



Original Article

Development and Validation of a Nomogram Based on Perioperative Factors to Predict Post-hepatectomy Liver Failure

Bin Xu^{1,2#} , Xiao-Long Li^{1,2#}, Feng Ye^{3#}, Xiao-Dong Zhu^{1,2}, Ying-Hao Shen^{1,2}, Cheng Huang^{1,2}, Jian Zhou^{1,2}, Jia Fan^{1,2}, Yong-Jun Chen^{3*}  and Hui-Chuan Sun^{1,2*} 

¹Department of Liver Surgery and Transplantation, Liver Cancer Institute, Zhongshan Hospital, Fudan University, Shanghai, China; ²Key Laboratory of Carcinogenesis and Cancer Invasion of Ministry of Education, Shanghai, China; ³Department of Hepatobiliary Surgery, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Received: 5 January 2021 | Revised: 11 February 2021 | Accepted: 17 February 2021 | Published: 15 March 2021

Abstract

Background and Aims: Post-hepatectomy liver failure (PHLF) is a severe complication and main cause of death in patients undergoing hepatectomy. The aim of this study was to build a predictive model of PHLF in patients undergoing hepatectomy. **Methods:** We retrospectively analyzed patients undergoing hepatectomy at Zhongshan Hospital, Fudan University from July 2015 to June 2018, and randomly divided them into development and internal validation cohorts. External validation was performed in an independent cohort. Least absolute shrinkage and selection operator (commonly referred to as LASSO) logistic regression was applied to identify predictors of PHLF, and multivariate binary logistic regression analysis was performed to establish the predictive model, which was visualized with a nomogram. **Results:** A total of 492 eligible patients were analyzed. LASSO and multivariate analysis identified three preoperative variables, total bilirubin ($p=0.001$), international normalized ratio ($p<0.001$) and platelet count ($p=0.004$), and two intraoperative variables, extent of resection ($p=0.002$) and blood loss ($p=0.004$), as independent predictors of PHLF. The area under receiver operating characteristic curve (referred to as AUROC) of the predictive model was 0.838 and

outperformed the model for end-stage liver disease score, albumin-bilirubin score and platelet-albumin-bilirubin score (AUROCs: 0.723, 0.695 and 0.663, respectively; $p<0.001$ for all). The optimal cut-off value of the predictive model was 14.7. External validation showed the model could predict PHLF accurately and distinguish high-risk patients. **Conclusions:** PHLF can be accurately predicted by this model in patients undergoing hepatectomy, which may significantly contribute to the postoperative care of these patients.

Citation of this article: Xu B, Li XL, Ye F, Zhu XD, Shen YH, Huang C, *et al.* Development and validation of a nomogram based on perioperative factors to predict post-hepatectomy liver failure. *J Clin Transl Hepatol* 2021;9(3):291–300. doi: 10.14218/JCTH.2021.00013.

Introduction

Hepatectomy is the main treatment for patients with benign or malignant liver lesions. However, patients undergoing liver resection are at increased risk for peri- and postoperative complications. Among these, post-hepatectomy liver failure (PHLF), defined as the impaired ability of the liver to maintain its synthetic, excretory and detoxifying functions, is one of the worst complications after hepatectomy and one of the major causes of perioperative mortality.^{1,2} Despite improvements in operative techniques, perioperative management and understanding of liver regeneration have improved the safety of liver resection over years, PHLF remains a challenge for patients undergoing hepatectomy and a concern of hepatic surgeons.³

Various assessment tools for liver function assessment and prediction of PHLF prior to surgery have been developed to reduce the incidence of PHLF and postoperative mortality. Indocyanine green retention rate at 15 min (ICG-R15) can measure the global liver function, and has been widely adopted in Eastern centers, whereas it is rarely used in Western countries due to its expensive cost and time-consuming requirement for performance.⁴ Clinic-biological scores like the model for end-stage liver disease (MELD) score, albumin-bilirubin (ALBI) score and platelet-albumin-bilirubin (PALBI) score are also adopted to evaluate the functional liver reserve,^{5–7} and are reported to accurately predict PHLF following hepatectomy.^{8–10} Volume and func-

Keywords: Hepatectomy; Post-hepatectomy liver failure; LASSO; Nomogram.

Abbreviations: ALB, albumin; ALBI, albumin-bilirubin; ALT, alanine aminotransferase; AUROC, area under receiver operating characteristic curve; CI, confidence interval; CT, computed tomography; DCA, decision curve analysis; FLR, future liver remnant; GGT, γ -glutamyl transpeptidase; HA, hyaluronic acid; Hb, hemoglobin; HBeAg, hepatitis B e antigen; HBsAg, hepatitis B surface antigen; HBV DNA, HBV DNA copy number; ICG-K, indocyanine green clearance rate constant; ICG-R15, indocyanine green retention rate at 15 min; INR, international normalized ratio; IQR, interquartile range; ISGLS, International Study Group of Liver Surgery; IV-col, type IV collagen; LASSO, least absolute shrinkage and selection operator; LN, laminin; LS, liver stiffness; MELD, model for end-stage liver disease; OR, odds ratio; PIIINP, procollagen III N-terminal peptide; P-ALB, pre-albumin; PALBI, platelet-albumin-bilirubin; PHLF, post-hepatectomy liver failure; PLT, platelet count; PPV, positive predictive value; ROC, receiver operating characteristics curve; TB, total bilirubin; WBC, white blood cell.

*These authors contributed equally to this work.

***Correspondence to:** Hui-Chuan Sun, Liver Cancer Institute and Zhongshan Hospital, Fudan University, 180 Fenglin Road, Shanghai 200032, China. ORCID: <https://orcid.org/0000-0003-3761-7058>. Tel: +86-21-3115-1990, Fax: +86-21-6403-7181, E-mail: sun.huichuan@zs-hospital.sh.cn; Yong-Jun Chen, Department of Hepatobiliary Surgery and Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Second Ruijin Road, Shanghai 200025, China. ORCID: <https://orcid.org/0000-0002-6486-2000>. Tel: +86-21-6431-4781, Fax: +86-21-6431-4781, E-mail: yongjunchen@yahoo.com

tion of the future liver remnant (FLR), as accessed by different imaging modalities, also have a superior ability to predict PHLF, but they could delay the time to surgery and also have financial constraints.^{11,12}

Intraoperative events can also influence the risk of PHLF.¹³ However, none of the models mentioned above include surgery-related factors, such as blood loss, extent of hepatectomy and intraoperative transfusions, to predict the probability of PHLF immediately after surgery.

The aim of this study was, therefore, to determine predictors of PHLF, including preoperative and intraoperative variables, and to build predictive models of PHLF in patients undergoing hepatectomy.

