Clinical Gastroenterology and Hepatology

Validation of a Model for Identification of Patients With Compensated Cirrhosis at High Risk of Decompensation

--Manuscript Draft--

Manuscript Number:	CGH-D-18-01904R2
Full Title:	Validation of a Model for Identification of Patients With Compensated Cirrhosis at High Risk of Decompensation
Article Type:	Clinical - Liver/Pancreas/Biliary
Section/Category:	Clinical Research (non-trial)
Corresponding Author:	Indra Neil Guha University of Nottingham Nottingham, UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Nottingham
Corresponding Author's Secondary Institution:	
First Author:	Indra Neil Guha
First Author Secondary Information:	
Order of Authors:	Indra Neil Guha
	Rebecca Harris
	Sarah Berhane
	Audrey Dillon
	Lisa Coffey
	Martin James
	Alessandro Cucchett
	David Harman
	Guru aithal
	Omar Elshaarawy
	Imam Waked
	Stephen Stewart
	Philip Johnson
Order of Authors Secondary Information:	
Abstract:	Abstract: Background & Aims: It is important to rapidly identify patients with advanced liver disease. Routine tests to assess liver function and fibrosis provide data that can be used to determine patients' prognoses. We tested the validated the ability of combined data from the ALBI and FIB-4 scoring systems to identify patients with compensated cirrhosis at highest risk for decompensation. Methods: We collected data from 145 patients with compensated cirrhosis (91% Child A cirrhosis and median MELD scores below 8) from a cohort in Nottingham, United Kingdom, followed for a median 4.59 years (development cohort). We collected baseline clinical features and recorded decompensation events. We used these data to develop a model based on liver function (assessed by the ALBI score) and extent of fibrosis (assessed by the FIB-4 index) to determine risk of decompensation. We

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

	validated the model in 2 independent external cohorts (1 in Dublin, Ireland and 1 in Menoufia, Egypt) comprising 234 patients. Results: In the development cohort, 19.3% of the patients developed decompensated cirrhosis. Using a combination of ALBI and FIB-4 scores, we developed a model that identified patients at low vs high risk of decompensation (hazard ratio [HR] for decompensation in patients with high risk score was 7.10). When we tested the scoring system in the validation cohorts, the HR for decompensation in patients with a high-risk score was 12.54 in the Ireland cohort and 5.10 in the Egypt cohort. Conclusion: We developed scoring system, based on a combination of ALBI and FIB-4 scores, that identifies patients at risk for liver decompensation. We validated the scoring system in 2 independent international cohorts (Europe and the Middle East), so it appears to apply to diverse populations.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Nottingham
Biomedical Research Centre

NHS
National Institute for
Health Research

Gastrointestinal and Liver Disorders Theme
NIHR Nottingham Biomedical Research Centre
at the Nottingham University Hospitals NHS Trust
and the University of Nottingham
E Floor West Block
Queen's Medical Centre
Derby Road
Nottingham
NG7 2UH

14/01/2019

Dear Professor Kanwal,

Re : Clinical Gastroenterology and Hepatology - CGH-D-18-01904R1

We are delighted that the journal has provisionally accepted our paper. The one outstanding issue was to perform an additional analysis using NRI and IDI. We have completed this analysis. We have added to the methods and have inserted the results into an amended table 3.

Yours sincerely

Dr Neil Guha
Clinical associate professor in Hepatology

Professor Philip Johnson
Professor in Translational Oncology

What You Need to Know

Background: We combined data from the ALBI and FIB-4 scoring systems to develop a system to identify patients with compensated cirrhosis at highest risk for decompensation.

Findings: Using a combination of ALBI and FIB-4 scores, we developed and validated a model that identified patients at low vs high risk of decompensation.

Implications for patient care: We validated the scoring system in 2 independent international cohorts (Europe and the Middle East), so it might be used worldwide in determining the prognoses of patients with cirrhosis.

Revised Manuscript in Word or RTF (with changes marked) Click here to access/download;Revised Manuscript in Word or RTF (with changes marked);ALBI paper revision_jan

Validation of a Model for Identification of Patients With Compensated Cirrhosis at High Risk of Decompensation

Guha IN*, Harris R*, Berhane S [#], Dillon A [≠], Coffey L[≠], James MW*, Cucchetti A[¥], Harman DJ *, Aithal GP*, Elshaarawy O[§], Waked I[§], Stewart S[≠], and Johnson PJ [#]

* NIHR NOTTINGHAM BIOMEDICAL RESEARCH CENTRE, NOTTINGHAM UNIVERSITY HOSPITALS NHS TRUST AND UNIVERSITY OF NOTTINGHAM, NOTTINGHAM, UK

[§] HEPATOLOGY & GASTROENTEROLOGY DEPARTMENT, NATIONAL LIVER INSTITUTE, MENOUFIA UNIVERSITY, EGYPT

[≠]CENTRE FOR LIVER DISEASE, MATER MISERICORDIAE UNIVERSITY HOSPITAL, 55 ECCLES STREET, DUBLIN 7, IRELAND

¥ DEPARTMENT OF MEDICAL AND SURGICAL SCIENCES - DIMEC; ALMA MATER STUDIORUM, UNIVERSITY OF BOLOGNA, MORGAGNI-PIERANTONI HOSPITAL, FORLÌ; ITALY

[#] DEPARTMENT OF MOLECULAR AND CLINICAL CANCER MEDICINE, UNIVERSITY OF LIVERPOOL, UK

Conflicts of interest: none

Corresponding author: Dr IN Guha (neil.guha@nottingham.ac.uk), NIHR Biomedical Research Centre , University of Nottingham , QMC campus, Derby Road, Nottingham , NG7 2UH, United Kingdom.

