French Editorial Commentary from the ACHBPT

Risk predictors of post-hepatectomy liver failure: unraveling complexities and navigating challenges in clinical application

Marc-Anthony Chouillard[1], Christian Hobeika[1,2]^

[1]Département de Chirurgie Hépato-Bilio-Pancréatique et de Transplantation Hépatique, Hôpital Beaujon, AP-HP, Université Paris-Cité, Clichy, France; [2]UMR Inserm 1275 CAP Paris-Tech, Hôpital Lariboisière, Université Paris-Cité, Paris, France

Correspondence to: Christian Hobeika, MD, PhD. Département de Chirurgie Hépato-Bilio-Pancréatique et de Transplantation Hépatique, Hôpital Beaujon, AP-HP, Université Paris-Cité, 100 Bd du Général Leclerc, 92110 Clichy, France; UMR Inserm 1275 CAP Paris-Tech, Hôpital Lariboisière, Université Paris-Cité, Paris, France. Email: hobeikachristian@hotmail.com.

Comment on: Santol J, Kim S, Gregory LA, *et al.* An APRI+ALBI Based Multivariable Model as Preoperative Predictor for Posthepatectomy Liver Failure. Ann Surg 2023. [Epub ahead of print]. doi: 10.1097/SLA.0000000000006127.

Keywords: Liver failure; modeling prediction; hepatectomy

Submitted Feb 10, 2024. Accepted for publication Apr 22, 2024. Published online May 24, 2024.
doi: 10.21037/hbsn-24-81
View this article at: https://dx.doi.org/10.21037/hbsn-24-81

Studies related to the prediction of post-hepatectomy liver failure (PHLF) have seen a surge in recent literature. A PubMed search using the terms ((“pred*” OR “nomogra*” OR “model*”) AND (“mortality” OR “liver failure” OR “PHLF”) AND (“hepatect*” OR “liver resect*”)) revealed 29 relevant studies on PHLF prediction between January 2020 and November 2023, with 20 adhering to grade B/C International Study Group of Liver Surgery (ISGLS) definitions (*Table 1*). These studies are primarily enrolling patients with hepatocellular carcinoma (HCC). This underscores the growing interest in applying such predictive scores in routine clinical practice. However, the extent to which these predictive models can be effectively implemented in clinical settings remains unclear (21,22). Indeed, all studies are retrospective, and only a limited number underwent external validation. It is crucial to recognize that these scores predominantly emerge within surgical cohorts, where patients underwent prior meticulous selection, leading to tailored surgical strategies and the exclusion of specific candidates (21).

The study conducted by Santol *et al.* (23) introduces a novel predictive model using logistic regression to estimate the risk of PHLF based on the ISGLS grade B/C definition. The uniqueness of this model lies in the incorporation of the sum of aspartate aminotransferase (AST) to platelet ratio index (APRI) + albumin-bilirubin score (ALBI) as a composite variable, purported to comprehensively reflect liver functional reserve and parenchymal changes across various clinical scenarios [including fibrosis/cirrhosis/metabolic dysfunction-associated liver disease (MASLD) and chemotherapy-associated liver injury (CALI)/sinusoidal obstruction syndrome (SOS)] (24,25). This composite variable with sex, age, tumor type, and the extent of hepatectomy are integrated into the newly developed predictive model. The model undergoes training on the National Surgical Quality Improvement Program (NSQIP) database, comprising over 12,000 patients undergoing liver resection, and validation in an international multicenter cohort involving 10 institutions and 2,525 patients. The study demonstrates validated discriminatory performance with an area under the curve (AUC) of 0.74. It is a well-conducted study with noticeable strengths; it proposes a simple, objective, non-invasive tool to refine PHLF risk assessment trained in a large cohort of patients using already implemented tools (APRI and ALBI). The score underwent external validation with substantial statistical power, and its discriminatory performances were conserved in the validation cohort. It incorporates an online tool

^ ORCID: 0000-0002-9592-2520.