Methods

Study population

Five hundred and five consecutive patients who underwent hepatectomy at Zhongshan Hospital, Fudan University (Zhongshan cohort, from July 2015 to June 2018), and 167 consecutive patients at Ruijin Hospital, Shanghai Jiao Tong University School of Medicine (Ruijin cohort, from January 2018 to October 2019) were included in this study. Thirteen (2.6%) of the total patients in the Zhongshan cohort were excluded because of incomplete data. The remaining 492 patients in the Zhongshan cohort were randomly divided into a development cohort ($n=344$) and an internal validation cohort ($n=148$) using simple random sampling, with a random number seed of 2,017,0307. All patients in the Ruijin cohort were used as an external validation cohort ($n=167$) (Fig. 1).

The inclusion criteria were as follows: (i) patients who received hepatectomy; (ii) patients who received contrast-enhanced computed tomography (CT) scans or magnetic resonance imaging (MRI) conducted 1 week before resection; and (iii) patients who received blood routine test, biochemical test, coagulation function test, hepatitis B serologic test, liver fibrosis test¹⁴ and liver stiffness (LS)¹⁵ assessed by shear wave elastography conducted within 1 week before surgery.

This study was approved by the Institutional Ethics Committee of the two hospitals and was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Informed consent was obtained from all patients.

Data collection and definition

Clinical characteristics, including 22 preoperative variables, 3 intraoperative variables and 2 clinical outcomes, were recorded (Table 1). In addition, the MELD, ALBI and PALBI scores were calculated as reported,⁵⁻⁷ to compare with the model established in this study. No missing data were found for any patient in any of the study cohorts.

PHLF was defined as postoperative deterioration of liver function with an increase in the international normalized ratio (INR) and concomitant hyperbilirubinemia on or after postoperative day 5, as proposed by the International Study Group of Liver Surgery (commonly known as the ISGLS).¹

Presence of gastroesophageal varices and splenomegaly were confirmed by CT scans or MRI report.¹⁶⁻²² The extent of resection was defined by number of Couinaud's segments. Extent of resection ≥ 3 Couinaud's segments was defined as major resection, otherwise it was minor resection. The extent of resection was characterized as an intraoperative variable because the extent of resection planned preoperatively could differ from the actual extent during the surgery. Hospital stay was calculated from the date of surgery to date of discharge.

Statistical analysis

Categorical variables were expressed as counts and percentages, and were compared using Pearson's χ^2 analysis, Fisher's exact test or Mann-Whitney U test, as appropriate. Continuous variables were expressed as mean (\pm standard deviation) or median (interquartile range [IQR]) and were compared using Student's t -test, Mann-Whitney U test or Kruskal-Wallis test, as appropriate. The p -values were adjusted by Holm's method for multiple comparisons.

The least absolute shrinkage and selection operator (LASSO) logistic regression model with 10-fold cross-validation was performed to select perioperative variables associated with PHLF. As the group of variables selected by LASSO is not completely consistent every time due to randomness of cross-validation,²³ we repeated the same LASSO algorithm with the same candidate variables 1,000 times, and the most frequent group of selected variables was accepted as significant variables.

A multivariate binary logistic regression model was then produced to identify significant independent predictors of PHLF, with a removal significance level of 0.05. No evidence of non-log-linear relationship was found for all continuous variables. All significant variables were reserved in the final model because multicollinearity was not found.

Predictive performance was assessed using the receiver operating characteristic (ROC) curve and compared by DeLong's test. The optimal cut-off value of the logistic model was determined using ROC by maximizing the Youden index (sensitivity plus specificity minus 1). Calibration curves were plotted to assess the calibration of the model. Decision curve analysis (DCA) was conducted to determine the clinical utility of the model.²⁴ A nomogram was established based on the predictive model for the development cohort.

Statistical testing was carried out at the 2-sided tailed α level of 0.05. Data were analyzed using R version 3.6.2 (Vienna, Austria). Variable selection with LASSO was performed by the *cv.glmnet* function in the *glmnet* package. Binary logistic regression modeling was performed by the *glm* function. The nomogram was plotted by the *nomogram* function in the *rms* package. Delong's test was produced by the *roc.test* function in the *pROC* package. Calibration curves and DCA were analyzed by the *calibrate* function in the *rms* package and the *decision_curve* function in the *rmda* package, respectively.

Results

Clinical characteristics

The clinical characteristics of patients in the Zhongshan cohort are listed in Table 1. The comparison of clinical characteristics between the Zhongshan cohort and the Ruijin cohort is shown in Supplementary Table 1. The clinical characteristics were similar between the development and internal validation cohorts.

In the Zhongshan cohort, hospital stay of patients without PHLF (median [IQR]: 8 [7-10] days) was shorter than that of grade A (10 [8-13] days), grade B (10 [8-13] days) and grade C (16.5 [10-29] days) PHLF patients ($p < 0.001$ for all).

Establishment of the predictive model in the development cohort

All variables listed in Table 1 were analyzed. The result of variable selection by LASSO is shown in Supplementary Ta-

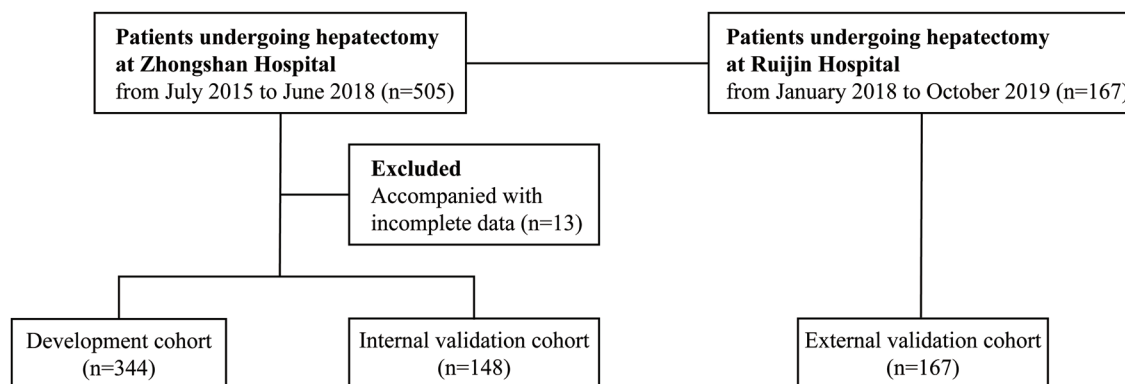


Fig. 1. Flowchart of this study’s design.

ble 2, which identified type IV collagen, total bilirubin (referred to as TB), albumin (ALB), INR, platelet count, extent of resection and blood loss as the most significantly related factors to PHLF.

The result of multivariate logistic regression analysis is shown in Table 2. These independent predictors were used to establish a predictive model, which was designated as the PHLF score, and visualized with a nomogram (Fig. 2).

Predictive accuracy and calibration of the PHLF score compared to other scores in the development cohort

The area under the ROC curve (AUROC) [95% confidence interval (CI)] of the PHLF score was 0.838 (0.790–0.885), which has better accuracy in predicting PHLF than the other three scores ($p < 0.001$ for all, compared by DeLong’s test): MELD score, 0.723 (0.664–0.782); ALBI score, 0.695 (0.630–0.758) and PALBI score, 0.663 (0.600–0.726), respectively (Fig. 3A). Calibration curves showed good agreement between prediction and observation (Fig. 4A). DCA revealed that the PHLF score provided superior net benefit over the other three scores (Fig. 3D).

