Annals of Surgery Publish Ahead of Print
DOI:10.1097/SLA.0000000000006127

OPEN

An APRI+ALBI based multivariable model as preoperative predictor for posthepatectomy liver failure

Jonas Santol MD[1,2,3]; Sarang Kim[4]; Lindsey A. Gregory[2]; Ruth Baumgartner MD[5]; Anastasia Murtha-Lemekhova MD[6]; Emrullah Birgin MD[7], Severin Gloor MD[8]; Eva Braunwarth MD[9]; Markus Ammann MD[10]; Johannes Starlinger MD, PhD[11]; David Pereyra MD, PhD[12]; Daphni Ammon MD[12]; Marijana Ninkovic MD[9]; Anna E. Kern[12]; Benedikt Rumpf MD[13]; Gregor Ortmayr MD[14]; Yannic Herrmann[4]; Yawen Dong MD[1]; Felix X. Huber[4]; Jeremias Weninger[4]; Cornelius A. Thiels DO, MBA[2]; Susanne G. Warner MD[2]; Rory L. Smoot MD[2]; Mark J. Truty MD[2]; Michael L. Kendrick MD[2]; David N. Nagorney MD[2]; Sean P. Cleary MD[2,15]; Guido Beldi MD[8]; Nuh N. Rahbari MD[7]; Katrin Hoffmann MD, MBA[6], Stefan Gilg MD, PhD[16]; Alice Assinger PhD[3]; Thomas Gruenberger MD[1]; Hubert Hackl PhD[17]; Patrick Starlinger MD, PhD[2,12,18]

1 Department of Surgery, HPB Center, Vienna Health Network, Clinic Favoriten and Sigmund Freud Private University, Vienna, Austria
2 Department of Surgery, Division of Hepatobiliary and Pancreatic Surgery, Mayo Clinic, Rochester, MN, USA.
3 Institute of Vascular Biology and Thrombosis Research, Center for Physiology and Pharmacology, Medical University of Vienna, Vienna, Austria.
4 Medical University of Vienna, Vienna, Austria
5 Department of Oncology, Karolinska University Hospital, Stockholm, Sweden
6 Department of General, Visceral, and Transplantation Surgery, Heidelberg University Hospital, Heidelberg, Germany.

7 Department of Surgery, University Medicine Mannheim, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany.
8 Department of Visceral Surgery and Medicine, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland
9 Department of Visceral, Transplantation and Thoracic Surgery, Medical University of Innsbruck, Innsbruck, Austria
10 Department of Surgery, State Hospital Wiener Neustadt, Wiener Neustadt, Austria
11 Howto Health GmbH, Berlin, Germany.
12 Department of General Surgery, Division of Visceral Surgery, Medical University of Vienna, Vienna, Austria
13 Hospital Barmherzige Schwestern, Department of Surgery, Vienna, Austria
14 Center for Cancer Research, Medical University of Vienna, Vienna, Austria
15 Division of General Surgery, University of Toronto
16 Department of HPB surgery, Karolinska University Hospital, Stockholm, Sweden.
17 Institute of Bioinformatics, Biocenter, Medical University of Innsbruck, Innsbruck, Austria.
18 Center for Physiology and Pharmacology, Medical University of Vienna, Vienna Austria.

Keywords: APRI+ALBI, ICG clearance, ALICE grade, 90 day mortality, liver surgery.

Corresponding Author: Patrick Starlinger MD, PHD; Mayo Clinic, Rochester, 200 First St. SW, Rochester, MN 55905, USA; +1 507-293-1632; Starlinger.patrick@mayo.edu

Disclosure of Conflicts: Patrick Starlinger and Johannes Starlinger were involved in the development of a freely available smartphone-first application (TELLAPRIALBI, https://tellaprialbi.howto.health), all other authors have no conflict of interest to declare.

Funding Statement: There are no sources of funding needing to be declared.

Author Contributions

Jonas Santol made substantial contributions to conception and design, acquisition of data and analysis and interpretation of data, participated in drafting the article, revising it critically for important intellectual content and gave final approval of the version to be published.

Sarang Kim made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Lindsey A. Gregory made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Ruth Baumgartner made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Anastasia Murtha-Lemekhova made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Emrullah Birgin made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Severin Gloor made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Eva Braunwarth made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Markus Amman made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Johannes Starlinger made substantial contributions to acquisition of data and analysis and interpretation of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

David Pereyra made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Daphni Ammon made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Marijana Ninkovic made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Anna E. Kern made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Benedikt Rumpf made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Gregor Ortmayr made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Yannic Herrmann made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Yawen Dong made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Felix X. Huber made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Jeremias Weninger made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Cornelius A. Thiels made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Susanne G. Warner made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Rory L. Smoot made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Mark J. Truty made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Michael L. Kendrick made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

David N. Nagorney made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Sean P. Cleary made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Guido Beldi made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Nuh N. Rahbari made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Katrin Hoffmann made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Stefan Gilg made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Alice Assinger made substantial contributions to conception and design, participated in drafting the article, revising it critically for important intellectual content and gave final approval of the version to be published.

Thomas Gruenberger made substantial contributions to acquisition of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Hubert Hackl made substantial contributions to acquisition of data and analysis and interpretation of data, participated in revising the article critically for important intellectual content and gave final approval of the version to be published.

Patrick Starlinger made substantial contributions to conception and design, acquisition of data and analysis and interpretation of data, participated in drafting the article, revising it critically for important intellectual content and gave final approval of the version to be published.

Structured Abstract

Objective and Background: Clinically significant posthepatectomy liver failure (PHLF B+C) remains the main cause of mortality after major hepatic resection. This study aimed to establish an APRI+ALBI, aspartate aminotransferase to platelet ratio (APRI) combined with albumin-bilirubin grade (ALBI), based multivariable model (MVM) to predict PHLF and compare its performance to indocyanine green clearance (ICG-R15 or ICG-PDR) and albumin-ICG evaluation (ALICE).

Methods: 12,056 patients from the National Surgical Quality Improvement Program (NSQIP) database were used to generate a MVM to predict PHLF B+C. The model was determined using stepwise backwards elimination. Performance of the model was tested using receiver operating characteristic curve analysis and validated in an international cohort of 2,525 patients. In 620 patients, the APRI+ALBI MVM, trained in the NSQIP cohort, was compared with MVM's based on other liver function tests (ICG clearance, ALICE) by comparing the areas under the curve (AUC).

Results: A MVM including APRI+ALBI, age, sex, tumor type and extent of resection was found to predict PHLF B+C with an AUC of 0.77, with comparable performance in the validation cohort (AUC 0.74). In direct comparison with other MVM's based on more expensive and time-consuming liver function tests (ICG clearance, ALICE), the APRI+ALBI MVM demonstrated equal predictive potential for PHLF B+C. A smartphone application for calculation of the APRI+ALBI MVM was designed.

Conclusion: Risk assessment via the APRI+ALBI MVM for PHLF B+C increases preoperative predictive accuracy and represents an universally available and cost-effective risk assessment prior to hepatectomy, facilitated by a freely available smartphone app.

Introduction

With an average incidence of 10-15%, posthepatectomy liver failure (PHLF) poses a significant risk for patients undergoing liver surgery. Responsible for nearly 50% of short-term postoperative (postOP) mortality after major liver resection, PHLF is the main cause of death after hepatic resection[1]. Besides simple volumetric analyses, the main challenge for preoperative (preOP) risk assessment is the significant heterogeneity of underlying liver diseases and concomitantly affected liver function in patients evaluated for hepatic resection, especially in patients with primary liver cancer. Further, in patients with metastatic disease to the liver, liver function is critically affected by neoadjuvant chemotherapy causing chemotherapy-associated liver injury (CALI).[2-5] Characteristics of CALI can vary in severity between patients, from steatosis to sinusoidal obstruction syndrome (SOS) or chemotherapy associated steatohepatitis (CASH).[6] PostOP liver function is also challenged by the recent rise in non-alcoholic steatohepatitis and subsequent non-alcoholic fatty liver disease, now frequently observed in patients undergoing hepatic resection.[7, 8] Particularly co-incidence of these liver diseases/injuries adds further complexity to preOP liver function assessment and PHLF risk-assessment. Measurement and metric expression of these conditions remain a major challenge, as we recently concluded in the European Consensus Guidelines for preoperative liver function testing.[9] Dynamic liver function measurement through indocyanine green (ICG) clearance and, building upon ICG-clearance, the albumin-indocyanine green evaluation (ALICE) grade, or non-invasive tests like fibrosis-4 (FIB-4) index and APRI+ALBI have all been evaluated for their ability to predict PHLF.[10-13] Based on routine laboratory parameters APRI+ALBI, the summative combination of the aspartate aminotransferase (AST) to platelet ratio (APRI) and the albumin - bilirubin (ALBI) grade, has been evaluated for its correlation with chronic liver disease and CALI.[10] APRI+ALBI could be shown to closely reflect the development of CALI after neoadjuvant chemotherapy, as well as the subsequent recovery of liver function in patients with colorectal cancer liver metastases (CRCLM) during the chemotherapy break before surgery.[10] Also, APRI+ALBI could predict postOP liver dysfunction after liver surgery in both patients with primary liver cancer and CRCLM, outperforming both APRI and ALBI alone.[10, 14]

The aim of this study was to develop an APRI+ALBI based preoperative multivariable model to predict PHLF and validate its performance in an independent international multicenter cohort. Further, we aimed to compare the predictive potential of this model to models based on more expensive, time consuming and sometimes even invasive tests such as ICG-

clearance, the ALICE score or the FIB-4 score.Ultimately, we aimed to develop a clinically easily accessable tool to allow accurate preoperative risk assessment.