Study concept and design: ING and PJ

Data acquisition: RH, DH, LC, AD, MJ, GPA, OE, IW and SS.

Data analysis: AC, SB and PJ

Data interpretation, critical revision for important intellectual content and final approval of the version to be published: all authors

Financial support for this study was received by National Institute for Health Research (NIHR) Nottingham Digestive Diseases Biomedical Research Centre based at Nottingham University Hospitals NHS Trust and University of Nottingham. The funder had no involvement in study design, writing the report or decision for publication.

No conflicts of interest for all authors listed.

Title: Validation of a Model for Identification of Patients With Compensated Cirrhosis at High Risk of Decompensation

Abstract:
Background & Aims: It is important to rapidly identify patients with advanced liver disease. Routine tests to assess liver function and fibrosis provide data that can be used to determine patients' prognoses. We tested the validated the ability of combined data from the ALBI and FIB-4 scoring systems to identify patients with compensated cirrhosis at highest risk for decompensation.

Methods: We collected data from 145 patients with compensated cirrhosis (91% Child A cirrhosis and median MELD scores below 8) from a cohort in Nottingham, United Kingdom, followed for a median 4.59 years (development cohort). We collected baseline clinical features and recorded decompensation events. We used these data to develop a model based on liver function (assessed by the ALBI score) and extent of fibrosis (assessed by the FIB-4 index) to determine risk of decompensation. We validated the model in 2 independent external cohorts (1 in Dublin, Ireland and 1 in Menoufia, Egypt) comprising 234 patients.

Results: In the development cohort, 19.3% of the patients developed decompensated cirrhosis. Using a combination of ALBI and FIB-4 scores, we developed a model that identified patients at low vs high risk of decompensation (hazard ratio [HR] for decompensation in patients with high risk score was 7.10). When we tested the scoring system in the validation cohorts, the HR for decompensation in patients with a high-risk score was 12.54 in the Ireland cohort and 5.10 in the Egypt cohort.

Conclusion: We developed scoring system, based on a combination of ALBI and FIB-4 scores, that identifies patients at risk for liver decompensation. We validated the scoring system in 2 independent international cohorts (Europe and the Middle East), so it appears to apply to diverse populations.

KEY WORDS: liver failure, prognostic factor, alcohol-associated liver disease outcome, NAFLD prediction

What You Need to Know

Background: We combined data from the ALBI and FIB-4 scoring systems to develop a system to identify patients with compensated cirrhosis at highest risk for decompensation.

Findings: Using a combination of ALBI and FIB-4 scores, we developed and validated a model that identified patients at low vs high risk of decompensation.

Implications for patient care: We validated the scoring system in 2 independent international cohorts (Europe and the Middle East), so it might be used worldwide in determining the prognoses of patients with cirrhosis.

INTRODUCTION

The progression of chronic liver disease (CLD), from fibrosis to clinical outcomes, and clinical sequelae including liver decompensation (ascites, variceal bleeding and encephalopathy) manifest relatively late in the natural history and are often the index presentation of liver disease (1). Identification of patients that need intensive monitoring and timely intervention is challenging. Robust prognostic tools using simple laboratory variables, with potential for implementation in non-specialist settings and across different health care systems, have significant appeal.

The Child Pugh and MELD score are extensively validated and easily applicable tools that have been used for decades in clinical practice; the caveats that limit their performance have been previously documented (2-4). Importantly, these scoring systems provide value after synthetic liver function has become significantly deranged and provide only short term prognostic value. Presently, there are no scores, performed in routine clinical practice, that provide robust prognostic stratification within early, compensated cirrhosis over the medium/long term.

Serum albumin and bilirubin have recently been combined in a statistical model as the 'ALBI score', a measure of liver function in patients with HCC (5-8). This model has been extensively validated and subsequently extended to patients with chronic liver disease (without HCC) in a variety of clinical settings (9-11). We hypothesised that the ALBI score, being a measure of hepatic reserve, would be related to the risk of subsequent liver decompensation and tested this hypothesis in the setting of a prospective, longitudinal cohort of patients with compensated cirrhosis in Nottingham, UK. We further explored the combination ALBI with of markers of liver fibrosis (tracking progression of liver disease) as measured by the Fib-4 index. (12-14). Finally, we sought to validate this combined assessment of fibrosis and function in two independent, external cohorts.

PATIENTS AND METHODS

Our primary analysis and model building was focussed on the prospective, longitudinal Nottingham (UK) cohort. We then tested the generalisability of our findings on two independent cohorts from Dublin (Ireland) and Menoufia (Egypt).

Prospective Nottingham (UK) cohort

Patients were consecutively recruited from the Nottingham compensated cirrhosis cohort study (3CN). The 3CN study is a prospective, longitudinal study initiated in 2010 focussed on the study of early compensated liver cirrhosis. The study was approved by a NHS ethics committee and standard regulatory requirements obtained (10/H0403/10). Inclusion criteria were patients between the ages of 18-75 years, an established diagnosis of cirrhosis obtained by at least one of the following criteria: histology, radiological or endoscopic evidence of portal hypertension, clinical evidence of cirrhosis with thrombocytopenia and validated non-invasive liver fibrosis test (Transient elastography >15 kPa). Exclusion criteria included the presence of hepatocellular carcinoma at baseline, portal or splenic vein thrombosis, clinical or radiological ascites at baseline visit, history of variceal haemorrhage and any previous episode of clinical encephalopathy. Patients at baseline had blood tests drawn for routine laboratory measures which included albumin, bilirubin and platelet count. Patients were followed up at six monthly visit appointments until the end of the study duration (10/08/2016). At each clinic visit, routine bloods were drawn for assessment of liver function, full blood count, clotting and renal function and assessment made for the appearance of a liver related clinical event in the intervening period. At the end of the study all patients were assessed for clinical outcomes using digital hospital records and contacting primary care physicians directly in those failing to attend secondary care.