Risk stratification based on the PHLF score in the development cohort

The optimal cut-off value of the PHLF score was determined to be 14.7 using ROC by maximizing the Youden index. The sensitivity, specificity, positive predictive value (referred to as PPV) and negative predictive value in predicting PHLF were 76.9%, 78.3%, 56.0%, and 90.4%, respectively.

Patients with PHLF score ≥ 14.7 were defined as the high-risk group, otherwise the patients were classified as the low-risk group. The incidence (55.6% vs. 9.6%, $p < 0.001$) and severity ($p < 0.001$) of PHLF were significantly different between the two groups (Table 3 and Fig. 5A).

Validation of the PHLF score in two independent cohorts

In the internal validation cohort, the AUROC of the PHLF score was 0.788 (0.693–0.884), which outperformed the other three scores in predicting PHLF (compared by DeLong’s test): MELD score ($p = 0.006$), ALBI score ($p = 0.010$) and PALBI score ($p = 0.002$), respectively (Fig. 3B). PHLF score showed good agreement between prediction and observation in calibration curve (Fig. 4B) and provided supe-

rior net benefit over other scores in the DCA curve (Fig. 3E). The incidence (42.6% vs. 12.9%, $p < 0.001$) and severity ($p < 0.001$) of PHLF were significantly different between high-risk and low-risk groups (Table 3 and Fig. 5B).

In the external validation cohort, the AUROC of the PHLF score was 0.750 (0.632–0.868), which was marginally superior to other three scores in predicting PHLF (compared by DeLong’s test): MELD score ($p = 0.103$), ALBI score ($p = 0.535$) and PALBI score ($p = 0.100$), respectively (Fig. 3C). PHLF score also provided superior net benefit over other scores in DCA analysis (Fig. 3F). The incidence (16.9% vs. 5.2%, $p = 0.013$) and severity ($p = 0.015$) of PHLF were also significantly different between the high-risk and low-risk groups (Table 3 and Fig. 5C).

Discussion

In this study, PHLF in patients undergoing hepatectomy could be accurately predicted immediately after surgery using routinely available variables, including three preoperative (TB, INR and platelet count) and two intraoperative (extent of resection and blood loss) factors. In addition, patients could be properly stratified in terms of the risk of PHLF, with a cut-off value of 14.7.

This study suggested that hepatic surgeons can take the optimized measures to prevent or manage PHLF perioperatively. On the basis that patients reserve good liver function, surgeons can calculate the maximum of intraoperative blood loss they can tolerate to prevent PHLF, because the extent of resection can be estimated by preoperative imaging data, and blood loss was the only unknown variable. This could remind surgeons to be more careful during surgery to reduce blood loss in order to prevent PHLF. Furthermore, surgeons could better inform patients and their families of the risk of PHLF after surgery. When the risk of PHLF is highly predicted, surgeons may suggest patients take medications to improve liver function and/or take systemic therapy to shrink the tumor as the best choice at that time, rather than surgery. Then, when the liver function or the tumor regression reaches a certain extent, surgery can be performed. If patients insist on performance of the surgery, surgeons can determine the appropriate level of postoperative care and extend the length of hospital stay, in addition performing a more careful operation.

The aim of this study was to establish a model to predict PHLF in patients undergoing hepatectomy. Many useful criteria and scores were demonstrated to predict the incidence of PHLF. One of the most classic models was “Makuuchi’s criteria”, representing a decision tree for selection of operative procedures in patients with impaired liver function

Table 1. Comparison of clinical characteristics between development and internal validation cohorts

Variables	Development cohort, n=344	Internal validation cohort, n=148	p-value
Age in years	56.4±11.2	57.0±10.9	0.625
Sex			0.714
Male	298 (86.6%)	130 (87.8%)	
Female	46 (13.4%)	18 (12.2%)	
Diabetes			0.878
No	291 (84.6%)	126 (85.1%)	
Yes	53 (15.4%)	22 (14.9%)	
HBsAg			0.078
–	56 (16.3%)	34 (23.0%)	
+	288 (83.7%)	114 (77.0%)	
HBeAg			0.763
–	282 (82.0%)	123 (83.1%)	
+	62 (18.0%)	25 (16.9%)	
HBV DNA			0.275
≤10 ³ /mL	198 (57.6%)	93 (62.8%)	
>10 ³ /mL	146 (42.4%)	55 (37.2%)	
Hb in g/L	143.0 (127.0–153.0)	142.0 (133.0–150.3)	0.948
WBC as ×10 ⁹ /L	5.3 (4.2–6.5)	5.3 (4.5–6.3)	0.587
PLT as ×10 ⁹ /L	148.0 (106.0–207.0)	162.5 (114–195.3)	0.400
TB in μmol/L	11.9 (8.8–15.9)	11.7 (9.2–16.5)	0.886
ALB in g/L	42.0 (39.0–45.0)	42.0 (39.0–45.0)	0.708
P-ALB in g/L	0.22 (0.17–0.26)	0.22 (0.18–0.26)	0.590
ALT in U/L	29.0 (20.0–43.0)	29.0 (20.8–42.3)	0.717
GGT in U/L	56.5 (33.0–108.0)	63.0 (34.8–115.5)	0.734
INR	1.01 (0.96–1.07)	1.03 (0.97–0.106)	0.345
HA in ng/mL	87.3 (64.2–135.2)	85.5 (60.0–135.4)	0.486
LN in ng/mL	50.0 (50.0–67.0)	50.0 (50.0–64.8)	0.536
PIIINP in ng/mL	6.5 (5.3–8.4)	6.7 (5.4–8.4)	0.829
IV-col in ng/mL	51.8 (50.0–83.9)	54.6 (50.0–79.6)	0.807
LS in kPa	12.0 (9.2–15.2)	11.4 (8.5–15.0)	0.240
Gastroesophageal varices			0.634
No	309 (89.8%)	135 (91.2%)	
Yes	35 (10.2%)	13 (8.8%)	
Splenomegaly			0.285
No	90 (26.2%)	32 (21.6%)	
Yes	254 (73.8%)	116 (78.4%)	
Extent of resection			0.395
Minor, <3 Couinaud's segments	250 (72.7%)	113 (76.4%)	
Major, ≥3 Couinaud's segments	94 (27.3%)	35 (23.6%)	
Hilar occlusion in min	15.0 (0.0–18.0)	14.5 (0.0–18.3)	0.740
Intraoperative blood loss in mL	200.0 (100.0–300.0)	200.0 (100.0–300.0)	0.816
Causes of hepatectomy			1
Malignant tumor	343 (99.7%)	148 (100%)	

(continued)

Table 1. (continued)