Table 1 Studies developing risk predictors of grade B/C PHLF according to the ISGLS definition between January 2020 and November 2023

Study	Country	Population	No. of patients[†]	No. events (ISGLS grade B/C PHLF) (%)[†]	Parameters included in the predictor	AUROC (95% confidence interval)
Fagenson *et al.* (1), 2020	USA	HCC	13,783	397 (2.9%)	ALBI	0.67
Yamamoto *et al.* [‡](2), 2020	Japan	HCC	876+250	92 (10.5%) + 27 (10.8%)	PLT, Alb, sFLR	0.749 (0.63–0.83)
Ye *et al.* [‡](3), 2020	China	HCC on HBV	1,200+387	154 (12.8%) + 78 (20.2%)	T-Bil, PLT, PreAlb, AST, PT, sFLR	0.820 (0.756–0.861)
Mai *et al.* (4), 2020	China	Hemi-hepatectomy for HCC	353	66 (18.7%)	Neural network, in order of importance: sFLR, T-Bil, PLT, AST, PT	0.876 (0.801–0.950)
Starlinger *et al.* (5), 2021	USA	NSQIP	12,055	96 (1.1%)[§]	ALBI + APRI	0.689
Dhir *et al.* (6), 2021	USA	NSQIP	10,808	316 (2.9%)	Age, BMI, sex, diabetes, dyspnea, ascites, corticosteroids, anticoagulation, biliary stent, chemotherapy, viral hepatitis, additional minor resections, biliary reconstructions, resection type, Na, Alb, T-Bil, INR	0.78
Wang *et al.* [‡](7), 2021	China	HCC	2,071+590	254 (9.5%) + 51 (8.6%)	T-Bil, Alb, GGT, PT, CSPH, major/minor resection	0.856 (0.803–0.909)
Zhong *et al.* (8), 2021	China	HCC	574	85 (14.8%)	Cirrhosis, blood loss, PALBI (Alb, T-Bil, PLT), FIB-4 major/minor resection, ascites	0.803 (0.723–0.883)
Cho *et al.* (9), 2022	South Korea	HCC	160	24 (15%)	ALBI, AFP, major/minor resection, liver stiffness (MRI)	0.871
Xiang *et al.* (10), 2021	China	HCC >10 cm	186	54 (29%)	Radiomics from CT, MELD, extent of resection	0.863 (0.750–0.975)
Takahashi *et al.* (11), 2022	Japan	HCC	361	39 (11%)	ALBI, sFLR	0.89 (0.83–0.96)
Alaimo *et al.* (12), 2022	International	HCC	1,785	106 (5.9%)	CCI, ALBI, TBS	0.67 (0.61–0.73)
Wang *et al.* (13), 2022	China	HCC	595	40 (6.7%)	C-P score, PLT, ALT, T-Bil, minor/major resection	0.753 (0.696–0.809)
Lei *et al.* [‡](14), 2022	China	HCC	668+192	93 (13.5%) + 18 (9.4%)	Age, sex, T-Bil, CSPH, PT	0.72 (0.65–0.78)
Xu *et al.* [‡](15), 2022	China	HCC >10 cm	514+97	52 (15.2%) + 23 (23.7%)	C-P score, blood loss, INR, cirrhosis, modified ALBI score	0.740 (0.624–0.856
Hobeika *et al.* [‡](16), 2022	France	HCC	323+165	19 (6.2%) + 22 (13.3%)	MELD, FIB-4, HCV, liver nodularity (CT), sFLR	0.867 (0.802–0.955)
Meng *et al.* (17), 2023	Asia	HCC	971	183 (18.8%)	Age, BMI, ascites, spleen/PLT ratio, blood loss, PreAlb, T-Bil	0.668

Table 1 (*continued*)

Table 1 (*continued*)

Study	Country	Population	No. of patients[†]	No. events (ISGLS grade B/C PHLF) (%)[†]	Parameters included in the predictor	AUROC (95% confidence interval)
Maehira *et al.* [¶](18), 2023	Japan	Major hepatectomy	65	21 (32%)	sFLR, ALT, PT	0.894
Long *et al.* [‡](19), 2023	China	HCC	223+43	59 (26.5%) + 7 (16%)	C-P score, sFLR, liver stiffness (elastometry), CSPH	0.845 (0.654–1.000)
Li *et al.* (20), 2023	China	HCC	276	65 (24%)	Radiomics from MRI, ICG-R15, ALBI	0.82 (0.72–0.91)