Validation cohorts

Prospective Dublin (Ireland) cohort

Patients were screened at general hepatology outpatients for inclusion criteria and following consent, were invited for a study visit. Enrolment was consecutive and occurred between August 2011 and April 2015. Cirrhosis at this site was confirmed by one of the following criteria: histology proving cirrhosis, endoscopic *or* radiological evidence of varices *and* thrombocytopenia (platelets $<150x10^9$) *or* radiological evidence of splenomegaly (spleen >11cm) or LSM >12 kPa. Patients were excluded if they had any of the following: diagnosis of hepatocellular carcinoma, splenic or portal vein thrombosis, brain injury or unconsciousness, dementia, active substance use that would preclude clinic visits, and decompensated liver disease (clinical or radiological ascites at baseline visit, history of variceal bleeding and overt clinical encephalopathy). The

date and nature of hepatic decompensation, occurring after enrolment in the study, were collected by a liver specialist and supported with either endoscopy reports, abdominal imaging or hospitalization report.

Retrospective Menoufia (Egypt) cohort

Patients were recruited from the outpatient clinic at the National Liver Institute, Menoufia University, Egypt. This was a retrospective cohort starting in 2006 and only patients with complete follow up data were included. Patients diagnosed with liver cirrhosis were included after fulfilling at least one of the following criteria: histopathological diagnosis by liver biopsy, radiological or endoscopic evidence of portal hypertension or transient elastography >14 kPa. Patients with missing data, hepatocellular carcinoma at baseline, history of any malignancies, evidence of portal or splenic vein thrombosis and any patient with history of hepatic decompensation at time of diagnosis (clinical or radiological ascites, variceal bleeding and any previous episode of clinical encephalopathy) were excluded from the study. Routine laboratory diagnostic tests were performed including serum albumin, bilirubin and platelet count. Hepatic decompensation, following enrolment in the study, were ascertained by clinicians in charge of their care and supported with either endoscopy reports, abdominal imaging or hospitalization report.

Statistical methods

Stata/SE 14.2 (StataCorp, Texas, USA) and R software (R 3.4.4)(15) were used to undertake the analysis. Continuous variables were presented as median (with interquartile range) and categorical variables were presented as percentages. Highly skewed variables were $\log_{10}$ transformed. The primary outcome was hepatic decompensation. This was defined using the clinical parameters of first episode of ascites (as defined by confirmation with ultrasonography and requiring treatment with diuretics or paracentesis) or the first variceal bleed (defined by requiring endoscopic intervention) or the first episode of encephalopathy , assessed by an experienced clinician and defined by Grade 3 / 4 West Haven classification ; whichever event occurred first. Time to decompensation (TTD) was calculated from date of entry to study until date of first recorded decompensation. Patients were censored if they underwent liver transplantation or died. Deaths, occurring in hospital and outside the hospital, were collected by using a combination of hospital patient records (Nottingham, Dublin and Egypt), family practitioner records (Nottingham and Dublin), death certificates (Nottingham, Dublin and Egypt) and telephone interviews with relatives (Egypt). When the prospective study in Nottingham was initiated we planned for 50 % decompensation rate at 4 years based on a systematic review by D'Amico and colleagues (16). Therefore, we estimated a cohort of n=150, followed for at least 4 years, would yield 75 events. Using the 10:1 rule of thumb for prediction models, we anticipated this would allow us to explore at least 7 prognostic variables. We did not use imputation for missing variables.

ALBI vs MELD

Liver function was initially measured by ALBI or MELD score. Univariable Cox regression was undertaken for both variables. Decompensation risk predictions by MELD and ALBI were compared using Harrell's C statistic (17), Gönen & Heller's K (18) and Royston & Sauerbrei's R^2D (19). Higher values of these measures translate to better survival prediction. Confidence intervals (C.I.) and p-values of the comparisons were estimated using the bootstrap method (1000 samples). Patients who have both ALBI and MELD recorded where included in this part of the analysis.

Association between ALBI, FIB-4 and TTD

TTD according to baseline ALBI grade were examined via Kaplan-Meier (KM) graphs and compared using the log-rank test. Association between ALBI score and FIB-4 (appendix 1) and TTD was tested using univariable Cox regression analysis.

Building of the ALBI-FIB4 Score

A multivariable Cox model combining ALBI score and FIB-4 was generated. Model fit was compared to the univariable models using the log-likelihood ratio (LR) test. Any violation in the proportional hazards assumption was tested by examining the Schoenfeld residuals. The formula for the new score (linear predictor) was produced using the coefficients of the model. A high risk group of patients were identified by applying a cut-off at the 85th centile of the score (patients within the top 15% risk). Our usual approach has been to define four classes based on cut-offs at the 15th, 50th and 85th centile of the linear predictor as proposed by Cox and Royston (20, 21). However in this study, the numbers of decompensation events was too small to justify a four class model and we therefore used the 85th centile figure as the cut-off.

Model performance and validation

The new model was validated in two independent external cohorts from Dublin (Ireland) and Menoufia (Egypt). Median survival, hazard ratios and percentage decompensation at 3 and 5 years were calculated for each risk group in both the derivation and validation sets. KM graphs according to the two risk groups were also plotted for the derivation and validation sets.

In order to make the formula applicable to a clinics setting we developed a simple online calculator (https://jscalc.io/calc/gdEJj89Wz5PirkSL).