Variables	Development cohort, n=344	Internal validation cohort, n=148	p-value
Benign tumor	1 (0.3%)	0 (0%)	
<i>Clinical outcomes</i>			
PHLF [†]			0.330
No	253 (73.5%)	115 (77.7%)	
Yes	91 (26.5%)	33 (22.3%)	
PHLF grade [‡]			0.300
0	253 (73.5%)	115 (77.7%)	
A	63 (18.3%)	24 (16.2%)	
B	19 (5.5%)	8 (5.4%)	
C	9 (2.6%)	1 (0.7%)	
Hospital stay as median (IQR) in days	8 (7–11)	8.5 (7–11)	0.863

[†]PHLF was defined as postoperative deterioration of liver function with an increase in the INR and concomitant hyperbilirubinemia on or after postoperative day 5, as proposed by the ISGLS. [‡]Following the ISGLS definition of PHLF grade. ALT, alanine aminotransferase; Hb, hemoglobin; HBeAg, hepatitis B e antigen; HBsAg, hepatitis B surface antigen; HBV DNA, hepatitis B virus DNA; HA, hyaluronic acid; IV-col, type IV collagen; LN, laminin; LS, liver stiffness; P-ALB, pre-albumin; PIIINP, procollagen III N-terminal peptide; PLT, platelet count; WBC, white blood cell; GGT, γ-glutamyl transpeptidase.

Table 2. Independent predictors of PHLF after multivariate logistic analysis

Variables	β	OR	95% CI	p-value
Intercept	-15.585			
TB in μmol/L	0.074	1.077	1.029–1.128	0.001
INR [†] , per 0.1 increase	1.332	3.788	2.531–5.867	<0.001
PLT, per 10 ⁹ /L increase	-0.007	0.993	0.989–0.998	0.004
Extent of resection				
Minor, <3 segments		1		
Major, ≥3 segments	1.059	2.883	1.471–5.716	0.002
Blood loss [‡] , per 100 mL increase	0.132	1.141	1.043–1.251	0.004

[†]INR was multiplied by 10 and put into the multivariate binary logistic regression model. [‡]Blood loss was divided by 100 and put into the multivariate binary logistic regression model. The score and predicted probability of PHLF can be calculated using the following formulas: PHLF score: 0.074 × TB + 1.332 × INR (multiplied by 10) - 0.007 × PLT (per 10⁹/L) + 1.059 × extent of resection (major=1; minor=0) + 0.132 × blood loss (divided by 100). The predicted probability of PHLF=1/(1+exp(-PHLF score + 15.585)). CI, confidence interval; OR, odds ratio.

reserve, which included three determining factors: ascites, serum TB value, and ICG-R15.²⁵ Imamura *et al.*²⁶ reported zero-mortality after hepatectomy and only one patient developed PHLF from among nine hundred and fifteen con-

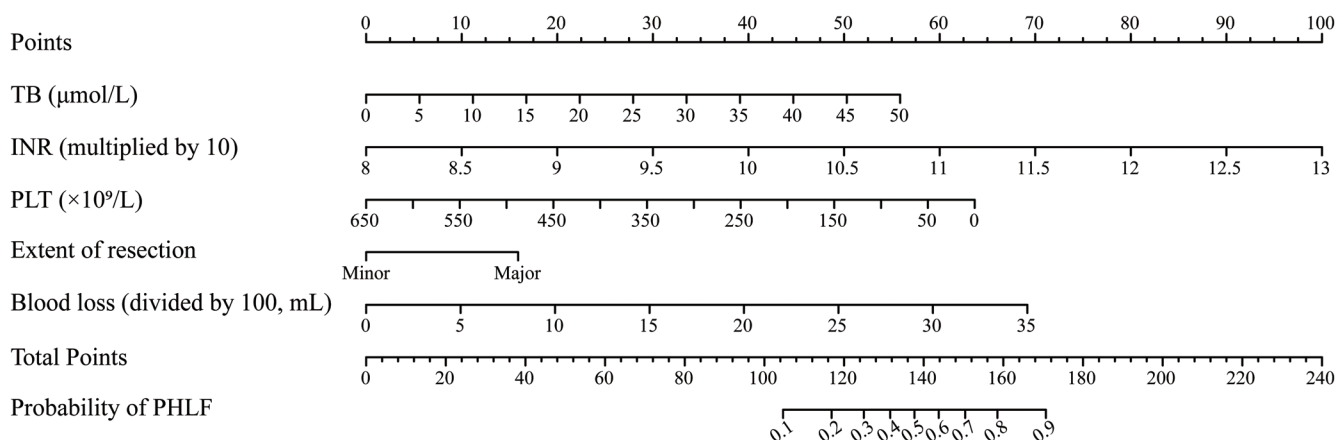


Fig. 2. Nomogram for the prediction of PHLF. The nomogram was established based on the development cohort. PHLF, post-hepatectomy liver failure.