[†], where applicable, numbers of patients/events are reported as follows: training cohort (including internal validation when appropriate) + external validation cohort; [‡], studies including external validation; [§], considering ISGLS grade C PHLF only; [¶], no internal or external validation. ALBI score = [log10 bilirubin (μmol/L) × 0.66] + [albumin (g/L) × (−0.0852)]. APRI = [AST (U/L)/upper limit of normal (U/L)] × 100/platelet count (10^9/L). FIB-4 = age (years) × AST (U/L)/[platelet count (10^9/L) × $ALT^{1/2}$ (U/L)]. TBS^2 = (maximal diameter)^2 + (number of lesions)^2. PHLF, post-hepatectomy liver failure; ISGLS, International Study Group of Liver Surgery; AUROC, area under the receiver operating characteristic; HCC, hepatocellular carcinoma; ALBI, albumin-bilirubin score; PLT, platelets; Alb, albumin; sFLR, estimation of the future liver remaining; HBV, hepatitis B virus; T-Bil, total bilirubin; PreAlb, prealbumin; AST, aspartate aminotransferase; PT, prothrombin time; NSQIP, National Surgical Quality Improvement Program; APRI, AST to platelet ratio index; BMI, body mass index; INR, international normalized ratio; GGT, gamma glutamyl transpeptidase; CSPH, clinically significant portal hypertension; FIB-4, fibrosis-4 score; AFP, alpha-fetoprotein; MRI, magnetic resonance imaging; CT, computed tomography; MELD, model for end-stage liver disease; CCI, Charlson comorbidity index; TBS, tumor burden score; ALT, alanine aminotransferase; HCV, hepatitis C virus; C-P score, Child-Pugh score; ICG-R15, indocyanine green retention after 15 minutes.

(TELLAPRIALBI) to facilitate its application in routine practice. This study also reinforces the relevance of combining multiple biomarkers to capture the multifaceted mechanisms of liver functional recovery following liver resection (21).

Despite these unquestionable strengths, this study illustrates the methodological issues that predictive models raise. The first point is the clinical representativeness of included populations that directly impact the generalization of the results. The potential selection and information bias related to registries such as the NSQIP database become apparent when the data are compared with the validation cohort. Concerns arise as the low APRI + ALBI score (median =−4.17), low overall morbidity (17.7%), and grade B/C PHLF rates (2.6% of cases, constituting 59% of all PHLF patients) suggest a low-risk profile of patients undergoing minor resections (61.2%), frequently for colorectal metastasis (43.4%)—patients less likely to pose a risk of PHLF in routine practice. In contrast, the validation cohort displays expected results in a cohort at risk of PHLF with a 5.1% mortality rate and 11.6% grade B/C PHLF against a median APRI + ALBI of −2.29. However, the absence of histological data restrains the interpretability of the results. Of note, even in the validation cohort, the rate of HCC patients remains low, and only 6.9% of the 620 patients in the validation cohort with data on histology had severe fibrosis; thus, generalization to patients with underlying liver diseases who represent a group of high risk of PHLF, is uncertain.

A second matter of discussion lies in the construction of predictive models. Santol *et al.* (23) used the sum of APRI + ALBI, but to what extent it is best to collapse these two tests remains to be determined. APRI + ALBI alone performs poorly in the NSQIP cohort (AUC =0.698, pseudo-R^2=0.044); one could argue that including ALBI and APRI separately in a model would capture better performances. Other limitations stem from the lack of granularity in NSQIP data, exposing it to a high risk of unobserved heterogeneity—particularly critical when considering the scarcity and likely multifactorial nature of PHLF. The model's variables semi-automatedly selected are likely to incompletely apprehend the whole clinical picture, including comorbidities, underlying liver disease, volume optimization strategies, future remnant liver volume, type of surgical approach, tumor size, and number, etc. Model specifications are questionable (i.e., handling of missing data, high Akaike information criteria, wide confidence intervals, etc.), which could explain curious associations such as patients with benign lesions being associated with

markedly increased risk of PHLF compared to colorectal liver metastasis (CRLM) patients (26).

