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 Sia*, 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.
The model showed the maximum diameter of nodes, PLT, excision site, iMELD score and FF score were independent risk factors of PHLF in HCC patients (OR=1.258, 0.986, 4.670, 1.237, 320.382; P<0.05, and feature coefficients shown as figure 2). Nomogram by Logistic Regression with predictors is shown in figure 5. Moreover, the AUC of LF scoring model was significantly higher than clinical-imaging model and single radiomics model. Meanwhile, the accuracy of the model in predicting overall survival and progression-free survival of HCC patients with PHLF were 0.771 and 0.762, respectively.

Conclusions The study indicates that LF scoring model has a good predictive value for the occurrence of PHLF in HCC patients and plays a vital role in predicting the clinical outcome of HCC patients who suffered PHLF.