Patients and Methods

Study cohort

For this study 12,056 patients from the American national surgery quality improvement program (NSQIP) database, who underwent elective hepatic resection and had preOP APRI+ALBI scores available, were included to calculate a multivariable prediction model, predicting clinically relevant PHLF grade B and C (PHLF B+C). This model was then validated in an international multicenter cohort of 2,525 patients from 10 different institutions. Participating institutions were Clinic Favoriten (Vienna, Austria), General Hospital Vienna (Vienna, Austria), Clinic Landstraße (Vienna, Austria), Mayo Clinic Rochester (Minnesota, USA), Karolinska Institute (Stockholm, Sweden), University Hospital Heidelberg (Heidelberg, Germany), University Hospital Mannheim (Mannheim, Germany), Inselspital University Hospital Bern (Bern, Switzerland), University Hospital Innsbruck (Innsbruck, Austria), State Hospital Wiener Neustadt (Wiener Neustadt, Austria). Lastly, in 620 patients out of our international multicenter cohort the multivariable APRI+ALBI score model (as trained in the NSQIP cohort) was directly compared to 4 different multivariable models. These models included other liver function tests, namely ICG clearance, the ALICE grade and the FIB-4 index. These were trained in the 620 patients from 10 different international centers. Characteristics of all cohorts are summarized in table 1 and 2. A flow chart of the study design can be found in the supplement. (Supp. fig. 1, Supplemental Digital Content 1, http://links.lww.com/SLA/E919)

All patients underwent elective minor or major hepatic resection between 2000 and 2021. Patient data were collected from prospectively maintained institutional databases or collected retrospectively. All patients had preoperative APRI+ALBI scores available. Underlying tumor entities included CRCLM, primary liver cancer (hepatocellular carcinoma, cholangiocellular carcinoma), benign tumors and other malignancies with metastases to the liver. Exclusion criteria included under 18 years of age, pregnancy and decompensated liver cirrhosis.

All patients either gave written informed consent or data was collected from national registries according to national laws. The study was approved by the institutional ethics committees of the participating institutions (National Surgical Quality Improvement Program (NSQIP) database: # 19-007654; Vienna: # EK 2032/2013; Rochester: # 21-006411;

Stockholm: # 2020-04493; Heidelberg: # S-429/2021; Mannheim: 2012-293N-MA; Bern: # 2018–01576; Innsbruck: 1076/2017, 1052/2019; Wiener Neustadt: GS4 – EK-4/568-2018).

Definition of liver resections

Liver resections were classified according to the International Hepato-Pancreato-Biliary Association Brisbane 2000 nomenclature as minor (3 < segments) and major hepatectomy (3 ≥ segements).[15]

Measurement of routine blood parameters

AST, alanine aminotransferase, albumin, serum bilirubin (SB), alcalic phosphatase, gamma-glutamyltransferase (GGT), prothrombin time (PT) and platelet counts were measured in appropriate samples by routine laboratory blood tests.

ICG Measurement

Perioperative indocyanine green (ICG) clearance testing was performed as previously described.[16] ICG clearance was measured in plasma disappearance rate (ICG-PDR) and retention 15 minutes after administration (ICG-R15).

Calculation of scores

APRI+ALBI, Albumin-ICG evaluation (ALICE) and fibrosis-4 (Fib-4) index were calculated according to previously published formulae.[10, 17, 18]

Definition of postoperative outcome parameters

Follow-up period was 90 days. Postoperative morbidity was defined as described by Dindo et al, with severe morbidity calssified as morbidity grade 3 or higher.[19] Posthepatectomy liver failure (PHLF) was defined and graded according to the criteria put forth by the international study group of liver surgery.[20] PHLF was classified as an elevation of SB and prolonged PT persisting on postoperative day (POD) 5. When deranged values of SB and PT were measured already prior to the operation (preOP), SB had to be higher and PT lower than the abnormal preOP values. If patients were excluded from routine blood workup due to good clinical performance or because of early discharge, patients were classified as no PHLF. To better represent the percentage of patients with clinically relevant, symptomatic PHLF, PHLF was defined as PHLF grades B and C (PHLF B+C) and no PHLF was defined as no PHLF or PHLF grade A.

Statistical methods

In order to identify the non-linear effect of APRI+ALBI scores on PHLF B+C a multivariable logistic regression model was learned on 12,056 patients (NSQIP cohort) using all available parameters. Using stepwise backwards feature elimination the best model based on the minimal Akaike information criterion (AIC) was determined and non-significant parameters were excluded from the model without compromising quality. The predicted PHLF B+C probabilities were compared with observed PHL B+C probabilities and model fit/performance parameters such as Nagelkerke Pseudo R^2 and Brier score were calculated using R packages rms, rcompanion. The prediction performance of the final model was assessed using receiver operating characteristics (ROC) and the area under curve (AUC) was calculated not only in the NSQIP cohort but also in the international multicenter validation cohort including 2,525 patients using the R package ROCR. In 620 patients out of these 2,525 patients (validation cohort) the prediction performance of the final model, trained in the NSQIP cohort, for PHLF B+C (PHLF C, 90 day mortality) were compared to multivariable models including the identical variables (sex, age, extent of resection, tumor type) but instead of APRI+ALBI including one of the liver function tests ICG-R15, ICG-PDR, ALICE grade, or FIB-4 index. For direct comparison, a univariate logistic regression model was trained for each of these parameters individually in the same cohort (620 patients). The R package pROC was used to calculate the 95%-confidence intervals for AUCs and to test differences between AUCs based on a resampling strategy using bootstrap analysis with 2,000 repetitions. The performance of the trained models were tested in different subgroups such as major resection or the respective tumor type. All analyses were performed using the statistical software environment R (R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/; v4.3.0) and SPSS (IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.).

Results

APRI+ALBI increases in nonlinear fashion and reflects risk increase for PHLF B+C

To evaluate the dynamic risk increase for PHLF B+C, a multivariable model for the prediction of clinically relevant PHLF B+C based on the APRI+ALBI score was calculated (Fig.1). All available parameters from the NSQIP cohort were used for calculation of the model. Stepwise backwards feature elimination was used to determine the optimal model.

The parameters included in the final model were the APRI+ALBI score, as well as sex, age, tumor type and extent of resection (Table 3). The model was trained using the NSQIP cohort (N = 12,056, Table 1) and then validated using an international multicenter cohort of 2,525 patients (Table 2). While observed PHLF B+C probability was higher in the validation cohort, it closely followed risk increase predicted by the model (Fig. 1A). We could observe that a higher APRI+ALBI score was associated with a concomitant exponential increasing risk for PHLF B+C. As major liver resections pose a particular high risk to develop PHLF, the model was tested specifically in the patient subgroup undergoing major liver resection (Fig. 1B). A comparable nonlinear risk increase with rising APRI+ALBI scores could be demonstrated. As expected, in comparison to the entire cohort, probability for PHLF B+C was higher in the major resection subgroup at lower APRI+ALBI deciles.

PHLF B+C risk is highly dependent on tumor type and extent of resection. To visualize this, probability for PHLF B+C development calculated by the multivariable model and the associated APRI+ALBI score in deciles is illustrated in Figure 1C. Patients are grouped for extent of resection (complete cohort, major resection subgroups) and tumor type (CRCLM, primary liver cancer, other malignancies). Depending on the patient group comparable APRI+ALBI values lead to a different probability for the occurrence of PHLF B+C calculated by the multivariable model (Fig. 1C).