A consequence of the previous point is the models' performance and clinical applicability—the score's discriminatory performance (AUC) reported by Santol *et al.* (23) could be qualified as acceptable. Still, uncertainties arise concerning its calibration in the validation cohort (it is unclear in which population the Brier score has been calculated, and no calibration curve is available) (27). Discrepancies between observed and predicted probabilities for APRI + ALBI alone are substantial, mainly when the predicted risk falls below 10% while the observed PHLF rate exceeds 35–40%, even in major hepatectomies. Most patients are comprised within the 4th to 7th deciles of the score, predicting a slight variation in PHLF probabilities (2.5% to 6.5%). Such discrepancies substantially limit the score's applicability, notably through TELLAPRIALBI. While the latter is an elegant tool, questions arise regarding the threshold for a tolerable risk (and accepted degree of misclassification) that would warrant proceeding with liver resection and how this risk would translate into clinical reality (22).

Predictive scores for PHLF show promise in enhancing perioperative assessment within specific contexts (i.e., already selected patients), and the study by Santol *et al.* (23) is no exception. Biases depend on patient selection, model construction, and validation. Prospective evaluations of existing scores are necessary to validate their use as alternatives to reference methods in refining surgical indications.

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned by the editorial office, *Hepatobiliary Surgery and Nutrition*. The article did not undergo external peer review.

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at https://hbsn.amegroups.com/article/view/10.21037/hbsn-24-81/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Fagenson AM, Gleeson EM, Pitt HA, et al. Albumin-Bilirubin Score vs Model for End-Stage Liver Disease in Predicting Post-Hepatectomy Outcomes. J Am Coll Surg 2020;230:637-45.
2. Yamamoto G, Taura K, Ikai I, et al. ALPlat criterion for the resection of hepatocellular carcinoma based on a predictive model of posthepatectomy liver failure. Surgery 2020;167:410-6.
3. Ye JZ, Mai RY, Guo WX, et al. Nomogram for prediction of the international study Group of Liver Surgery (ISGLS) grade B/C Posthepatectomy liver failure in HBV-related hepatocellular carcinoma patients: an external validation and prospective application study. BMC Cancer 2020;20:1036.
4. Mai RY, Lu HZ, Bai T, et al. Artificial neural network model for preoperative prediction of severe liver failure after hemihepatectomy in patients with hepatocellular carcinoma. Surgery 2020;168:643-52.
5. Starlinger P, Ubl DS, Hackl H, et al. Combined APRI/ALBI score to predict mortality after hepatic resection. BJS Open 2021;5:zraa043.
6. Dhir M, Samson KK, Yepuri N, et al. Preoperative nomogram to predict posthepatectomy liver failure. J Surg Oncol 2021;123:1750-6.
7. Wang YY, Xiang BD, Ma L, et al. Development and Validation of a Nomogram to Preoperatively Estimate Post-hepatectomy Liver Dysfunction Risk and Long-term Survival in Patients With Hepatocellular Carcinoma. Ann Surg 2021;274:e1209-17.
8. Zhong W, Zhang F, Huang K, et al. Development and Validation of a Nomogram Based on Noninvasive Liver Reserve and Fibrosis (PALBI and FIB-4) Model to Predict Posthepatectomy Liver Failure Grade B-C in Patients with Hepatocellular Carcinoma. J Oncol 2021;2021:6665267.
9. Cho HJ, Ahn YH, Sim MS, et al. Risk Prediction Model Based on Magnetic Resonance Elastography-Assessed

