzigotti, A. (2017). Non-invasive evaluation of portal hypertension using ultrasound elastography. J. Hepatol. *67*, 399–411.
26. Wang, H., Wen, B., Chang, X., Wu, Q., Wen, W., Zhou, F., et al. (2021). Baveno VI criteria and spleen stiffness measurement rule out high-risk varices in virally suppressed HBV-related cirrhosis. J. Hepatol. *74*, 584–592.
27. Qi, X., Liu, F., Li, Z., Chen, S., Liu, Y., Yang, Y., and Hou, J. (2018). Insufficient accuracy of computed tomography-based portal pressure assessment in hepatitis B virus-related cirrhosis: an analysis of data from CHESS-1601 trial. J. Hepatol. *68*, 210–211.
28. Qi, X., Xu, M., Li, Z., and Yang, C. (2014). Virtual portal pressure from anatomic CT angiography. J. Hepatol. *61*, 180–181.
29. Ohta, T., Sakaguchi, K., Fujiwara, A., Fujioka, S., Iwasaki, Y., Makino, Y., Araki, Y., and Shiratori, Y. (2006). Simple surrogate index of the fibrosis stage in chronic hepatitis C patients using platelet count and serum albumin level. Acta Med. Okayama *60*, 77–84.
30. Bossuyt, P.M., Reitsma, J.B., Bruns, D.E., Gatsonis, C.A., Glasziou, P.P., Irwig, L., Lijmer, J.G., Moher, D., Rennie, D., de Vet, H.C.W., et al. (2015). STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ *351*, h5527.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Python 3.7	Python	https://www.python.org
PyTorch 0.4.1	PyTorch	https://pytorch.org/
TPOT	Le et al.[13]	https://github.com/EpistasisLab/tpot
MedCalc v19.0.4	MedCalc	N/A
SPSS 22.0	SPSS	N/A
Spleen and liver segmentation code	This paper	https://github.com/vanziaa/aHVPG
*a*HVPG code and statistics results	This paper	https://github.com/vanziaa/aHVPG
Pyradiomics	Griethuysen JJM et al.[12]	https://github.com/AIM-Harvard/pyradiomics

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Shenghong Ju (jsh@seu.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The confidential medical records data reported in this study cannot be deposited in a public repository. To request access, contact the Lead Author. In addition, statistics results from these data have been deposited at https://github.com/vanziaa/aHVPG and are publicly available as of the date of publication.

All original code has been deposited at https://github.com/vanziaa/aHVPG and is publicly available as of the date of publication.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

This study involved human subjects and detailed inclusion/exclusion criteria was described in method details section, as well as in Figure 1. Demographic findings, including age and sex for all patients enrolled in this study is included in Table 1 and Table S2. The sample size consideration is shown in Table S1. This study retrospectively enrolled patients from six Chinese hospitals (Centre A: The Fifth Medical Center of PLA General Hospital, Centre B: Beijing Shijitan Hospital, Centre C: The Third Xiangya Hospital of Central South University, Centre D: Beijing You'an Hospital, Centre E: Xingtai People's Hospital, Centre G: Southeast university Zhongda hospital) and a Turkish hospital (Centre F: Ankara University). The study was approved by all local institutional review boards (The main centre is the Fifth Medical Center of PLA General Hospital, IRB Number: 2015068D, and this study was registered and approved in other centres following the main centre). Written informed consent was obtained from all HVPG measurement participants. All centres (Centres A-G) are registered as collaborators in clinical trials the Chinese Portal Hypertension Alliance (CHESS) 1701 (ClinicalTrials.gov, NCT03138915) and CHESS1802 (NCT03766880).

METHOD DETAILS

Study design and participants

This study followed the Standards for Reporting of Diagnostic Accuracy (STARD) reporting guidelines.[30]

aHVPG development and internal test datasets

We formed the training and internal test datasets using prospective patients with cirrhosis undergoing clinically-indicated transjugular HVPG measurement and contrast-enhanced abdominal CT in five Chinese hospitals (Centres A-E) and a Turkish hospital (Centre F) from August 2016 to April 2019,[7,8] to develop the deep learning network for liver and spleen segmentation and the *a*HVPG for HVPG estimation in cirrhosis patients.