An APRI+ALBI multivariable model accurately predicts development of PHLF B+C

Next, we aimed to evaluate the multivariable model's performance in predicting PHLF B+C. Performance of the model was tested using ROC curve analysis and the area under the curve (AUC) (Fig. 2A). To validate the APRI+ALBI multivariable model trained in the NSQIP cohort, we tested PHLF B+C prediction (AUC) of the multivariable model in the international multicenter cohort. The model showed comparable predictive performance for development of PHLF B+C in the validation cohort (Fig. 2B).

The APRI+ALBI score shows comparable results with established liver function tests in the prediction of adverse outcome after liver resection

In 620 patients of our international multicenter cohort (validation cohort) ICG-clearance, ALICE grade and the FIB-4 index, as 3 established liver function tests, were available for direct comparison to the APRI+ALBI score. For a descriptive comparison of the predictive potential of the different liver function tests, univariate models for the prediction of PHLF B+C were trained in these 620 patients for APRI+ALBI, ICG-R15, ICG-PDR, ALICE grade,

and FIB-4 index. The performance of all models for predicting PHLF B+C was evaluated using ROC curve analysis. The AUC for APRI-ALBI was then tested against the AUC of all other parameters by bootstrap analyses (Fig. 3A). To evaluate the discriminatory potential of all models for fulminant PHLF and short-term postOP mortality, ROC curves analysis was performed for PHLF grade C and 90 day mortality as well (Fig. 3B-C). The performance (AUC) of the APRI+ALBI model for prediction of PHLF B+C (and also prediction of PHLF C and 90 day mortality) was higher compared to that of ALICE, ICG-R15, ICG-PDR, and FIB-4 index when tested in the same training data, but were not significantly different when using resampling (Fig. 3A-C), indicating that APRI+ALBI showed at least equal performance to the other liver function tests. While predominantly not statistically significant, there was a tendency for superior predictive potential for PHLF B+C when comparing APRI+ALBI with ICG-R15, ICG-PDR and FIB-4 index (Fig. 3A). Similarly, while not statistically significant, APRI+ALBI also showed a tendency for superior discriminatory potential for PHLF C and 90 day mortality when compared with ICG-R15 and ICG-PDR (Fig. 3B-C).

An APRI+ALBI score based multivariable model shows equal performance for prediction of adverse outcome after liver resection in comparison with other multivariable models associated with more time consuming liver function tests

In the 620 patients, with detailed liver function assessment, we further aimed to compare our established APRI+ALBI model with models based on available liver function tests. Accordinlgy, we trained multivariable models utilizing ALICE, ICG-R15, ICG-PDR and FIB-4 respectively, as well as the same parameters used in the APRI+ALBI multivariable model (age, sex, tumor type, extent of resection) for their predictive potential for PHLF B+C. A detailed description of the different models can be found in Table 3. ROC curve analysis was calculated and AUC was compared between the models using bootstrap analysis (Fig. 4A-C). The APRI+ALBI model trained in the NSQIP cohort showed similar performance in the prediction of PHLF B+C when compared with the models based on the other liver function tests (Fig. 4A). Similar results were observed for PHLF C (Fig. 4B) as well as 90 day mortality (Fig. 4C).

As different tumor types are associated with different risk for development of PHLF B+C, we also evaluated predictive potential for PHLF B+C in different tumor subgroups. In direct comparison with the models based on other liver function tests the APRI+ALBI multivariable model showed similar predictive performance for PHLF B+C diagnosis in patients with

CRCLM, primary liver cancer and liver metastases from other malignancies (Supp. fig. 2, Supplemental Digital Content 1, http://links.lww.com/SLA/E919).

Design of a smartphone application towards clinical implementation of the APRI+ALBI PHLF B+C multivariable model

Using the APRI+ALBI score-based multivariable prediction model for PHLF B+C, a freely available smartphone-first application was designed (TELLAPRIALBI, https://tellaprialbi.howto.health). TELLAPRIALBI allows calculation of the APRI+ALBI score based multivariable prediction model upon input of the underlying parameters. Based on the APRI+ALBI score, age, sex, corresponding tumor subgroup and extent of resection patient specific PHLF B+C probability is identified (%).

Discussion

PHLF B+C remains the most common immediate cause of death after liver resection, with almost 50% of 90-day mortality after surgery related to PHLF.[1, 21] With no postOP treatment available, preOP risk stratification is critical. Aiming at moving towards personalized risk-assessment for patients undergoing hepatic resection, we calculated a multivariable prediction model for clinically significant PHLF B+C in 12,056 patients from the NSQIP database (model generation cohort). The APRI+ALBI multivariable model was then validated using an international cohort of 2,525 patients out of 10 different centers (validation cohort). Individual scores and dynamic liver function tests have often been evaluated for their ability to accurately predict PHLF, but rarely have they been directly compared. Therefore, in a sub-cohort of 620 patients, similar models were trained based on ICG-clearance, ALICE and FIB-4 respectively and compared to the APRI+ALBI multivariable model for PHLF B+C prediction. Despite the APRI+ALBI score being calculated using simple routine laboratory tests and the APRI+ALBI multivariable model being trained in a different much larger patient cohort (NSQIP cohort), the APRI+ALBI model showed comparable predictive potential for PHLF B+C, PHLF grade C and 90 day mortality. As ICG clearance is time consuming, more expensive and not universally available, we believe that these very robust results are critically relevant for preoperative PHLF prediction. We ultimately developed a smart phone application to allow for easy calculation of the APRI+ALBI multivariable model and clinically meaningful patient specific risk-assessment.

APRI and ALBI scores have both been associated with a variety of different liver pathologies. The APRI score was originally developed in the setting of chronic liver disease, as a non-invasive test for fibrosis and cirrhosis in hepatitis C patients.[22] Further, APRI has been shown to closely correlate with CALI. In particular, several studies have shown APRI to reflect SOS after oxaliplatin based chemotherapy regimens.[23, 24] The ALBI score was initially compared to the Child-Pugh score for the assessment of liver function in HCC patients, with similar results.[25] It shows a close correlation with fibrosis and cirrhosis in HCC patients.[26] Both scores have previously been evaluated for their ability to predict PHLF, demonstrating significant predictive potential of the individual parameters on their own.[27, 23, 28] Recently, in several studies, we compared the predictive potential of the combined APRI+ALBI score to APRI and ALBI alone for their predictive potential for PHLF or postOP mortality, documenting improved predictive potential of the combined score.[10, 14] This might be caused by broad detection of the multiple liver pathologies seen in patients undergoing hepatic resection. Development of PHLF has different causes, depending on the underlying liver disease and tumor type. PHLF risk is usually increased due to CALI in CRCLM patients after neoadjuvant chemotherapy.[5] HCC patients on the other hand are more likely to suffer from chronic liver disease caused by alcoholic steatohepatitis, non-alcoholic steatohepatitis or viral hepatitis.[29] The reason why APRI+ALBI could reflect chronic liver disease or liver injury, might be due to the parameters APRI+ALBI is comprised of, which provide a comprehensive evaluation on liver function. Hepatocyte demise is represented by AST, liver function is reflected in albumin and bilirubin and the inclusion of platelets mirrors the endocrine function of the liver as well as portal hypertension. It is important to note that the APRI+ALBI multivariable mode introduced in this study includes among other parameters tumor type and planned extent of resection. Both factors are known to significantly affect postoperative outcome, which could also be observed in our analyses. In combination with the holistic assessment of liver function via the APRI+ALBI score, this model was found to be suitable to predict postOP outcome for multiple patient subgroups, suggesting its relevance in a variety of different indications for hepatic resection.

While volumetric analyses are critical to avoid PHLF, postOP liver function recovery is also critically affected by underlying liver disease.[30] While we rely on crude and poorly validated cutoffs for volumetry (e.g., 20-25% in healthy liver, 30% after chemotherapy and 40% for cirrhotic livers), we underestimate the relevance of quantifying hepatic function.[9, 31] We believe that our analyses provide a very strong basis to move forward with integrative models

also including volumetry, possibly enabling a patient specific assessment of the required future liver remnant volume.