Model comparisons

The discriminatory performance of ALBI-FIB4 score was compared to the other scores, namely ALBI, FIB-4, MELD, and Child-Pugh score using Harrell's C statistic, Gönen & Heller's K and Royston & Sauerbrei's R^2D.

Models were also compared using integrated discrimination improvement (IDI) and category-free net reclassification improvement ($NRI^{>0}$). IDI measures the difference in the discrimination slopes between the new model ALBI-FIB4 and the older models (ALBI, FIB4, MELD and Child-Pugh).(22) Discrimination slope is a measure of the separation in predicted probabilities for events and non-events.(23) NRI measures the amount of correct reclassifications by the new model (ALBI-FIB4) compared to the older models (ALBI, FIB4, MELD and Child-Pugh) based on calculated predicted risk probabilities.(22-24). In this analysis, an extension of IDI and NRI that incorporates censored survival data was used.(24, 25) Confidence intervals (C.I.) and p-values for IDI and NRI were estimated using perturbation-resampling (1000 iterations). Model comparisons was undertaken in both the derivation and validation sets. Due to the low number of decompensation events and in order to improve the statistical power, both validation sets were combined.

RESULTS

The Nottingham cohort comprised 145 patients with cirrhosis (91% Child Pugh A grade). The most common aetiology was alcohol (44.8%) followed by non-alcoholic fatty liver disease (29.7%) and other aetiologies as shown in Table 1. The median length of follow-up was 4.59 years (95% CI 4.02, 4.90), with an overall decompensation rate of 19.3% and an overall survival of 82.1%. The most common decompensation event was ascites development (71.4 %) followed by variceal bleeding (21.4 %) and grade 3/4 encephalopathy (7.2 %). The median ALBI score at baseline was -2.43 (-2.72 to -2.18).

The Dublin cohort comprised of 141 patients with cirrhosis (90% Child Pugh Grade A). The most common aetiology was alcohol (39.7%) followed by HCV (29.8%). The overall decompensation rate was 12.1% over a maximum time of 6.4 years. The most common decompensation event was variceal haemorrhage (47.1%), followed by ascites (29.4%) and encephalopathy (both 23.5%). The median ALBI score at baseline was -2.48 (-2.71 to -2.13).

The Menoufia cohort comprised of 93 patients with cirrhosis (88% Child Pugh Grade A). The most common aetiology was NAFLD (47.3%) followed by HCV (34.4%). The overall decompensation rate was 38.7 % over a maximum time of 10.6 years. The most common decompensation event was ascites (57.1%), followed by variceal bleeding (22.9%), followed by encephalopathy (14.3%). The median ALBI score at baseline was -2.52 (-2.93 to -2.05)

ALBI vs MELD

Both baseline ALBI score and MELD were significantly influenced decompensation with hazard ratios (HR) 7.90 (95% C.I. 3.37, 18.54), p<0.0001 and 1.20 (95% C.I. 1.07, 1.35), p=0.002 respectively. The Harrell's C statistic, Gönen & Heller's K and Royston & Sauerbrei's R^2D for ALBI vs MELD were as follows: 0.755 (95% C.I. 0.663, 0.835) vs 0.715 (95% C.I. 0.621, 0.804), 0.699 (95% C.I. 0.634, 0.747) vs 0.583 (95% C.I. 0.531, 0.651) and 0.374 (95% C.I. 0.164, 0.560) vs 0.213 (95% C.I. 0.058, 0.411) respectively. In all three measures, ALBI showed higher values compared to MELD, although some of these were not significant (p=0.480, p=0.001 and p=0.122 respectively). Furthermore, adding MELD to FIB-4 did not improve the fit of the model (LR test, p=0.4058) compared to adding ALBI score to FIB-4 (p=0.0078). ALBI was therefore considered a better measure for liver function compared to MELD for the purpose of this analysis.

Association between ALBI, FIB-4 and TTD

As expected in a cohort of compensated cirrhosis, the majority of patients were in ALBI grade 1 (34.5%) or ALBI grade 2 (63%) with only a minority in grade 3 (3%). Figure 1 shows TTD according to baseline ALBI grade 1 and 2. There were significant (p=0.0102) differences in TTD curves between the two grades. Univariable Cox regression analysis showed that both ALBI score (HR 6.09 [95% C.I. 2.87, 12.96], p<0.0001) and FIB-4 (HR 1.26 [95% C.I. 1.14, 1.40], p<0.0001) strongly influenced TTD.

Building of the ALBI-FIB4 Score

A multivariable Cox regression model that combined ALBI and FIB-4 scores had a better fit compared to when either of the variables were taken alone (likelihood ratio test p=0.0167, upon adding FIB-4 to ALBI score and p=0.0069 vice versa). In the multivariable model, the hazard ratio, 95% C.I. and p-value for the ALBI score and FIB-4 were 3.79 (1.53, 9.36), p=0.004 and 1.18 (1.05, 1.33), p=0.008 respectively.

The formula for this new ALBI-FIB4 score was as follows: (ALBI score*1.331) + (FIB-4*0.165). Patients with a score greater than -1.822 were considered high risk whereas those equal to or below -1.822 as low risk.

Model performance and validation

Nottingham, UK (derivation cohort)

Using the ALBI-FIB4 score, 85% of the cohort were stratified in the low risk group and 15% stratified in the high risk group. The groups showed clear separation (p<0.0001) in those reaching a clinical end point of decompensation (Figure 2A). In the low risk group at baseline, 9.3 % reached a decompensation event within 3 years. This is in contrast to those in a high risk group where 48.4 % reached a decompensation event rate

within 3 years (Table 2). The hazard ratios of reaching a decompensation event was 7.10 (95% C.I. 3.07 to 16.42) in the high risk group compared to the low risk group.