The inclusion criteria were as follows: (1) Confirmed cirrhosis (diagnostic criteria including 1) Laboratory tests, such as liver malfunction and hepatitis viruses; 2) Imaging methods, such as CT, MR, and US, which showed the signs of cirrhosis like liver nodules, ascites,

or splenomegaly; 3) Physical examination findings, including the history of HBV or other viral hepatitis, jaundice, ascites, or splenomegaly; 4) A proportion of patients underwent biopsy to confirm the diagnosis.); (2) Patients who had abdominal contrast-enhanced CT scan within 14 days before HVPG measurement; (3) Adult patients (age from 18 to 75 years); (4) Written informed consent.

The exclusion criteria were as follows: (1) Patients previously underwent any surgical procedures of liver or spleen (e.g., TIPS, liver transplantation, splenectomy, and partial splenic embolization); (2) Patients with hepatocellular carcinoma; (3) Acute portal hypertension in cases of acute-on-chronic liver failure; (4) Technical reasons (e.g., abnormal CT parameters, artifacts).

The segmentation task enrolled 100 patients based on experience, including 82 patients selected randomly from Center A-E for training and internal test and all 18 patients from Center F for external test).

The radiomics task involved all enrolled patients. By stratified random sampling, 60% of patients in each HVPG level (0-10, 10-12, 12-16, and $\geq$20 mmHg) were randomly selected for training, and the rest for internal testing. There was no data overlap between the training and test datasets, and the segmentation and classification tasks are independent.

External test dataset

We retrospectively enrolled patients with cirrhosis undergoing HVPG measurement and CT from the Centre G using the same criteria from January 2020 to June 2021 for the external test of *a*HVPG.

Healthy control dataset

To further investigate whether the model might identify a normal CT scan as CSPH, we retrospectively enrolled a cohort of healthy participants with abdominal contrast-enhanced CT from routine checkup, in Centre G from January 2021 to June 2021. This cohort would be only used in the negative control test.

CT examinations

CT examinations were performed using standard contrast-enhanced abdominal protocols in each institution (Table S9) within 14 days before the catheterization. All images were extracted from two cohorts and sent to a core laboratory (Southeast University, centre G). The portal-venous phase from the contrast-enhanced abdominal CT were analysed. All the images were reviewed to exclude those that were of abnormal quality by radiologists.

Transjugular HVPG measurement

The details of the transjugular HVPG measurement process were reported in our previous study.[7,8,15] All transjugular HVPG measurements were performed by trained interventional radiologists following the standard operating procedure. We used the balloon catheter with a pressure transducer at the tip. A zero measurement was conducted before the study. The free hepatic venous pressure was measured as the balloon catheter was placed close to the inferior vena cava (approximately 1–3 cm). At the ostium of the right hepatic vein, we measured the free hepatic venous pressure, and then the wedged hepatic venous pressure was measured while the balloon was inflated for total occlusion. Continuous recording was performed until the pressure reached a plateau. The HVPG was the difference between the wedged hepatic venous pressure and the free hepatic venous pressure.

In this study, each HVPG measurement was according to 3 times repeated measurements of WHVP and FHVP, and then the average of at least 2 measurements with difference < 1mmHg was obtained and HVPG was calculated. HVPG measurements were performed in high-volume liver centers, and in each center, one independent interventional radiologist with more than ten-year experience leaded the HVPG measurement and was responsible for quality control.

From 2016, patients with cirrhosis were enrolled in CHESS1701 and 1802 projects and underwent clinically-indicated invasive HVPG measurement to assess the patient's portal pressure, risk of complications or treatment results. In Center G, from 2020, HVPG have been measured in some patients hospitalized with cirrhosis and gastrointestinal bleeding or high-risk esophageal varices and recorded in the medical record system.

Liver stiffness and clinical features

Liver stiffness was measured by FibroScan (Echosens, France) in patients without contraindications. 84 of 372 enrolled patients underwent FibroScan examination.