Previous research has assessed a multitude of different metrics for their ability to reliably predict PHLF. Within our analyses, 4 multivariable models based on established preOP liver function tests (ICG-clearance, the ALICE score and the FIB-4 index) were compared with an APRI+ALBI model. All models were found to have a similar predictive potential for PHLF B+C, PHLF C and 90 day mortality as compared to the APRI+ALBI model. Equal performance of the APRI+ALBI model, in comparison with the models based on the other liver function tests, was especially remarkable, as the APRI+ALBI model was trained in another cohort (N = 12,056), eliminating the risk of overfitting.

ICG-clearance, a dynamic liver function test, has shown association with PHLF and postOP mortality, pre- and intraoperatively and for many represents the gold standard for liver function testing prior to hepatic resection.[11, 32, 33] Several studies could show a direct correlation of ICG-clearance with portal hypertension and cirrhosis.[34, 35] Importantly, liver perfusion critically affects ICG-clearance, as ICG-R15 has been shown to be directly influenced by changes in portal flow, as well as cholestasis, making its assessment challenging in patients with preOP hyperbilirubinemia and changes in portal venous flow.[36] In regards to CALI, data on the effects of neoadjuvant chemotherapy on ICG-clearance is limited.[37] This lack of association of ICG-clearance with CALI might in part explain why the APRI+ALBI score appeared to have a higher predictive potential than ICG-clearance for PHLF B+C in our analyses. In this context we do believe that the association of APRI+ALBI with a wide range of different etiologies of liver disease, represents one of the key elements for its excellent predictive potential for PHLF and postOP 90 day mortality. Further, APRI+ALBI score is available at a fraction of the costs of ICG clearance measurement, exhibits none of its invasive features and eliminates the risk of allergic reaction to ICG dye components.

A combination of ICG-clearance and albumin, the so called ALICE grade, has been evaluated for its associations with short term postOP outcome for patients with HCC, CCA, CRCLM and hepatic alveolar echinococcosis.[12, 17, 38, 39] However, studies directly comparing ICG-clearance with ALICE grade are rare and limited in sample size.[39] The inclusion of ICG-R15 in the formula for calculation of the ALICE grade introduces the limitations of ICG-clearance,

described above. While the predictive ability is clearly improved through the inclusion of albumin, as seen in an overall increase of the AUCs in our ROC curve analysis, calculation of the ALICE grade remains an invasive test and is more expensive than a parameter solely based on routine laboratory values, like the APRI+ALBI score.

FIB-4 index was originally developed and validated as a non-invasive test for significant fibrosis in human immunodeficiency or hepatitis C virus related chronic liver disease.[18, 40] In the current literature, only a few studies have examined FIB-4 in a preOP setting, mainly in HCC patients.[41] Very rarely have studies assessed the outcome after liver surgery in CRCLM patients depending on FIB-4 scores.[42] Predictive potential of FIB-4 has been compared to Child-Pugh score and APRI, showing improved or comparable results.[41, 43] When compared with APRI+ALBI, FIB-4 appeared to show similar predictive potential for PHLF B+C. It is however worth mentioning that predictive performance of the FIB-4 multivariable model and respectively the performance of the models utilizing ALICE and ICG clearance is limited by the size of their training cohort. Also, the APRI+ALBI model was validated using an independent patient cohort. To accurately assess the predictive performance of the models based on ALICE, ICG clearance and the FIB-4 index, going beyond a descriptive comparison of the models, further validation is needed.

It was previously documented that risk assessment for PHLF strikingly depends on underlying tumor types. For example, a certain APRI+ALBI value might be associated with moderate risk for CRCLM patients but with high risk in HCC patients. This phenomenon can be observed for basically every preOP liver function test and poses a very relevant challenge for the application of cutoffs in clinical routine. To address this issue, we developed a novel multivariable model for the prediction of PHLF B+C probability, with different tumor subgroups, as well as the APRI+ALBI score, age, sex and planned type of resection included as variables. This tool, built on more than 12,000 patients and now integrated in the TELLAPRIALBI smart phone application, will allow for the first time to perform risk-assessment of PHLF B+C in different patient subgroups. This multivariable model is based solely on routinely available blood parameters and basic patient characteristics available prior to every planned liver resection.

Interestingly, observed PHLF risk was higher in the international multicenter cohort, when compared to predicted risk in the NSQIP database. PHLF incidence in NSQIP analyses has

been reported at about 5%, in comparison to the reported incidence of approximately 10% in prospective international multicenter studies.[14, 44] PHLF incidence increases in low-volume centers and major liver resections.[45, 46] Our study included high- as well as low-volume centers, but more importantly, approximately 70% of the patients underwent major liver resection (Table 2), 30% more than in the NSQIP database cohort (Table 1). This could in part explain a difference in observed and predicted PHLF incidence in the logistic regression-based prediction model (Fig. 1A). However, while significantly reduced, this increase could also be observed when we only assessed patients undergoing major resection, suggesting that there are clearly other factors involved for this increase in PHLF in the international multicenter cohort. Obviously, limited granularity of nationwide databases (NSQIP cohort) as compared to prospectively maintained databases (international multicenter cohort) might be an important aspect accounting for these differences as well.

In conclusion, we were able to document and validate a high predictive potential of a novel APRI+ALBI score based preOP multivariable model for multiple postOP outcome measures, particularly clinically significant PHLF and 90 day mortality, in a cohort of > 14,000 patients. Importantly, this routine laboratory parameter-based score showed equal performance to other multivariable models based on well-established, costly and time-consuming tests, such as ICG-clearance or ALICE grade in the prediction of clinically significant PHLF. We also created a freely available smartphone application to calculate the multivariable model and patient specific individual PHLF B+C probability.

Acknowledgements

J.S. collected and analyzed the data, prepared the figures and wrote the manuscript; S.K., L.G., R.B., A.M.-L., E.B., S.G., E.B., M.A., D.P., D.A., M.N., A.E.K., B.R., G.O., Y.H., F.X.H., Y.D. and J.W. collected and analyzed the data; J.S., C.T., S.W., R.S., M.T., M.K., D.N., S.C., G.B., N.R., K.H., S.G., A.A., T.G. provided resources; H.H. analyzed the data, prepared the figures and wrote the manuscript; P.S. conceived the study, analyzed and interpreted the data, provided resources and wrote the manuscript.

Figure 1: A multivariable model based on the APRI+ALBI score, aspartate to aminotransferase ratio (APRI) combined with albumin-bilirubin grade (ALBI), for the prediction of posthepatectomy liver failure (PHLF) grade B and C (B+C). The model is tested in the entire cohort (A) and tested in only patients undergoing major resection (B), documenting the non-linear increase of PHLF B+C with rising APRI+ALBI score. The model was calculated using the NSQIP cohort (predicted). Predictive performance of the model was validated using the validation cohort (international multicenter cohort, observed). APRI+ALBI score is given in deciles on the x-axis. PHLF B+C risk is given in % on the y-axis (A, B). To visualize the probability for PHLF B+C development for the complete cohort, a major resection subgroup, a colorectal cancer with liver metastases (CRCLM), primary liver cancer and other malignancies with metastases in the liver subgroup and respectively subgroups for patients undergoing major liver resection in the different tumor subroups a bubble plot was generated (C). APRI+ALBI score is given in deciles and bubble size at each decile and for each patient groups indicates number of patients. PHLF B+C probability is reflected in bubble colour and is calculated by the APRI+ALBI based multivariable model. Associated PHLF B+C probability for each bubble is explained in the figure legend (C).