Dublin, Ireland and Menoufia, Egypt (external validation cohorts)

In the external validation cohorts the ALBI-FIB4 index stratified 81% of the Egyptian cohort and 82% of the Irish cohort into the low risk group. The high and low risk groups again showed clear separation ($p<0.001$) in both validation cohorts (Figure 2B and 2C). The hazard ratio of reaching decompensation between the low and high risk groups was 12.54 (95% C.I. 4.25 to 36.93) in the Irish cohort and 5.10 (95% C.I. 2.07 to 12.59) in the Egyptian cohort. Percentage of patients reaching decompensation event within the first 3 or 5 years in the two validation cohorts are shown in Table 2.

Model comparisons

Table 3 displays survival prediction comparisons between ALBI-FIB4 and other scores, namely ALBI, FIB-4, MELD and Child-Pugh score using Harrell's C statistic, Gönen & Heller's K,Royston & Sauerbrei's R^2D, IDI and NRI in both the derivation and validation sets. The table shows that overall ALBI-FIB4 and in some instances, ALBI score, gave higher values of the five measures compared to the rest, although there may not be enough statistical power to detect significant differences in many of the comparisons due to low number of decompensation events.

DISCUSSION

We have validated the ALBI score in a prospectively accrued data set as a measure of liver dysfunction in patients with chronic liver disease. In addition, we created a new score, 'ALBI-FIB4' which can effectively stratify patients for the risk of future liver decompensation. The ALBI-FIB4 score identified a high risk group more effectively than the MELD score and maintained performance in two external cohorts with distinct differences in aetiology of cirrhosis.

The Nottingham cohort represents early compensated cirrhosis as evidenced by the fact that 91% of the cohorts were Child Pugh stage A (9% child Pugh stage B) with a median MELD score of less than 8 at baseline. This explains why the cohort had predominance of ALBI grade 1 (35%) or ALBI grade 2 (63%) at baseline; with only a small fraction of ALBI grade 3 (3%). When followed prospectively over 5 years the ALBI grade showed impressive prognostic separation; a decompensation rate of 11.5 % versus 27.3 % in grade 1 versus grade 2 respectively. The validation of ALBI in a compensated cirrhosis cohort of mixed aetiology supports previous

findings in compensated cirrhosis and chronic hepatitis B (11). Prior to this, the body of evidence for ALBI was focussed on patients with hepatocellular carcinoma (5-8).

FIB-4 has been extensively validated as a non-invasive marker of fibrosis and shown to have a prognostic ability to predict clinical outcomes across chronic liver disease including HCV, HBV and NAFLD (12-14). The ALBI score is postulated to show early changes in synthetic function and the use of the ratio results in more sensitivity than simply observing for the individual parameters to cross the upper limit of normal. The combination has biological plausibility, by looking at markers of fibrosis and function simultaneously. The combined score showed superiority within the multivariable cox regression model but this was not translated to statistical differences in Harrell's C statistic or IDI/NRI.

An important strength of our study is that we utilised the robust clinical end points of ascites, encephalopathy and variceal bleeding as originally described in the seminal natural history studies of cirrhosis (26, 27). We did not use jaundice or change in Child Pugh classification as these are directly influenced by prognostic variables we chose to study. The Nottingham cohort is a mixed aetiology cohort with compensated cirrhosis followed prospectively for clinical events. This cohort represents a formal, prospective, protocol driven study designed a priori to document the clinical history of patients with compensated cirrhosis. The collection of hard clinical outcomes, using a variety of source documents, was a central aspect of this study but we accept that within each cohort intrinsic limitations exist; e.g. a patient migrating to another health region during the period of study. The prevalence of the different aetiologies, with alcohol the dominant aetiology followed by NAFLD and viral hepatitis is very similar to a large population study of UK patients with cirrhosis (28). This implies the data has direct application to a population in which simple markers such as ALBI and FIB-4 has the greatest relevance. The validation cohort from Ireland, driven by the aetiologies of alcohol and HCV was in contrast to the cohort from Egypt with the predominant aetiologies of HCV and NAFLD. This provides reassurance that the model has generalisability for stratifying liver disease at an international level. A limitation of stratification based on the ALB-FIB-4 score is that only 15% are classified in the high risk group). We deliberately focussed on the previously validated scores of ALBI and FIB-4 in this study as our aim was to create a tool that could be used within a community setting or low resource countries.

A frequently levelled criticism of algorithms such as ALBI-FIB-4 is that they are too complicated to be applied routinely in the clinical setting. To overcome this problem we developed a simple online calculator which can be accessed using the following link: https://jscalc.io/calc/gdEJj89Wz5PirkSL.With respect to comparable models which are routinely performed in clinical practice, the ALBI-FIB-4 score was numerically

superior to the other scores including the MELD score and Child Pugh Score. However a limitation of the low number of events was that we were not able to show consistent statistical superiority or look at a wide number of variables in the analysis.

We have shown that routinely available laboratory variables, combined in a novel algorithm ALBI-FIB-4, can stratify patients with cirrhosis for future risk of liver decompensation. The ability to do this in the context of early, compensated cirrhosis with preserved liver synthetic function whilst also predicting long term clinical outcomes has clinical utility for international health care systems.

Acknowledgements

We acknowledge the support of the 3CN research team including nurses and support staff; Andrea Bennett, Tracey Wildsmith, Louise James, Antonella Ghezzi, Fozia Naushahi, Melanie Lingaya, Yirga Falcone, Denise Marshall and Govind Jumbu. We thank the patients and their relatives at Nottingham, Menoufia and Dublin.

.