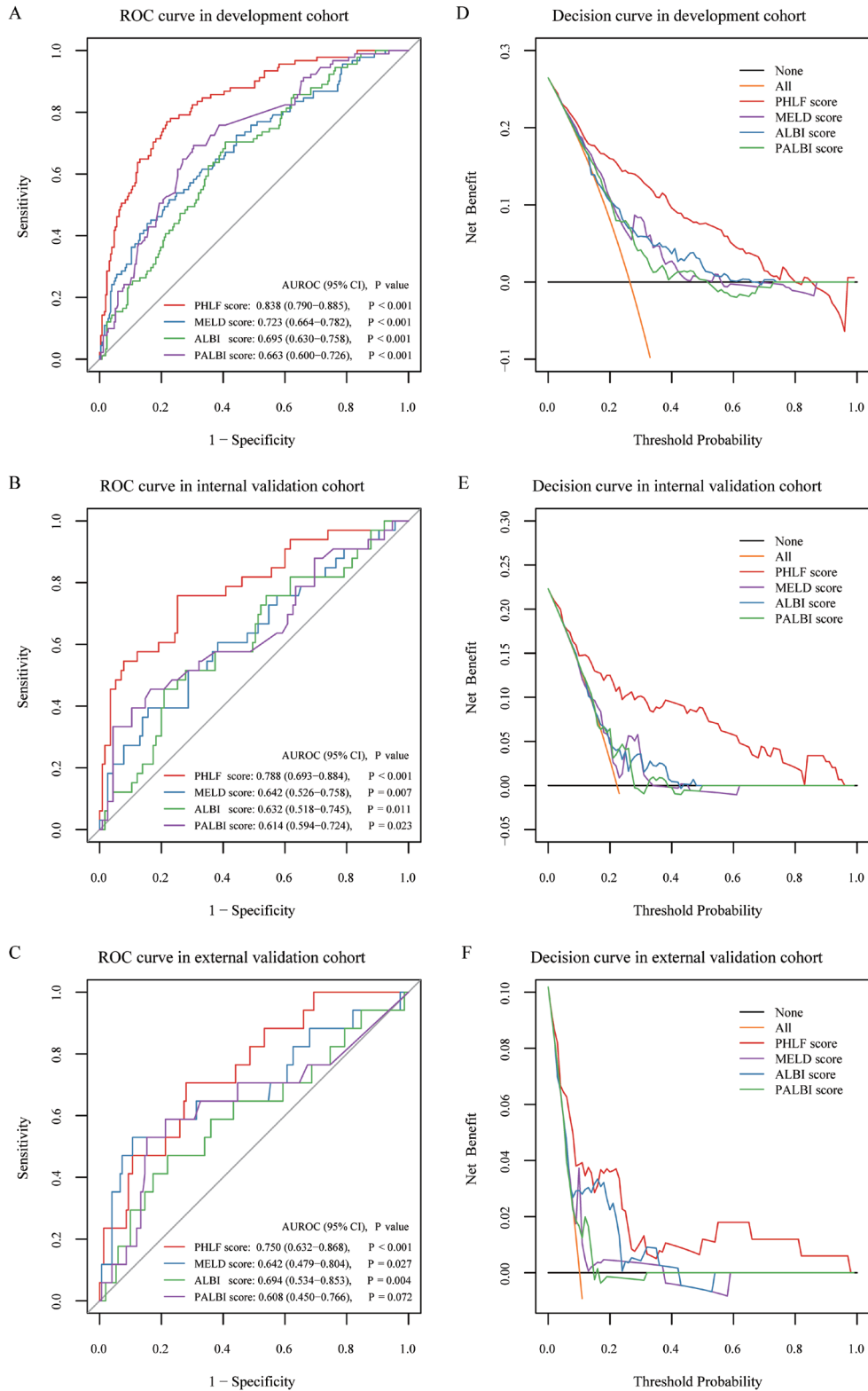


Fig. 3. ROC curves and decision curves for the prediction of PHLF. ROC curves of PHLF score, MELD score, ALBI score and PALBI score in the (A) development cohort, (B) internal validation cohort and (C) external validation cohort. Decision curves of PHLF score, MELD score, ALBI score and PALBI score in the (D) development cohort, (E) internal validation cohort and (F) external validation cohort. The orange line indicates the net benefit of assuming that all patients have PHLF. The black line indicates the net benefit of assuming no patients have PHLF. ALBI, albumin-bilirubin; MELD, model for end-stage liver disease; PALBI, platelet-albumin-bilirubin; PHLF, post-hepatectomy liver failure; ROC, receiver operating characteristics curve.

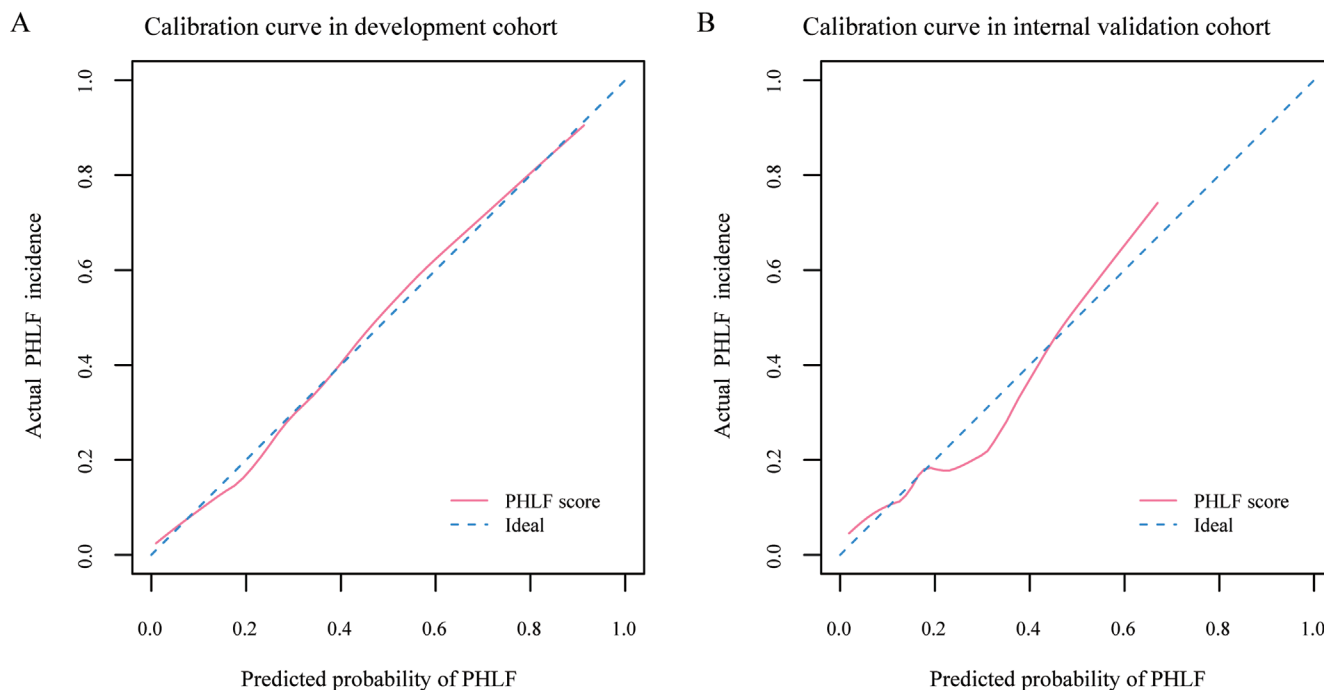


Fig. 4. Calibration curves for the prediction of PHLF. Calibration curves of the PHLF score in (A) development cohort and (B) internal validation cohort. The diagonal blue dashed line represents a perfect prediction by an ideal model. The pink solid line represents the performance of the predictive model, of which a closer fit to the diagonal blue dashed line represents a better prediction. PHLF, post-hepatectomy liver failure.

secutive patients within the criteria. However, within each category of Makuuchi’s criteria, there is a relatively wide range of hepatic function reserve and it does not take into account the individual variation in the FLR volume.²⁷

MELD,⁸ ALBI^{10,28} and PALBI²⁹ scores were previously reported to be accurate for the prediction of PHLF in patients with hepatocellular carcinoma. However, the ALBI score and PALBI score were based on a relatively low proportion of 727 (28.0%) patients undergoing hepatectomy.^{6,7} A study showed neither ALBI nor PALBI could predict survival of patients following transjugular intrahepatic portosystemic shunt creation,³⁰ which may suggest that they were not the

most suitable to predict PHLF for patients undergoing surgery. All patients in the current study underwent hepatectomy and their indications of surgery did not only include hepatocellular carcinoma, but all had reasoned hepatectomy. Our model may be more suitable for this target population and perform better in such.