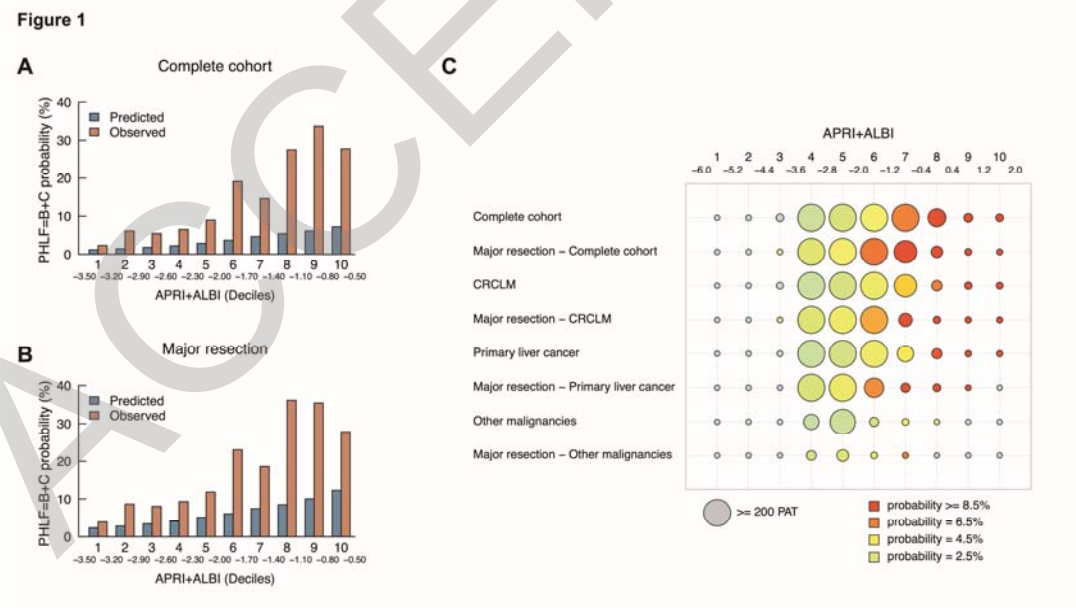


Figure 2: Evaluation of the performance of an APRI+ALBI, aspartate to aminotransferase ratio (APRI) combined with albumin-bilirubin grade (ALBI), based multivariable model (APRI+ALBI, age, sex, tumor type, extent of resection) for posthepatectomy liver failure grade B and C (PHLF B+C) prediction. Performance of the model is evaluated using receiver operating characteristic (ROC) curve analysis and area under the curve (AUC) (A). Validation of the model is done via ROC curve analysis and AUC calculation as well (B).

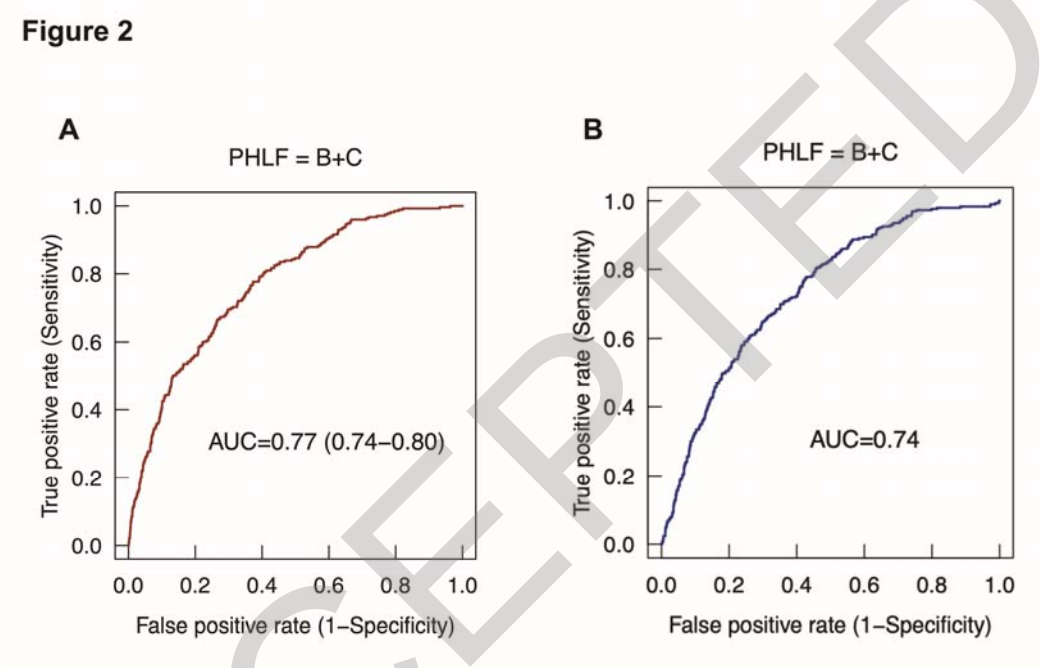


Figure 3: Descriptive comparison of the performance of several univariate models for prediction of posthepatectomy liver failure grade B and C (PHLF B+C) (A), PHLF C (B) and 90 day mortality (C). The different models are calculated using APRI+ALBI, aspartate to aminotransferase ratio (APRI) combined with albumin-bilirubin grade (ALBI), indocyanine green clearance (ICG) retention 15 minutes after administration (ICG-R15) and plasma dissaperance rate (ICG-PDR), albumin-indocyanine green evaluation (ALICE) grade and fibrosis-4 (FIB-4) index. Models are calculated out of 620 patients from an international multicenter cohort. Performance of each model is indicated as area under curve (AUC) from receiver operating characteristic (ROC) curve analysis. 95%-confidence intervals and tests between the AUC of APRI+ALBI versus the AUC of each of the other parameters have been performed by a bootstrap resampling analysis (lower panels).

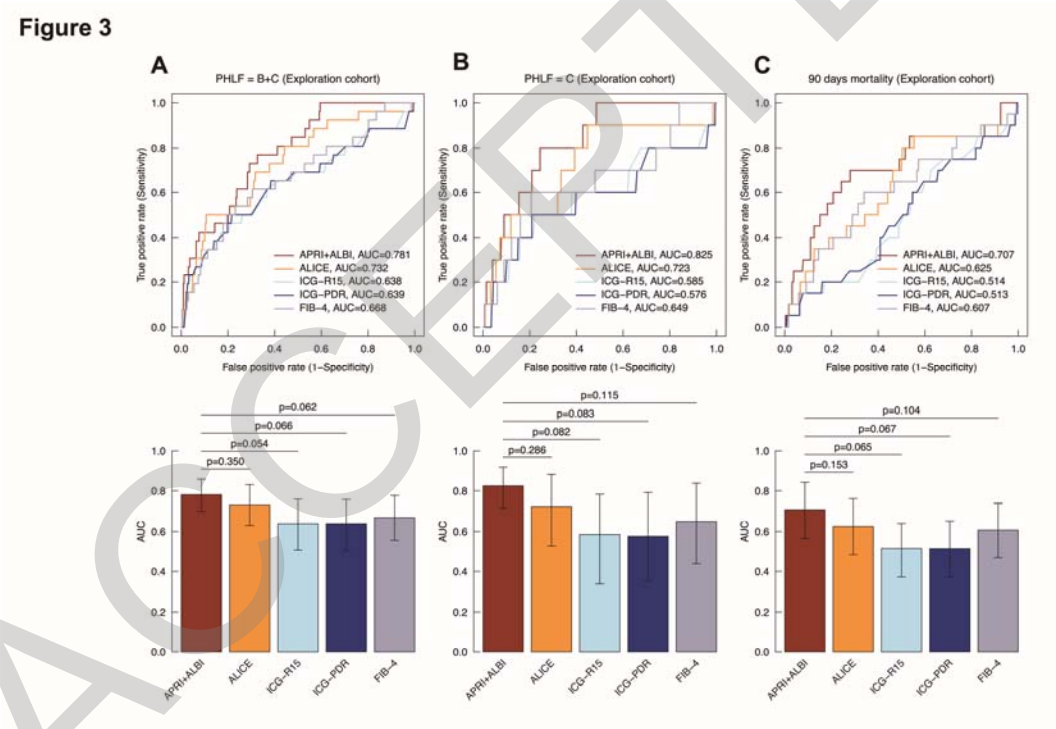
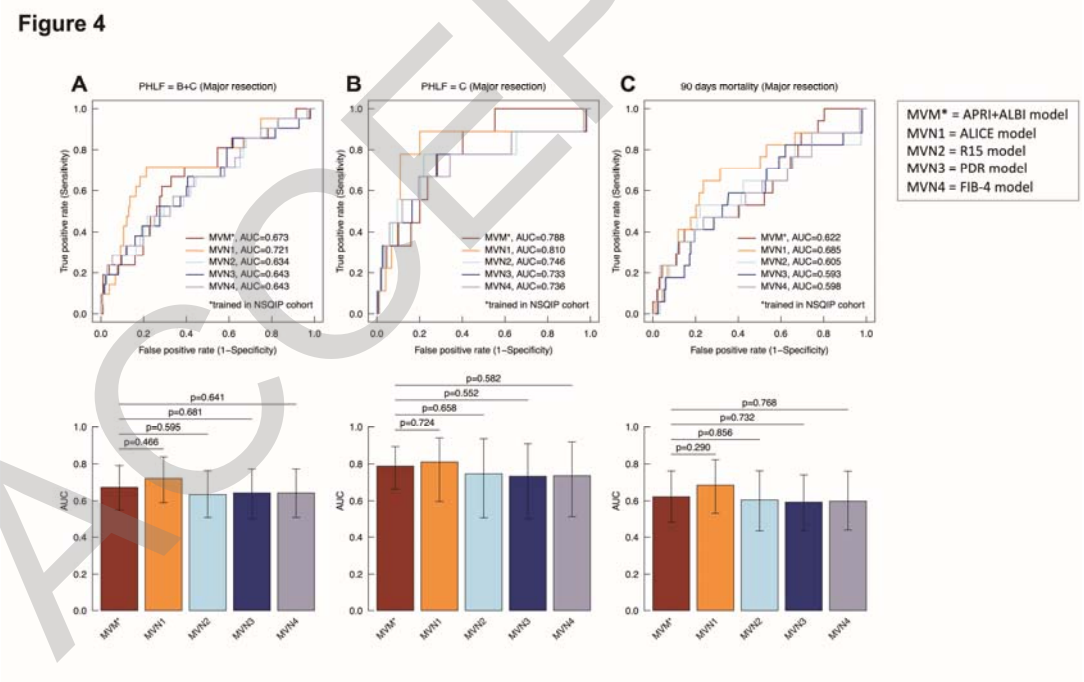