Liver Stiffness for Predicting Posthepatectomy Liver Failure in Patients with Hepatocellular Carcinoma. Gut Liver 2022;16:277-89.
10. Xiang F, Liang X, Yang L, et al. CT radiomics nomogram for the preoperative prediction of severe post-hepatectomy liver failure in patients with huge (≥ 10 cm) hepatocellular carcinoma. World J Surg Oncol 2021;19:344.
11. Takahashi K, Gosho M, Kim J, et al. Prediction of Posthepatectomy Liver Failure with a Combination of Albumin-Bilirubin Score and Liver Resection Percentage. J Am Coll Surg 2022;234:155-65.
12. Alaimo L, Endo Y, Lima HA, et al. A comprehensive preoperative predictive score for post-hepatectomy liver failure after hepatocellular carcinoma resection based on patient comorbidities, tumor burden, and liver function: the CTF score. J Gastrointest Surg 2022;26:2486-95.
13. Wang J, Zhang Z, Shang D, et al. A Novel Nomogram for Prediction of Post-Hepatectomy Liver Failure in Patients with Resectable Hepatocellular Carcinoma: A Multicenter Study. J Hepatocell Carcinoma 2022;9:901-12.
14. Lei Z, Cheng N, Si A, et al. A Novel Nomogram for Predicting Postoperative Liver Failure After Major Hepatectomy for Hepatocellular Carcinoma. Front Oncol 2022;12:817895.
15. Xu MH, Xu B, Zhou CH, et al. An mALBI-Child-Pugh-based nomogram for predicting post-hepatectomy liver failure grade B-C in patients with huge hepatocellular carcinoma: a multi-institutional study. World J Surg Oncol 2022;20:206.
16. Hobeika C, Guyard C, Sartoris R, et al. Performance of non-invasive biomarkers compared with invasive methods for risk prediction of posthepatectomy liver failure in hepatocellular carcinoma. Br J Surg 2022;109:455-63.
17. Meng XQ, Miao H, Xia Y, et al. A nomogram for predicting post-hepatectomy liver failure in patients with hepatocellular carcinoma based on spleen-volume-to-platelet ratio. Asian J Surg 2023;46:399-404.
18. Maehira H, Iida H, Mori H, et al. Preoperative Predictive Nomogram Based on Alanine Aminotransferase, Prothrombin Time Activity, and Remnant Liver Proportion (APART Score) to Predict Post-Hepatectomy Liver Failure after Major Hepatectomy. Eur Surg Res 2023;64:220-9.
19. Long H, Peng C, Ding H, et al. Predicting symptomatic post-hepatectomy liver failure in patients with hepatocellular carcinoma: development and validation of a preoperative nomogram. Eur Radiol 2023;33:7665-74.
20. Li C, Wang Q, Zou M, et al. A radiomics model based on preoperative gadoxetic acid-enhanced magnetic resonance imaging for predicting post-hepatectomy liver failure in patients with hepatocellular carcinoma. Front Oncol 2023;13:1164739.
21. Primavesi F, Maglione M, Cipriani F, et al. E-AHPBA-ESSO-ESSR Innsbruck consensus guidelines for preoperative liver function assessment before hepatectomy. Br J Surg 2023;110:1331-47.
22. 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.
23. Santol J, Kim S, Gregory LA, et al. An APRI+ALBI Based Multivariable Model as Preoperative Predictor for Posthepatectomy Liver Failure. Ann Surg 2023. [Epub ahead of print]. doi: 10.1097/SLA.0000000000006127.
24. Shi JY, Sun LY, Quan B, et al. A novel online calculator based on noninvasive markers (ALBI and APRI) for predicting post-hepatectomy liver failure in patients with hepatocellular carcinoma. Clin Res Hepatol Gastroenterol 2021;45:101534.
25. Pereyra D, Starlinger P. ASO Author Reflections: APRI+ALBI: A Novel Tool for Estimating Chemotherapy-Associated Liver Injury in Patients with Colorectal Cancer Liver Metastasis Undergoing Liver Resection. Ann Surg Oncol 2019;26:598-9.
26. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128-38.
27. de Hond AAH, Steyerberg EW, van Calster B. Interpreting area under the receiver operating characteristic curve. Lancet Digit Health 2022;4:e853-5.

Cite this article as: Chouillard MA, Hobeika C. Risk predictors of post-hepatectomy liver failure: unraveling complexities and navigating challenges in clinical application. HepatoBiliary Surg Nutr 2024;13(3):500-504. doi: 10.21037/hbsn-24-81