Furthermore, ALB, which was included in both the ALBI and PALBI scores, was not included in our model. ALBI and PALBI scores were determined patients with data of ALB level, reported as median (IQR) of 35 (31–39) g/L, but the ALB level of patients in this study was 42 (39–45) g/L. Hence, the ability of ALB to predict PHLF was not as impor-

Table 3. Incidences of PHLF of high-risk and low-risk groups with a cut-off value of 14.7 by the PHLF score in development and two validation cohorts

	Development cohort, n=344			Internal validation cohort, n=148			External validation cohort, n=167		
	High-risk group, n=126	Low-risk group, n=218	p-value	High-risk group, n=47	Low-risk group, n=101	p-value	High-risk group, n=71	Low-risk group, n=96	p-value
PHLF [†]			<0.001			<0.001			0.013
No	56 (44.4%)	197 (90.4%)		27 (57.4%)	88 (87.1%)		59 (83.1%)	91 (94.8%)	
Yes	70 (55.6%)	21 (9.6%)		20 (42.6%)	13 (12.9%)		12 (16.9%)	5 (5.2%)	
PHLF grade [‡]			<0.001			<0.001			0.015
0	56 (44.4%)	197 (90.4%)		27 (57.4%)	88 (87.1%)		59 (83.1%)	91 (94.8%)	
A	48 (38.1%)	15 (6.9%)		12 (25.5%)	12 (11.9%)		9 (12.7%)	3 (3.1%)	
B	15 (11.9%)	4 (1.8%)		7 (14.9%)	1 (1.0%)		1 (1.4%)	2 (2.1%)	
C	7 (5.6%)	2 (0.9%)		1 (2.1%)	0 (0%)		2 (2.8%)	0 (0%)	

[†]PHLF was defined as postoperative deterioration of liver function with an increase in the INR and concomitant hyperbilirubinemia on or after postoperative day 5 as proposed by the ISGLS. [‡]Following the ISGLS definition of PHLF grade.

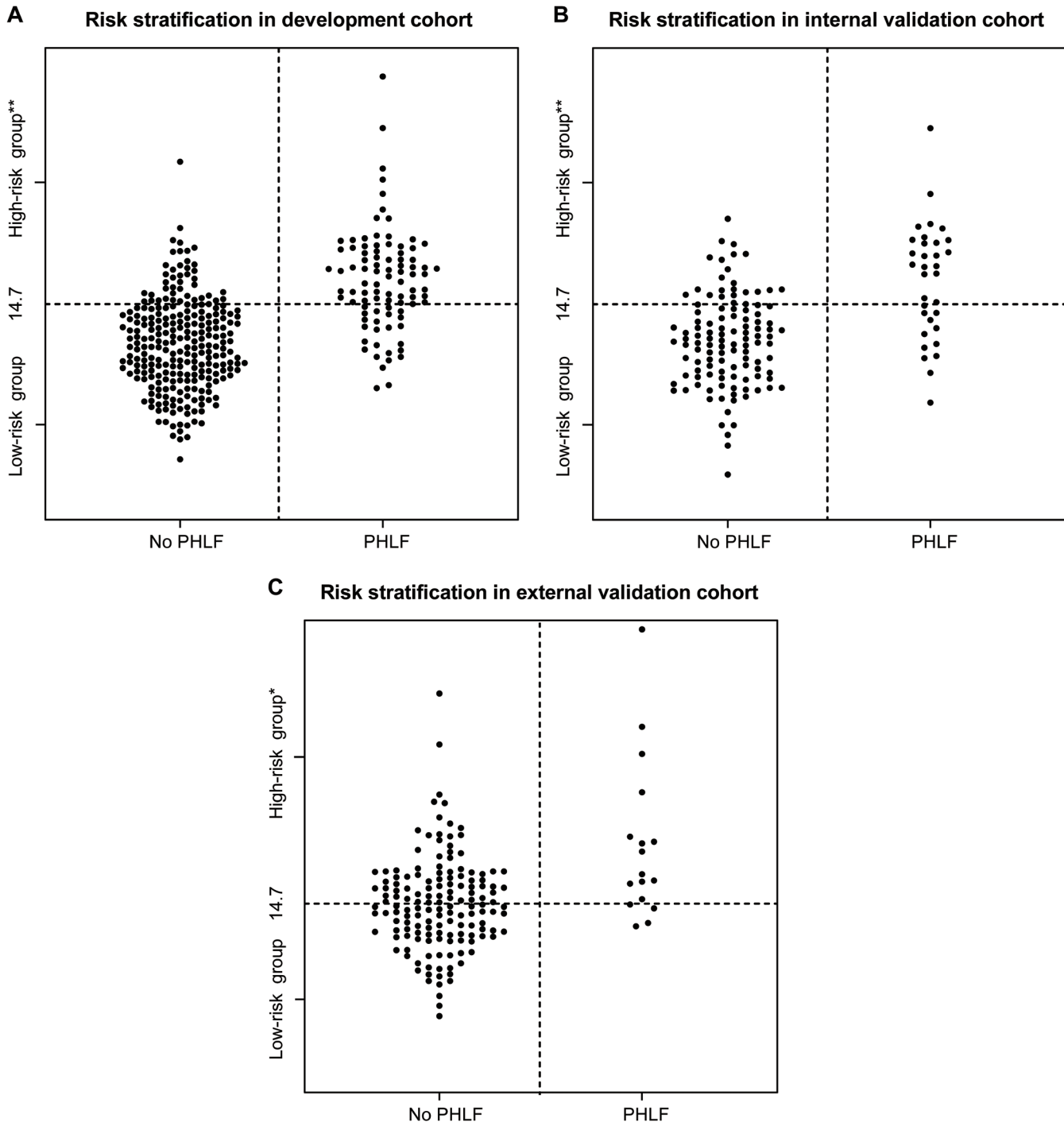


Fig. 5. Relationship between the value of the PHLF score and occurrence of PHLF in (A) development cohort, (B) internal validation cohort and (C) external validation cohort. The horizontal dotted line indicates the cut-off value of the PHLF score (14.7) for the prediction of PHLF. Patients with PHLF score ≥ 14.7 belong to the high-risk group, otherwise patients were classified in the low-risk group. * $p < 0.05$, ** $p < 0.001$, in comparison with the low-risk group. PHLF, post-hepatectomy liver failure.

tant as in patients with advanced diseases. The indocyanine green clearance rate constant (referred to as ICG-K) and ICG-R15 have been widely adopted in Eastern centers to measure liver function, but neither of them is a routine test in our center. Hwang *et al.*³¹ established a quantified model combined with ICG-K and FLR to predict PHLF,

which was similar to our model, both containing factors representing liver function and resected liver volume. However, surgery-related factors were not included in the model by Hwang *et al.*³¹ and in none of the models mentioned above. As PHLF could be influenced by surgery-related factors like blood loss, extent of hepatectomy,^{13,32} our model with intra-

operative variables could predict PHLF more accurately.