Figure 4: Comparison of different multivariable models for the prediction of posthepatectomy liver failure grade B and C (PHLF B+C), PHLF grade C (PHLF C) and 90 day mortality in 620 patients out of an international multicenter cohort. Models included in the comparison are an APRI+ALBI score based multivariable model trained in the National Surgery Quality Improvement Program (NSQIP) cohort, as well as models trained in 620 patients out of 10 international centers and based on indocyanine green clearance (ICG) retention 15 minutes after administration (ICG-R15) and plasma dissaperance rate (ICG-PDR), albumin-indocyanine green evaluation (ALICE) grade and fibrosis-4 (FIB-4) index respectively. All models include age, sex, tumor type and extent of resection as variables, as well as one of the respective liver function tests (APRI+ALBI, ICG-R15, ICG-PDR, ALICE, FIB-4). Performance of each model is indicated as area under curve (AUC) from receiver operating characteristics (ROC). 95%-confidence intervals and tests between the AUC of APRI+ALBI versus the AUC of each of the other parameters have been performed by a bootstrap resampling analysis (lower panels).



References

1. Gilg S, Sandström P, Rizell M, et al. The impact of post-hepatectomy liver failure on mortality: a population-based study. *Scand J Gastroenterol* 2018; 53(10-11):1335-1339.
2. Gruenberger T, Bridgewater J, Chau I, et al. Bevacizumab plus mFOLFOX-6 or FOLFOXIRI in patients with initially unresectable liver metastases from colorectal cancer: the OLIVIA multinational randomised phase II trial. *Ann Oncol* 2015; 26(4):702-708.
3. Nordlinger B, Sorbye H, Glimelius B, et al. Perioperative FOLFOX4 chemotherapy and surgery versus surgery alone for resectable liver metastases from colorectal cancer (EORTC 40983): long-term results of a randomised, controlled, phase 3 trial. *Lancet Oncol* 2013; 14(12):1208-15.
4. Nordlinger B, Sorbye H, Glimelius B, et al. Perioperative chemotherapy with FOLFOX4 and surgery versus surgery alone for resectable liver metastases from colorectal cancer (EORTC Intergroup trial 40983): a randomised controlled trial. *Lancet* 2008; 371(9617):1007-16.
5. Zorzi D, Laurent A, Pawlik TM, et al. Chemotherapy-associated hepatotoxicity and surgery for colorectal liver metastases. *Br J Surg* 2007; 94(3):274-86.
6. Vauthey JN, Pawlik TM, Ribero D, et al. Chemotherapy regimen predicts steatohepatitis and an increase in 90-day mortality after surgery for hepatic colorectal metastases. *J Clin Oncol* 2006; 24(13):2065-72.
7. Koh YX, Tan HJ, Liew YX, et al. Liver Resection for Nonalcoholic Fatty Liver Disease-Associated Hepatocellular Carcinoma. *J Am Coll Surg* 2019; 229(5):467-478.e1.
8. Reddy SK, Marsh JW, Varley PR, et al. Underlying steatohepatitis, but not simple hepatic steatosis, increases morbidity after liver resection: a case-control study. *Hepatology* 2012; 56(6):2221-30.
9. 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(10):1331-1347.
10. Pereyra D, Rumpf B, Ammann M, et al. The Combination of APRI and ALBI Facilitates Preoperative Risk Stratification for Patients Undergoing Liver Surgery After Neoadjuvant Chemotherapy. *Ann Surg Oncol* 2019; 26(3):791-799.
11. Schwarz C, Plass I, Fitschek F, et al. The value of indocyanine green clearance assessment to predict postoperative liver dysfunction in patients undergoing liver resection. *Scientific Reports* 2019; 9(1):8421.
12. Miyazaki Y, Kokudo T, Amikura K, et al. Albumin-Indocyanine Green Evaluation Grading System Predicts Post-Hepatectomy Liver Failure for Biliary Tract Cancer. *Dig Surg* 2019; 36(1):13-19.
13. 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.
14. Starlinger P, Ubl DS, Hackl H, et al. Combined APRI/ALBI score to predict mortality after hepatic resection. *BJS Open* 2021; 5(1).
15. Strasberg SM. Nomenclature of hepatic anatomy and resections: a review of the Brisbane 2000 system. *J Hepatobiliary Pancreat Surg* 2005; 12(5):351-5.
16. Krieger PM, Tamandl D, Herberger B, et al. Evaluation of chemotherapy-associated liver injury in patients with colorectal cancer liver metastases using indocyanine green clearance testing. *Ann Surg Oncol* 2011; 18(6):1644-50.

17. Kokudo T, Hasegawa K, Amikura K, et al. Assessment of Preoperative Liver Function in Patients with Hepatocellular Carcinoma - The Albumin-Indocyanine Green Evaluation (ALICE) Grade. *PLoS One* 2016; 11(7):e0159530.
18. Vallet-Pichard A, Mallet V, Nalpas B, et al. FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. *Hepatology* 2007; 46(1):32-6.
19. Dindo D, Demartines N, Clavien P-A. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Annals of surgery* 2004; 240(2):205-213.
20. Rahbari NN, Garden OJ, Padbury R, et al. Posthepatectomy liver failure: a definition and grading by the International Study Group of Liver Surgery (ISGLS). *Surgery* 2011; 149(5):713-24.
21. Sparrelid E, Olthof PB, Dasari BVM, et al. Current evidence on posthepatectomy liver failure: comprehensive review. *BJS Open* 2022; 6(6).
22. Wai C-T, Greenson JK, Fontana RJ, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 2003; 38(2):518-526.
23. Amptoulach S, Gross G, Sturesson C, et al. Preoperative Aspartate Aminotransferase-to-Platelet Ratio Index Predicts Perioperative Liver-Related Complications Following Liver Resection for Colorectal Cancer Metastases. *Scandinavian Journal of Surgery* 2017; 106(4):311-317.
24. Hubert C, Sempoux C, Humblet Y, et al. Sinusoidal obstruction syndrome (SOS) related to chemotherapy for colorectal liver metastases: factors predictive of severe SOS lesions and protective effect of bevacizumab. *HPB (Oxford)* 2013; 15(11):858-64.
25. Johnson PJ, Berhane S, Kagebayashi C, 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):550-8.
26. Dong Z-R, Zou J, Sun D, et al. Preoperative Albumin-Bilirubin Score for Postoperative Solitary Hepatocellular Carcinoma within the Milan Criteria and Child-Pugh A Cirrhosis. *Journal of Cancer* 2017; 8(18):3862-3867.
27. Demirtas CO, D'Alessio A, Rimassa L, et al. ALBI grade: Evidence for an improved model for liver functional estimation in patients with hepatocellular carcinoma. *JHEP Reports* 2021; 3(5).
28. Cheng J, Zhao P, Liu J, et al. Preoperative aspartate aminotransferase-to-platelet ratio index (APRI) is a predictor on postoperative outcomes of hepatocellular carcinoma. *Medicine (Baltimore)* 2016; 95(48):e5486.
29. Estes C, Razavi H, Loomba R, et al. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. *Hepatology* 2018; 67(1):123-133.
30. Yokoyama Y, Nishio H, Ebata T, et al. Value of indocyanine green clearance of the future liver remnant in predicting outcome after resection for biliary cancer. *Br J Surg* 2010; 97(8):1260-8.
31. Guglielmi A, Ruzzenente A, Conci S, et al. How Much Remnant Is Enough in Liver Resection? *Digestive Surgery* 2012; 29(1):6-17.
32. Sato N, Kenjo A, Suzushino S, et al. Predicting Post-Hepatectomy Liver Failure Using Intra-Operative Measurement of Indocyanine Green Clearance in Anatomical Hepatectomy. *World J Surg* 2021; 45(12):3660-3667.
33. de Liguori Carino N, O'Reilly DA, Dajani K, et al. Perioperative use of the LiMON method of indocyanine green elimination measurement for the prediction and early