NIHR disclaimer

This is a summary of independent research supported by the National Institute for Health Research (NIHR) Nottingham Digestive Diseases Biomedical Research Centre based at Nottingham University Hospitals NHS Trust and University of Nottingham. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

REFERENCES

1. Ratib S, Fleming KM, Crooks CJ, Aithal GP, West J. 1 and 5 year survival estimates for people with cirrhosis of the liver in England, 1998-2009: a large population study. J Hepatol. 2014;60(2):282-9.

2. Cholongitas E, Papatheodoridis GV, Vangeli M, Terreni N, Patch D, Burroughs AK. Systematic review: The model for end-stage liver disease--should it replace Child-Pugh's classification for assessing prognosis in cirrhosis? Aliment Pharmacol Ther. 2005;22(11-12):1079-89.

3. Cholongitas E, Senzolo M, Triantos C, Samonakis D, Patch D, Burroughs AK. MELD is not enough--enough of MELD? J Hepatol. 2005;42(4):475-7; author reply 8-9.

4. Bernardi M, Gitto S, Biselli M. The MELD score in patients awaiting liver transplant: strengths and weaknesses. J Hepatol. 2011;54(6):1297-306.

5. Toyoda H, Lai PB, O'Beirne J, Chong CC, Berhane S, Reeves H, et al. Long-term impact of liver function on curative therapy for hepatocellular carcinoma: application of the ALBI grade. Br J Cancer. 2016;114(7):744-50.

6. Johnson PJ, Berhane S, Kagebayashi C, Satomura S, Teng M, Reeves HL, et al. Assessment of liver function in patients with hepatocellular carcinoma: a new evidence-based approach-the ALBI grade. J Clin Oncol. 2015;33(6):550-8.

7. Hiraoka A, Kumada T, Michitaka K, Toyoda H, Tada T, Ueki H, et al. Usefulness of albumin-bilirubin grade for evaluation of prognosis of 2584 Japanese patients with hepatocellular carcinoma. J Gastroenterol Hepatol. 2016;31(5):1031-6.

8. Pinato DJ, Sharma R, Allara E, Yen C, Arizumi T, Kubota K, et al. The ALBI grade provides objective hepatic reserve estimation across each BCLC stage of hepatocellular carcinoma. J Hepatol. 2017;66(2):338-46.

9. Peng Y, Qi X, Tang S, Deng H, Li J, Ning Z, et al. Child-Pugh, MELD, and ALBI scores for predicting the in-hospital mortality in cirrhotic patients with acute-on-chronic liver failure. Expert Rev Gastroenterol Hepatol. 2016;10(8):971-80.

10. Zou D, Qi X, Zhu C, Ning Z, Hou F, Zhao J, et al. Albumin-bilirubin score for predicting the in-hospital mortality of acute upper gastrointestinal bleeding in liver cirrhosis: A retrospective study. Turk J Gastroenterol. 2016;27(2):180-6.

11. Chen RC, Cai YJ, Wu JM, Wang XD, Song M, Wang YQ, et al. Usefulness of albumin-bilirubin grade for evaluation of long-term prognosis for hepatitis B-related cirrhosis. J Viral Hepat. 2017;24(3):238-45.

12. Vergniol J, Foucher J, Terrebonne E, Bernard PH, le Bail B, Merrouche W, et al. Noninvasive tests for fibrosis and liver stiffness predict 5-year outcomes of patients with chronic hepatitis C. Gastroenterology. 2011;140(7):1970-9, 9 e1-3.

13. de Ledinghen V, Vergniol J, Barthe C, Foucher J, Chermak F, Le Bail B, et al. Non-invasive tests for fibrosis and liver stiffness predict 5-year survival of patients chronically infected with hepatitis B virus. Aliment Pharmacol Ther. 2013;37(10):979-88.

14. Angulo P, Bugianesi E, Bjornsson ES, Charatcharoenwitthaya P, Mills PR, Barrera F, et al. Simple noninvasive systems predict long-term outcomes of patients with nonalcoholic fatty liver disease. Gastroenterology. 2013;145(4):782-9 e4.

15. Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018 https://www.R-project.org/.

16. D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. J Hepatol. 2006;44(1):217-31.

17. Newson RB. Comparing the predictive powers of survival models using Harrell's C or Somers' D. Stata Journal. 2010;10(3):339.

18. Gönen M, Heller G. Concordance probability and discriminatory power in proportional hazards regression. Biometrika. 2005;92(4):965-70.

19. Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. Statistics in medicine. 2004;23(5):723-48.

20. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol. 2013;13:33.

21. Cox DR. Note on Grouping. Journal of the American Statistical Association. 1957;52(280):543-7.

22. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008;27(2):157-72; discussion 207-12.

23. Pencina MJ, D'Agostino RB, Vasan RS. Statistical methods for assessment of added usefulness of new biomarkers. Clin Chem Lab Med. 2010;48(12):1703-11.

24. Pencina MJ, D'Agostino RB, Sr., Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med. 2011;30(1):11-21.

25. Uno H, Tian L, Cai T, Kohane IS, Wei LJ. A unified inference procedure for a class of measures to assess improvement in risk prediction systems with survival data. Stat Med. 2013;32(14):2430-42.

26. Saunders JB, Walters JR, Davies AP, Paton A. A 20-year prospective study of cirrhosis. Br Med J (Clin Res Ed). 1981;282(6260):263-6.

27. Gines P, Quintero E, Arroyo V, Teres J, Bruguera M, Rimola A, et al. Compensated cirrhosis: natural history and prognostic factors. Hepatology. 1987;7(1):122-8.

28. Ratib S, West J, Crooks CJ, Fleming KM. Diagnosis of liver cirrhosis in England, a cohort study, 1998-2009: a comparison with cancer. Am J Gastroenterol. 2014;109(2):190-8.

