LAPS and the ALPPS was 85.7% (6/7) and 78.6% (11/14). The incidence of major complications was 36.4% (4/11) of the ALPPS group and 50.0% (3/6) of the LAPS group after the 2 stages operation. One patient died of the ALPPS group. Additionally, the median increase in FLR, median operative time and blood loss during the two stages of the LAPS were similar to those subjected to ALPPS.

Conclusions LAPS has a potential advantage in eliminating major complications of PHLF associated with classic ALPPS. LAPS may achieve the same effect of promoting significant growth of the FLR in patients with HBV-related HCC, albeit at the cost of longer interval time.

**IDDF2020-ABS-0078**

**IMMUNOSCORE CLASSIFICATION FROM HEPATOCELLULAR CARCINOMA HISTOPATHOLOGY IMAGES USING DEEP LEARNING: A PRELIMINARY STUDY**

**Background**

Immunotherapy is a recent advance for the treatment of hepatocellular carcinoma (HCC). Immunoscore assessment plays a critical role in precision immunotherapy and can predict prognosis in patients with HCC. This study aims to develop a deep-learning model to automated analyze histopathology images for classification of immunoscore (CD3 or CD8, 0–2 vs. 3–4) in HCC.

**Methods**

We trained a patch-based deep convolutional neural network (Resnet-18) on whole-slide images to automatically classify immunoscore into 0–2 or 3–4. The data were randomly split into a training and testing dataset. The performance was first estimated on the training dataset with ninefold cross-validation and then further validated on the testing dataset. Cross-entropy was used as a model-optimized loss function and the accuracy as well as the area under the curve (AUC) were calculated for the identification values. Heatmaps were also generated by our model to visualize the regions the most associated with the classification.

**Results**

We included 28 images from a study cohort of 28 HCC patients for training (18 images) and testing (10 images) the model. After iterative training, an optimized architecture...
achieved an AUC of 0.71 was used as our final model. For validation on the testing dataset, the model yielded an accuracy of 90% and AUC of 0.93 (95% CI: 0.76 to 1.00) while the percentage of patches positively classified, and outperforms average of the probabilities of the corresponding patches (accuracy 70%; AUC 0.79, 95% CI: 0.50 to 1.00) using the same optimal threshold of 0.33. The heatmaps show that almost all of patches are highly identified to show the regions of immunoscore (figure 1) A. Immunoscore of 3–4 [positive], B. Immunoscore of 0–2 [negative].

Conclusions The automated deep-learning model achieved good performance and could potentially assist clinicians in the identification of HCC patients who are more likely to respond to immunotherapy, or at least, providing second opinions on therapeutic decision-making.

IDDF2020-ABS-0088 COMPARING NON-INVASIVE TESTS FOR PREDICTION OF FIBROSIS IN NON-ALCOHOLIC FATTY LIVER DISEASE

Wilson Su*, Yusuke Onishi, Ashis Mukhopadhya. Department of Digestive Disorders, Aberdeen Royal Infirmary, UK

10.1136/gutjnl-2020-IDDF.151

Background Non-alcoholic fatty liver disease (NAFLD) is becoming the most common type of chronic liver disease with an estimated worldwide prevalence of 25%. It is a spectrum of disease that ranges from simple steatosis to advanced fibrosis. Non-invasive tests play an important role in identifying patients with fibrosis that require further investigation and follow up. The aim of the study is to evaluate the diagnostic accuracy of different non-invasive scoring tests.

Methods Patients with NAFLD who underwent liver stiffness measurement (LSM) by FibroScan at Aberdeen Royal Infirmary between 2013 and 2016 were retrospectively included in our study. Patients’ demographic, clinical and laboratory data were collected closest from the date of the FibroScan. NAFLD fibrosis score (NFS), APRI, FIB-4 and BARD scores were calculated. For this study, clinically significant fibrosis (CSF) is defined as LSM > 7 kPa and advanced fibrosis is defined as LSM >12 kPa. The diagnostic accuracy of the four fibrosis scores was examined by calculating the area under the receiver operating characteristic curve (AUROC). The sensitivity, specificity, positive predictive value and negative value were calculated using optimal cut-offs calculated by Youden index.

Results Of the 863 patients included in this study, 498 (57.7%) were male and the mean age was 54.4 years (SD=14.7). The mean BMI was 32.6 (SD=6.4). 48% of patients had CSF with LSM >7 kPa and 28% had advanced fibrosis with LSM>12kPa. For CSF, AUROC curve values were: NFS 0.77 (95% CI, 0.73–0.80), APRI 0.74 (95% CI, 0.70–0.78) and BARDs 0.65 (95% CI, 0.65–0.74). For advanced fibrosis, the AUROC curve values were: NFS 0.83 (95% CI, 0.80 − 0.87), FIB-4 0.79 (95% CI, 0.75 –0.83), APRI 0.75 (95% CI, 0.71 to 0.79) and BARDs 0.75 (95% CI, 0.71–0.79).

Conclusions All fibrosis scores were superior at detecting advanced fibrosis (LSM> 12kPa) compared to CSF (LSM >7kPa). NFS showed a superior diagnostic accuracy of fibrosis compared to other scores.

IDDF2020-ABS-0092 PREDICTION OF POST-HEPATECTOMY LIVER FAILURE IN PATIENTS WITH HEPATOCELLULAR CARCINOMA BASED ON GD-EOB-DTPA-ENHANCED MRI: A LF SCORING MODEL

Yuyan Chen*. The First Affiliated Hospital of Sun Yat-sen University, China

10.1136/gutjnl-2020-IDDF.152

Background The purpose of this study was to establish a pre-operative clinical-radiomics prediction model of post-hepatectomy liver failure (PHLF) in patients with hepatocellular carcinoma (HCC) and to predict clinical outcome of HCC patients who suffered PHLF.

Methods The study included 555 HCC patients who underwent hepatectomy first time from January 2015 to December 2019 in The First Affiliated Hospital of Sun Yat-sen University and Sun Yat-sen University Cancer Center. Gd-EOB-DTPA-Enhanced MRI was performed within 30 days before surgery. Patients in this study didn’t have other tumors or serious organic diseases and were followed up after liver resections for 90 days. We obtained 60720 MR images from 555 patients (an ROI is shown as figure 1), including 390 cases as training data and 165 cases as independent testing data with the standard of PHLF as ISGLS. Through the multiple-model fusion algorithm, we extracted 1044 features per patient from his MR images, screening out feature sets of high contribution by RFE-SVM algorithm and transforming them to FF scores. Clinical indicators, radiologic features and FF scores were included in our LF scoring model through LDA (Linear discriminant Analysis) algorithms.

Results The AUC of LF scoring model reached 0.953 (95%CI 0.953–0.963) and 0.945 (95%CI 0.941–0.980) in the training