detection of post-hepatectomy liver failure. *European Journal of Surgical Oncology (EJSO)* 2009; 35(9):957-962.
34. Pind M-LL, Bendtsen F, Kallemose T, et al. Indocyanine green retention test (ICG-r15) as a noninvasive predictor of portal hypertension in patients with different severity of cirrhosis. *European Journal of Gastroenterology & Hepatology* 2016; 28(8).
35. Møller S, la Cour Sibbesen E, Madsen JL, et al. Indocyanine green retention test in cirrhosis and portal hypertension: Accuracy and relation to severity of disease. *J Gastroenterol Hepatol* 2019; 34(6):1093-1099.
36. Janssen MW, Druckrey-Fiskaaen KT, Omidi L, et al. Indocyanine green R15 ratio depends directly on liver perfusion flow rate. *J Hepatobiliary Pancreat Sci* 2010; 17(2):180-5.
37. Wang LJ, Yan XL, Li J, et al. Indocyanine Green Clearance Test for the Preoperative Assessment of Chemotherapy-Related Hepatic Injury in Patients with Colorectal Liver Metastasis. *Cancer Manag Res* 2020; 12:3237-3245.
38. Kokudo T, Shirata C, Toda T, et al. Assessment of preoperative liver function in patients undergoing liver resection – The ALICE grade. *HPB* 2018; 20:S204.
39. Liang Y, Zhang Z, Dai Z, et al. Comparative study of indocyanine green (ICG)-R15 and Albumin-Indocyanine Green Evaluation (ALICE) grading system in the prediction of posthepatectomy liver failure and postoperative mortality in patients with hepatic alveolar echinococcosis. *BMC Gastroenterol* 2022; 22(1):297.
40. Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 2006; 43(6):1317-25.
41. Zhang ZQ, Yang B, Zou H, et al. ALBI/ST ratio versus FIB-4 and APRI as a predictor of posthepatectomy liver failure in hepatocellular carcinoma patients. *Medicine (Baltimore)* 2019; 98(15):e15168.
42. Akiyama T, Miyamoto Y, Imai K, et al. Fibrosis-4 Index, a Noninvasive Fibrosis Marker, Predicts Survival Outcomes After Hepatectomy for Colorectal Cancer Liver Metastases. *Ann Surg Oncol* 2020; 27(9):3534-3541.
43. Zhou P, Chen B, Miao XY, et al. Comparison of FIB-4 Index and Child-Pugh Score in Predicting the Outcome of Hepatic Resection for Hepatocellular Carcinoma. *J Gastrointest Surg* 2020; 24(4):823-831.
44. Sultana A, Brooke-Smith M, Ullah S, et al. Prospective evaluation of the International Study Group for Liver Surgery definition of post hepatectomy liver failure after liver resection: an international multicentre study. *HPB (Oxford)* 2018; 20(5):462-469.
45. Gilg S, Sparrelid E, Isaksson B, et al. Mortality-related risk factors and long-term survival after 4460 liver resections in Sweden-a population-based study. *Langenbecks Arch Surg* 2017; 402(1):105-113.
46. Filmann N, Walter D, Schadde E, et al. Mortality after liver surgery in Germany. *Br J Surg* 2019; 106(11):1523-1529.

Table 1: Patient demographics for the National Surgical Quality Improvement Program (NSQIP) cohort. * = the NSQIP database lacks granularity regarding specific grades of postoperative morbidity. N = number; IQR = interquartile range; CRCLM = colorectral cancer with liver metastases; HCC = hepatocellular carcinoma; CCA = cholangiocellular carcinoma; PHLF = posthepatectomy liver failure; APRI = aspartate aminotransferase to platelet ratio; ALBI = albumin bilirubin grade; SB = serum bilirubin; AST = aspartate aminotransferase.

NSQIP Cohort	
	Entire Cohort (N = 12,056)
Parameter	**Median (IQR) / N (%)**
Age (years)	60 (50-68)
Sex	
Female	5997 (49.7)
Male	6059 (50.3)
Tumor entity	
CRCLM	5,227 (43.4)
HCC	2,271 (18.8)
CCA	1,343 (11.1)
Benign	1,269 (10.5)
Other	1,666 (13.8)
Hepatic Resection	
Minor	7,377 (61.2)
Major	4,679 (38.8)
Morbidity*	
No morbidity	9,924 (82.3)
Morbidity	2,131 (17.7)
Posthepatectomy liver failure	
no PHLF	11,515 (95.5)
PHLF total	540 (4.5)
ISGLS A	228 (1.9)
ISGLS B	178 (1.5)
ISGLS C	134 (1.1)
Preoperative Parameters	
APRI+ALBI score	-4.17 (-4.46- -3.80)
SB in mg/dl	8.55 (5.1-12.0)
AST in U/l	26 (20-37)
Albumin in g/l	4.10 (3.7-4.3)
Platelets in G/l	219 (174-272)

Table 2: Patient demographics for the international multicenter cohort, the validation cohort (N = 2525), and patient demographics for 620 patients out of the validation cohort used for comparison of different liver function tests. N = number; IQR = interquartile range; PHLF = posthepatectomy liver failure; CRCLM = colorectral cancer liver metastases; HCC = hepatocellular carcinoma; CCA = cholangiocellular carcinoma; ISGLS = international study group of liver surgery; APRI = aspartate aminotransferase to platelet ratio; ALBI = albumin bilirubin grade; ICG = indocyanine green; ICG-PDR = plasma disappearance rate; ICG-R15 = retention after 15 minutes; ALICE = albumin-indocyanine green evaluation; FIB-4 = fibrosis-4; SB = serum bilirubin; PT = prothrombin time; AP = alcaline phosphatase; AST = aspartate aminotransferase; ALT = alanine aminotransferase; GGT = gamma-glutamyl transferase; CHE = cholinesterase.

Validation Cohort						
Parameters	**Validation Cohort (N = 2.525)**			**Validation Cohort, direct comparison of liver funciton tests (N = 620)**		
	Entire Cohort (N = 2525)	**no PHLF (N = 2226)**	**PHLF B-C (N = 229)**	**Entire Cohort (N = 620)**	**no PHLF B+C (N = 584)**	**PHLF B+C (N = 36)**
	Median (IQR) / N (%)	**Median (IQR) / N (%)**	**Median (IQR) / N (%)**	**Median (IQR) / N (%)**	**Median (IQR) / N (%)**	**Median (IQR) / N (%)**
Age (years)	64 (56-72)	64 (55-572)	65 (58-72)	63.05 (55.41 - 70.47)	62.93 (55.11 - 70.15)	66.66 (57.56 - 74.75)
Sex						
Female	1056 (41.9)	957 (43)	99 (33.1)	258 (41.6)	247 (42.3)	11 (30.6)
Male	1469 (58.2)	1269 (57.0)	200 (66.9)	362 (58.4)	337 (57.7)	25 (69.4)
Tumor entity						
CRCLM	1251 (49.6)	1147 (51.4)	104 (34.8)	402 (64.8)	383 (65.6)	19 (52.8)
HCC	340 (13.5)	295 (13.2)	45 (15.1)	53 (8.5)	47 (8)	6 (16.7)
CCA	450 (17.8)	340 (15.4)	110 (36.8)	43 (6.9)	37 (6.3)	6 (16.7)

benign tumors	197 (7.8)	184 (8.2)	16 (5.4)	28 (4.5)	27 (4.6)	1 (2.8)
other malignancies	284 (11.3)	259 (11.7)	24 (8.0)	94 (15.2)	90 (15.4)	4 (11.1)
Hepatic Resection						
Minor	711 (27.6)	687 (30.5)	24 (7.7)	213 (34.4)	208 (35.6)	5 (13.9)
Major	1814 (72.4)	1539 (68.5)	275 (92.3)	390 (62.9)	359 (61.5)	31 (86.1)
Histology						
no Fibrosis				221 (35.6)	204 (34.9)	17 (47.2)
Fibrosis grade I				280 (45.2)	268 (45.9)	12 (33.3)
Fibrosis grade II				63 (10.2)	61 (10.4)	2 (5.6)
Fibrosis grade III				15 (2.4)	14 (2.4)	1 (2.8)
Fibrosis grade IV				28 (4.5)	25 (4.3)	3 (8.3)
Morbidity						
No morbidity	1025 (40.3)	1007 (45.2)	15 (5.1)	326 (52.5)	326 (55.8)	0 (0)
Morbidity any	1500 (59.6)	1219 (54.8)	284 (94.9)			
I				65 (10.5)	62 (10.6)	3 (8.3)
II				80 (12.9)	73 (12.5)	7 (19.4)
IIIa				54 (8.7)	49 (8.4)	5 (13.9)
IIIb				60 (9.7)	51 (8.7)	9 (25)
IVa				6 (1.0)	3 (0.5)	3 (8.3)
IVb				0 (0.0)	0 (0)	0 (0)
V				29 (4.7)	20 (3.4)	9 (25)
Severe morbidity						