Appendix 1

ALBI-score:

(log_{10} bilirubin[μmol/l] x 0.66) + (albumin[g/l] x - 0.085)

FIB-4 score : Age × AST (IU/L)/ platelet count ($\times 10^9$/L) × √ALT (IU/L)

Legends for tables and figures

Table 1: Baseline characteristics of cohorts

Table 2: Summary of median decompensation, percentage decompensation at 3 and 5 years with hazard ratios of the different categories

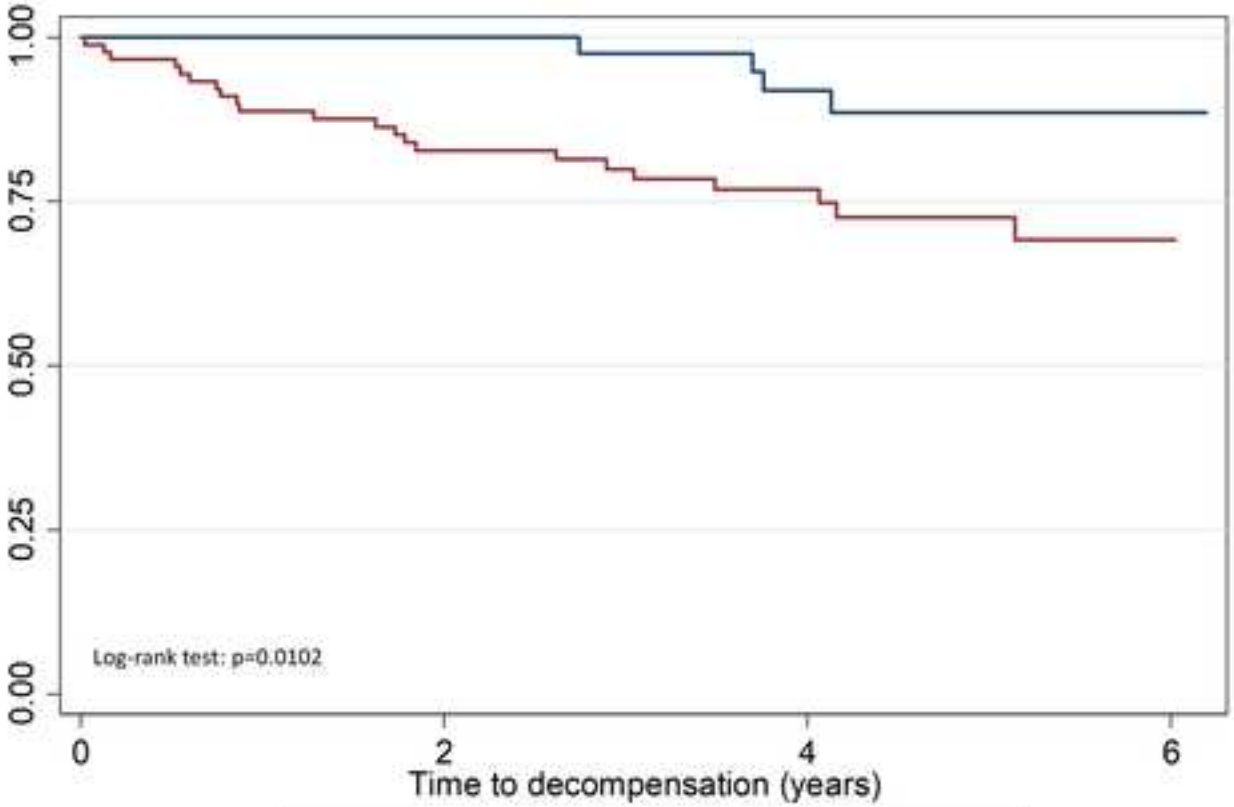
Table 3: Comparing ALBI-FIB4 model to other scores

Figure 1: Decompensation using ALBI grade at baseline in Nottingham cohort

Figure 2: Time to decompensation according to the (A) UK (B) Ireland and (C) Egypt cohorts using combined ALBI-FIB-4 score

Click here to access/download;Figure;updated figure 1.tif

Figure 1



Nottingham				
ALBI grade	N	Median (95% CI), years	% decompensation at 3-years	% decompensation at 5-years
1	50	Not reached	2.4 (0.4, 16.1)	11.5 (4.4, 28.1)
2	91	Not reached	20.0 (12.9, 30.3)	27.3 (18.4, 39.4)