This study included almost all indicators of laboratory tests, important clinical signs available before surgery, and three intraoperative factors. To develop a predictive model based on as many as 25 candidate variables, we employed LASSO, which has been developed to overcome the limitations when too many predictors are needing to be analyzed, to guarantee the objectivity of variables included in the model.³³ In addition, DCA was performed to compare the clinical utility of different models, visualizing the clinical consequences of a diagnostic strategy.³⁴ Traditional metrics of diagnostic performance, such as AUROC, sensitivity and specificity only measure the accuracy of one prediction model against another, but fail to consider whether patients will really benefit from a specific model with the high predictive accuracy.²⁴

In addition, the PHLF score was validated externally and demonstrated satisfactory predictive accuracy and clinical utility. Furthermore, PHLF score can also stratify patients undergoing hepatectomy in terms of risk of PHLF in the external validation cohort.

The relatively low PPV indicates that patients who were actually at high risk of PHLF were not assigned to the high-risk group, and some factors that caused high risk of PHLF, such as repeated resection and tumor-related factors, were not included. Patients who underwent repeated resection were at higher risk of developing PHLF. In addition, this study enrolled patients either with benign or malignant lesions, so that tumor-related factors such as tumor number, size and biomarker were not included in the analysis.

This study has several limitations. First, we retrospectively investigated a group of patients with a relatively low proportion of grade B or C PHLF. Grade A PHLF represents a transient deterioration in liver function that does not require extra treatment. However, the hospital stay of grade A PHLF patients was longer than those without PHLF, indicating that this model is of clinical significance. Second, the predictive model cannot assess the severity of PHLF and make a classification according to the ISGLS grade, where different grades of PHLF are subject to different treatments. Third, the extent of resection was defined as minor (<3 segments) or major (≥3 segments) in this study, which could not exactly reflect the FLR volume. Because the volume of segment II+III (left lateral section) is significantly smaller than that of segment VII+VIII (right lateral section), and the latter may exceed the volume of segment II+III+IV (left liver).³⁵ In addition, inadequate FLR volume can lead to PHLF.³⁶ The performance of the predictive model could be improved through measuring FLR volume and FLR function by three-dimensional CT reconstruction or other image fusion techniques preoperatively.^{12,37}

In conclusion, this study showed that PHLF after hepatectomy can be accurately predicted by five simple and readily available perioperative variables, which may significantly contribute to the postoperative care of those patients and improving clinical outcomes.

Funding

This work was supported by the Leading Investigator Program of the Shanghai municipal government (17XD1401100), the National Key Basic Research Program (973 Program; 2015CB554005) from the Ministry of Science and Technology of China, and the National Natural Science Foundation of China (81672326 and 81871928, and 81871929).

Conflict of interest

The authors have no conflict of interests related to this publication.

Author contributions

Study concept and design (BX, XLL, FY, YJC and HCS), acquisition of data (BX, XLL, FY, XDZ, CH and YHS), analysis of data (BX, XLL and FY), interpretation of data (BX, XLL, FY, XDZ, CH, YHS, YJC and HCS), drafting of the manuscript (BX, XLL and FY), critical revision of the manuscript for important intellectual content (YJC and HCS), administrative, technical, or material support, study supervision (JZ and JF).

Data sharing statement

The data collected for this study will not be shared because the Institutional Ethics Committee of the two hospitals required that data involved in this study should be kept confidential. Other documents such as statistical analysis plan and informed consent form will be shared.

References

- [1] Rahbari NN, Garden OJ, Padbury R, Brooke-Smith M, Crawford M, Adam R, *et al*. Posthepatectomy liver failure: a definition and grading by the International Study Group of Liver Surgery (ISGLS). *Surgery* 2011;149(5):713–724. doi:10.1016/j.surg.2010.10.001.
- [2] Citterio D, Facciorusso A, Sposito C, Rota R, Bhoori S, Mazzaferro V. Hierarchic interaction of factors associated with liver decompensation after resection for hepatocellular carcinoma. *JAMA Surg* 2016;151(9):846–853. doi:10.1001/jamasurg.2016.1121.
- [3] Schreckenbach T, Liese J, Bechstein WO, Moench C. Posthepatectomy liver failure. *Dig Surg* 2012;29(1):79–85. doi:10.1159/000335741.
- [4] Cescon M, Colecchia A, Cucchetti A, Peri E, Montrone L, Ercolani G, *et al*. Value of transient elastography measured with FibroScan in predicting the outcome of hepatic resection for hepatocellular carcinoma. *Ann Surg* 2012;256(5):706–712; discussion 712–713. doi:10.1097/SLA.0b013e3182724ce8.
- [5] Kamath PS, Kim WR, Advanced Liver Disease Study Group. The model for end-stage liver disease (MELD). *Hepatology* 2007;45(3):797–805. doi:10.1002/hep.21563.
- [6] Johnson PJ, Berhane S, Kagebayashi C, Satomura S, Teng M, Reeves HL, *et al*. Assessment of liver function in patients with hepatocellular carcinoma: a new evidence-based approach—the ALBI grade. *J Clin Oncol* 2015; 33(6):550–558. doi:10.1200/JCO.2014.57.9151.
- [7] Roayaie S, Jibara G, Berhane S, Tabrizian P, Park J-W, Yang J, *et al*. PALBI-an objective score based on platelets, albumin & bilirubin stratifies HCC patients undergoing resection & ablation better than child's classification. *Hepatology* 2015;62(1, SI):631A–632A.
- [8] Kong FH, Miao XY, Zou H, Xiong L, Wen Y, Chen B, *et al*. End-stage liver disease score and future liver remnant volume predict post-hepatectomy liver failure in hepatocellular carcinoma. *World J Clin Cases* 2019;7(22):3734–3741. doi:10.12998/wjcc.v7.i22.3734.
- [9] Andreatos N, Amini N, Gani F, Margonis GA, Sasaki K, Thompson VM, *et al*. Albumin-bilirubin score: predicting short-term outcomes including bile leak and post-hepatectomy liver failure following hepatic resection. *J Gastrointest Surg* 2017;21(2):238–248. doi:10.1007/s11605-016-3246-4.
- [10] Zhang ZQ, Xiong L, Zhou JJ, Miao XY, Li QL, Wen Y, *et al*. Ability of the ALBI grade to predict posthepatectomy liver failure and long-term survival after liver resection for different BCLC stages of HCC. *World J Surg Oncol* 2018;16(1):208. doi:10.1186/s12957-018-1500-9.
- [11] Araki K, Harimoto N, Kubo N, Watanabe A, Igarashi T, Tsukagoshi M, *et al*. Functional remnant liver volumetry using Gd-EOB-DTPA-enhanced magnetic resonance imaging (MRI) predicts post-hepatectomy liver failure in resection of more than one segment. *HPB (Oxford)* 2020;22(2):318–327. doi:10.1016/j.hpb.2019.08.002.
- [12] Li C, Zhu A. Application of image fusion in diagnosis and treatment of liver cancer. *Applied Sciences* 2020;10(3):1171–1171. doi:10.3390/app10031171.
- [13] Kauffmann R, Fong Y. Post-hepatectomy liver failure. *Hepatobiliary Surg Nutr* 2014;3(5):238–246. doi:10.3978/j.issn.2304-3881.2014.09.01.
- [14] Shen Y, Shi G, Huang C, Zhu X, Chen S, Sun H, *et al*. Prediction of postoperative liver dysfunction by serum markers of liver fibrosis in hepatocellular carcinoma. *PLoS One* 2015;10(10):e0140932. doi:10.1371/journal.pone.0140932.
- [15] Shen Y, Zhou C, Zhu G, Shi G, Zhu X, Huang C, *et al*. Liver stiffness assessed by shear wave elastography predicts postoperative liver failure in patients with hepatocellular carcinoma. *J Gastrointest Surg* 2017;21(9):1471–1479. doi:10.1007/s11605-017-3443-9.
- [16] Perri RE, Chiorean MV, Fidler JL, Fletcher JG, Talwalkar JA, Stedheim L, *et al*. A prospective evaluation of computerized tomographic (CT) scanning as a screening modality for esophageal varices. *Hepatology* 2008;47(5):1587–1594. doi:10.1002/hep.22219.
- [17] Kim SH, Kim YJ, Lee JM, Choi KD, Chung YJ, Han JK, *et al*. Esophageal