Clinical and laboratory characteristics on patient admission were collected including ascites, ALB, INR, AST, ALT, and PLT. Liver and spleen volume were calculated based on DL segmentation.

Deep learning network for liver and spleen segmentation

The 3D FCN is based on V-Net architecture. The coding side in the left with four stages performs feature extraction and resolution reduction. Each stage consists of one to three 3×3×3 convolution layers, one 2×2×2 convolution layer with stride 2 to reduce resolution (2-8 times downsampling) and uses a residual structure. The decoding side in the right also with four stages, performs feature fusion, segmentation and output a two-channel volumetric segmentation. Each stage consists of one 2×2×2 deconvolution layer with stride 2 to increase input size (Except for decoding stage 1), three to one 3×3×3 convolution layers, and residual functions. Fine grained features were forwarded from corresponding stages in the left to the right. PReLu non linearities are applied throughout the network. The DICE loss is employed in loss calculation.

Python 3.7 and PyTorch 0.4.1 were used to train and test the models with the following predefined parameters: initial learning rate of 10e-4, 1200 iterations, and a batch size of 1. The layer thickness was adjusted to 3 mm, and the greyscale values were limited to

-350 to 350 for all the images. Random rotation and zooming were performed during training. The liver and spleen segmentation models were trained independently.

The deep learning network was redeveloped. Private patients' CT images of cirrhosis were used to train the model, and we made changes in data preparation, training, and test process to enable the model to perform well in patients with cirrhosis. We also added the inference function to generate segmentation.

Performance analysis of deep learning network

The segmentation results were evaluated using the Dice metric (DM), Jaccard coefficient, and positive predictive values (PPVs) in the test dataset, which were calculated using the following formulas:

1) Dice metric (DM): measured the overlap between ground truth area and predicted area

$$DM(C_m, C_a) = 2 \times \frac{C_m \cap C_a}{C_m + C_a}$$

2) Jaccard coefficient: measured the size of the intersection divided by the size of the union

$$J(C_m, C_a) = \frac{C_m \cap C_a}{C_m \cup C_a} = \frac{C_m \cap C_a}{C_m + C_a - (C_m \cap C_a)}$$

C_m: the region of manual ground truth
C_a: the region of automatic segmentation result

3) Positive predictive values (PPV):

$$PPV_c = \frac{TP_c}{TP_c + FP_c}$$

TP : True Positive of class *c*
FP : False Positive of class *c*

Diagnostic accuracy assessment of *a*HVPG

The overall diagnostic performance was assessed in the training and internal test dataset. First, the correlation between *a*HVPG and transjugular HVPG was evaluated. Then, the diagnostic performance of *a*HVPG was evaluated in each HVPG stage, that is, $\geq$10 mmHg (CSPH), $\geq$12 mmHg, $\geq$16 mmHg, and $\geq$20 mmHg (HRPH). The receiver operating characteristic (ROC) curve, area under the ROC curve (AUC), sensitivity, specificity, PPV, and negative predictive value (NPV), positivity HVPG missed/all positivity cases, and F2-score were used to assess the diagnostic performance of *a*HVPG in each HVPG stage.

QUANTIFICATION AND STATISTICAL ANALYSIS

Continuous variables are presented as medians (interquartile ranges [IQRs]). Based on the data distributions, Student's t-test or Mann–Whitney U tests, as appropriate, were used to compare differences between groups. Categorical variables are presented as counts (%s) and were compared using χ^2 or Fisher's exact test.

Spearman's correlation analysis was performed to assess the correlation between *a*HVPG and transjugular HVPG. The cut-off values for each HVPG stage were evaluated in the training dataset by the cut-off at sensitivity $\geq$95%. DeLong tests were used to compare AUCs in the model performance and robustness assessment phase. A 2-tailed $p < 0.05$ was considered statistically significant.

We performed model development and multiple linear regression in Python 3.7, mainly including TPOT, sklearn, and statsmodels. MedCalc Statistical Software v19.0.4 (MedCalc Software bvba, Ostend, Belgium) was used for AUC analysis and the DeLong test. SPSS 22.0 software (SPSS, Inc., Chicago, IL, USA) was used for descriptive analysis and Spearman correlation analysis.