No severe morbidity	1786 (71.0)	1709 (76.7)	77 (25.3)	471 (75.9)	461 (78.9)	10 (27.8)
Severe morbidity	739 (29.3)	517 (23.3)	222 (74.7)	149 (24.1)	123 (21.1)	26 (72.2)
90 day mortality						
no 90 day mortality	2410 (94.9)	2167 (97.3)	235 (78.4)	591 (95.3)	564 (96.6)	27 (75)
90 day mortality	121 (5.1)	59 (2.7)	64 (21.6)	29 (4.7)	20 (3.4)	9 (25)
Posthepatectomy liver failure						
no PHLF	2034 (80.6)			553 (89.2)		
PHLF total	483 (19.1)			66 (10.6)		
ISGLS A	163 (7.4)			31 (5.0)		
ISGLS B	118 (5.4)			22 (3.5)		
ISGLS C	136 (6.2)			13 (2.1)		
Preoperative Parameters						
APRI+ALBI score	-2.29 (-2.65 - -1.70)	-2.24 (-2.61 - -1.67)	-1.60 (-2.21 - -0.86)	-2.57 (-2.82 - -2.32)	-2.59 (-2.85 - -2.34)	-2.25 (-2.42 - -1.53)
ICG-PDR in %/min				20.5 (17 - 24.6)	20.75 (17.23 - 24.78)	17.85 (14.20 - 22.08)
ICG-R15 in %				4.7 (2.35 - 7.95)	4.5 (2.1 - 7.6)	7.45 (3.33 - 15.95)
ALICE grade				-2.59 (-2.85 - -2.33)	-2.62 (-2.86 - -2.35)	-2.13 (--2.53 - -1.91)
Fib-4 index				1.67 (1.11 - 2.47)	1.65 (1.11 - 2.38)	2.56 (1.41 - 3.53)
SB in µmol/L	8.60 (6.00 -	8.60 (6.00	11.00	10.26	10.09 (7.70	11.80

	12.83)	- 12.00)	(7.00 - 18.72)	(7.70 - 15.22)	– 14.88)	(9.58 – 18.81)
PT in %	103 (97 - 110)	102 (97 - 110)	103 (97 - 109)	105 (91.25 - 120)	105 (92 - 120)	95 (80 - 119)
AP in U/l	100 (75 - 142)	96 (75 - 133)	135 (90 - 242)	94 (72 - 123)	93 (71 - 121.5)	115 (87.5 - 154.75)
AST in U/l	30 (23 - 43)	30 (23 - 43)	39 (28 - 65)	30 (23 - 40)	30 (23 - 39)	42 (27 - 649
ALT in U/l	29 (19 - 46)	28 (19 - 43)	35 (23 - 66)	27 (19 - 41.25)	27 (19 - 40.25)	31.5 (21 - 61.75)
GGT in U/l	67 (33.00 - 150.00)	60 (32 - 129)	144 (69 - 304)	56 (32 - 108)	53 (31 - 104)	121 (70 - 215)
Albumin in g/l	39 (34.90 - 43.10)	39.30 (35.00 - 43.20)	38.00 (33.10 - 42.30)	42 (39.60 - 44.20)	42.35 (39.9 - 44.3)	38.55 (35.65 - 41.48)
Platelets in G/l	236 (185 - 292)	237 (186 - 292)	235 (174 - 297)	226 (170 - 275)	226 (171 - 278)	200 (147 - 265)

Table 3: Multivariable Analysis calculated in the NSQIP cohort and comparison of different multivariable models in 620 patients out of the international multicenter cohort. The APRI+ALBI multivariable model was calculated using all available parameters from the NSQIP cohort. The model was created using backwards feature elimination, parameters proving non-significant were excluded without compromising quality of the model. Brier score for the APRI+ALBI multivariable model was 0.025. The APRI+ALBI multivariable model, trained in the NSQIP cohort was then compared to multivariable models tested in the international multicenter cohort (N = 620). These models included the same paramters as the APRI+ALBI multivariable model (age, sex, tumor type, extent of resection), as well as other liver function tests (ALICE, ICG-R15, ICG-PDR, FIB-4). For a descriptive analysis models of established liver function tests alone were compared separately. Tumor type reference level = CRCLM. OR = odds ratio, CI = confidence interval, AUC = area under receiver operating characteristic curve, 95%-CI¶ = confidence interval by bootstrap analysis with 2000 iterations, AIC = Akaike information criterion, Pseudo-R2‡ = Pseudo-R2 based on Nagelkerke (Cragg and Uhler), MVM, multivariable model; UVM, univariate model, † = Final model trained on the NSQIP cohort, PHLF B+C = posthepatectomy liver failure grade B and C, CRCLM = colorectal carcinoma with liver metastases;, °p<0.1, *p<0.05, **p<0.01, ***p<0.001

NSQIP Cohort

Name	**Outcome**	**Explaining variable(s)**	**AUC**	**95%-CI¶**	**AIC**	**Pseudo-R^2 ‡**
MVM	PHLF B+C	Sex***, Age*, Resection***, Tumor type***, APRI+ALBI***	0.771	0.743-0.796	2,278.9	0.114
			Coefficient	***P***	**OR**	**OR 95%-CI**
		Intercept	-4.3256	<0.0001	0.01	0.00-63.6
		Sex (Male)	0.4415	0.00081	1.56	0.65-3.69
		Age (Years)	0.0116	0.027	1.01	0.99-1.04
		Tumor type (Primary liver cancer)	0.0074	0.96	1.01	0.99-1.02
		Tumor type (Benign)	0.7602	<0.0001	2.14	0.48-9.49
		Tumor type (Other malignancies)	-0.2981	0.52	0.74	0.41-1.33
		Resection (Major)	1.3518	<0.0001	3.86	0.27-54.6
		APRI+ALBI	0.5037	<0.0001	1.66	0.62-4.44
Name	**Outcome**	**Explaining variable(s)**	**AUC**	**95%-CI¶**	**AIC**	**Pseudo-R^2 ‡**
UVM	PHLF B+C	APRI+ALBI***	0.698	0.666-0.730	2,426.2	0.044
			Coefficient	***P***	**OR**	**OR 95%-CI**
		APRI+ALBI	0.5536	<0.0001	1.74	0.59-5.15

International Multicenter Cohort

Name	**Outcome**	**Explaining variable(s)**	**AUC**	**95%-CI¶**	**AIC**	**Pseudo-R^2 ‡**

MVM[†]	PHLF B+C	Sex, Age, Resection, Tumor type, APRI+ALBI	0.725	0.634-0.810	--	--
MVN1	PHLF B+C	Sex, Age, Resection, Tumor type[o], ALICE***	0.790	0.692-0.873	176.04	0.171
MVN2	PHLF B+C	Sex, Age, Resection, Tumor type[o], ICG-R15*	0.744	0.651-0.826	183.75	0.122
MVN3	PHLF B+C	Sex, Age, Resection, Tumor type*, ICG-PDR	0.726	0.627-0.815	185.73	0.109
MVN4	PHLF B+C	Sex, Age, Resection, Tumor type[o], FIB-4*	0.723	0.627-0.813	182.34	0.131
APRI+ ALBI	PHLF B+C	APRI+ALBI**	0.781	0.695-0.856	181.64	0.056
ALICE	PHLF B+C	ALICE***	0.732	0.623-0.831	173.08	0.113
ICG-R15	PHLF B+C	ICG-R15*	0.638	0.501-0.759	184.08	0.039
ICG-PDR	PHLF B+C	ICG-PDR*	0.639	0.508-0.762	185.41	0.030
FIB-4	PHLF B+C	FIB-4**	0.668	0.544-0.781	181.70	0.055