varices in patients with cirrhosis: multidetector CT esophagography-comparison with endoscopy. *Radiology* 2007;242(3):759–768. doi:10.1148/radiol.2423050784.

[18] Lipp MJ, Broder A, Hudesman D, Suwandhi P, Okon SA, Horowitz M, *et al*. Detection of esophageal varices using CT and MRI. *Dig Dis Sci* 2011;56(9):2696–2700. doi:10.1007/s10620-011-1660-8.

[19] Shin SU, Lee JM, Yu MH, Yoon JH, Han JK, Choi BI, *et al*. Prediction of esophageal varices in patients with cirrhosis: usefulness of three-dimensional MR elastography with echo-planar imaging technique. *Radiology* 2014;272(1):143–153. doi:10.1148/radiol.14130916.

[20] Ronot M, Lambert S, Elkrief L, Doblaz S, Rautou PE, Castera L, *et al*. Assessment of portal hypertension and high-risk oesophageal varices with liver and spleen three-dimensional multifrequency MR elastography in liver cirrhosis. *Eur Radiol* 2014;24(6):1394–1402. doi:10.1007/s00330-014-3124-y.

[21] Bezerra AS, D'Ippolito G, Faintuch S, Szejnfeld J, Ahmed M. Determination of splenomegaly by CT: is there a place for a single measurement? *AJR Am J Roentgenol* 2005;184(5):1510–1513. doi:10.2214/ajr.184.5.01841510.

[22] Gunes SO, Akturk Y. Determination of splenomegaly by coronal oblique length on CT. *Jpn J Radiol* 2018;36(2):142–150. doi:10.1007/s11604-017-0704-1.

[23] Homrighausen D, McDonald DJ. Risk consistency of cross-validation with lasso-type procedures. *Statistica Sinica* 2018;27(3):1017–1036. doi:10.5707/ss.202015.0355.

[24] Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis. *JAMA* 2015;313(4):409–410. doi:10.1001/jama.2015.37.

[25] Makuuchi M, Kosuge T, Takayama T, Yamazaki S, Kakazu T, Miyagawa S, *et al*. Surgery for small liver cancers. *Semin Surg Oncol* 1993;9(4):298–304. doi:10.1002/ssu.2980090404.

[26] Imamura H, Seyama Y, Kokudo N, Maema A, Sugawara Y, Sano K, *et al*. One thousand fifty-six hepatectomies without mortality in 8 years. *Arch Surg* 2003;138(11):1198–1206; discussion 1206. doi:10.1001/archsurg.138.11.1198.

[27] Kobayashi Y, Kiya Y, Sugawara T, Nishioka Y, Hashimoto M, Shindoh J. Expanded Makuuchi's criteria using estimated indocyanine green clearance rate of future liver remnant as a safety limit for maximum extent of liver resection. *HPB (Oxford)* 2019;21(8):990–997. doi:10.1016/j.hpb.2018.12.001.

[28] Fagenson AM, Gleeson EM, Pitt HA, Lau KN. Albumin-bilirubin score vs model for end-stage liver disease in predicting post-hepatectomy outcomes. *J Am Coll Surg* 2020;230(4):637–645. doi:10.1016/j.jamcollsurg.2019.12.007.

[29] Lu LH, Zhang YF, Mu-Yan C, Kan A, Zhong XP, Mei J, *et al*. Platelet-albumin-bilirubin grade: risk stratification of liver failure, prognosis after resection for hepatocellular carcinoma. *Dig Liver Dis* 2019;51(10):1430–1437. doi:10.1016/j.dld.2019.04.006.

[30] Khabbaz RC, Lokken RP, Chen YF, Lipnik AJ, Bui JT, Ray CE Jr, *et al*. Albumin-bilirubin and platelet-albumin-bilirubin grades do not predict survival after transjugular intrahepatic portosystemic shunt creation. *Cardiovasc Intervent Radiol* 2018;41(7):1029–1034. doi:10.1007/s00270-018-1923-2.

[31] Hwang S, Ha TY, Song GW, Jung DH, Ahn CS, Moon DB, *et al*. Quantified risk assessment for major hepatectomy via the indocyanine green clearance rate and liver volumetry combined with standard liver volume. *J Gastrointest Surg* 2015;19(7):1305–1314. doi:10.1007/s11605-015-2846-8.

[32] Pruvot FR, Truant S. Major hepatic resection: from volumetry to liver scintigraphy. *HPB (Oxford)* 2016;18(9):707–708. doi:10.1016/j.hpb.2016.08.001.

[33] Guo P, Zeng F, Hu X, Zhang D, Zhu S, Deng Y, *et al*. Improved variable selection algorithm using a LASSO-type penalty, with an application to assessing hepatitis B infection relevant factors in community residents. *PLoS One* 2015;10(7):e0134151. doi:10.1371/journal.pone.0134151.

[34] Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ* 2016;352:i6. doi:10.1136/bmj.i6.

[35] Abdalla EK, Denys A, Chevalier P, Nemr RA, Vauthey JN. Total and segmental liver volume variations: implications for liver surgery. *Surgery* 2004;135(4):404–410. doi:10.1016/j.surg.2003.08.024.

[36] Kim HJ, Kim CY, Park EK, Hur YH, Koh YS, Kim HJ, *et al*. Volumetric analysis and indocyanine green retention rate at 15 min as predictors of post-hepatectomy liver failure. *HPB (Oxford)* 2015;17(2):159–167. doi:10.1111/hpb.12295.

[37] Theilig D, Steffen I, Malinowski M, Stockmann M, Seehofer D, Pratschke J, *et al*. Predicting liver failure after extended right hepatectomy following right portal vein embolization with gadoteric acid-enhanced MRI. *Eur Radiol* 2019;29(11):5861–5872. doi:10.1007/s00330-019-06101-2.