Click here to access/download;Figure;updated figure 2.tif

Figure 2

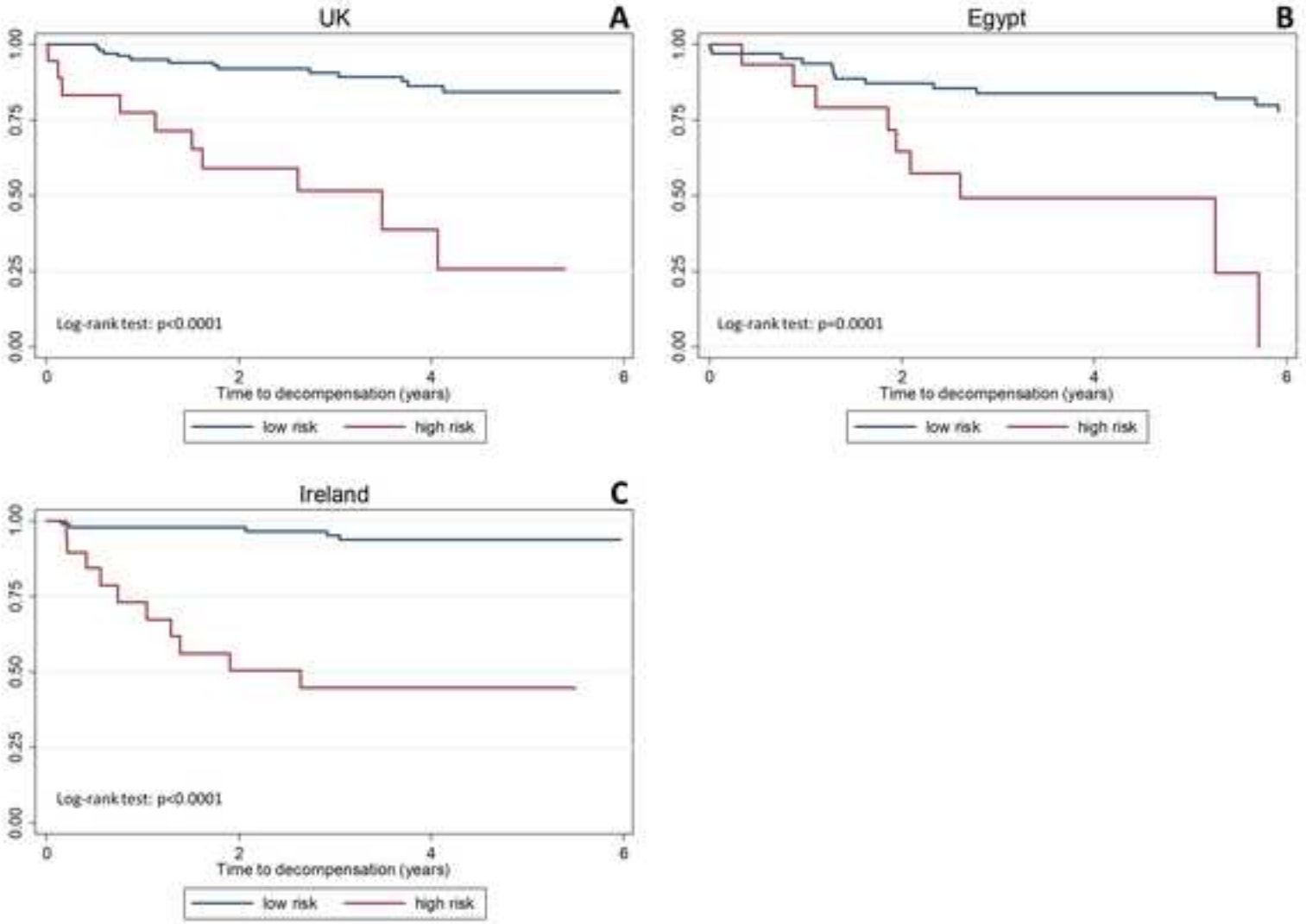


Table 1

TABLE 1

Variables	UK (N=145)	Egypt (N=93)	Ireland (N=141)
Male, n(%)	95 (65.5), n=145	70 (75.3), n=93	99 (70.2)
Age, years	61 (55, 66), n=145	53 (46, 59), n=93	54 (45, 61), n=140
Aetiology of disease , n(%)	n=145	n=93	n=141
ALD	65 (44.8)	7 (7.5)	56 (39.7)
NAFLD/NASH	43 (29.7)	44 (47.3)	12 (8.5)
HCV	17 (11.7)	32 (34.4)	42 (29.8)
HBV	1 (0.7)	3 (3.2)	7 (5.0)
PBC	5 (3.5)	0 (0)	5 (3.6)
PSC	2 (1.4)	0 (0)	0 (0)
Autoimmune hepatitis	4 (2.8)	1 (1.1)	0 (0)
Cryptogenic cirrhosis	3 (2.1)	6 (6.5)	3 (2.1)
Haemochromatosis	5 (3.5)	0 (0)	0 (0)
Other	0 (0)	0 (0)	16 (11.4)
Child-Pugh grade, n(%)	n=142	n=93	n=141
A	129 (90.9)	82 (88.2)	127 (90.1)
B	13 (9.2)	11 (11.8)	13 (9.2)
C	0 (0)	0 (0)	1 (0.7)
Bilirubin (µmol/L)	12 (9, 18), n=145	17.1 (10.3, 27.4), n=93	15.5 (11, 22), n=138
Albumin (g/l)	37 (34, 40), n=145	39 (35, 42.5), n=93	38 (35, 41), n=138
ALBI score	-2.43 (-2.72, -2.18), n=145	-2.52 (-2.93, -2.05), n=93	-2.48 (-2.71, -2.13), n=138
FIB-4 Score	2.78 (1.86, 3.83), n=121	3.84 (1.68, 6.57), n=93	3.03 (1.94, 4.88), n=126
Platelets	145 (111, 209), n=145	176 (119, 240), n=93	136 (97, 194), n=139
INR	1.0 (1.0, 1.1), n=141	1.0 (0.9, 1.1), n=93	1.1 (1.1, 1.2), n=139
ALBI grade, n(%)	n=145	n=93	n=138
1	50 (34.5)	40 (43.0)	51 (37.0)
2	91 (62.8)	47 (50.5)	84 (60.9)
3	4 (2.8)	6 (6.5)	3 (2.2)
MELD	7.50 (6.43, 8.47), n=142	6.98 (6.43, 10.52), n=91	8.45 (7.50, 9.77), n=138

Abbreviations: AFP, alpha-fetoprotein; ALBI, Albumin-Bilirubin; ALD, alcoholic liver disease; HBV, hepatitis B virus; HCV, hepatitis C virus; INR, international normalised ratio; MELD, Model for End-Stage Liver Disease; ALBI grades 1, 2, and 3 were defined as ALBI ≤−2.60, ALBI > −2.60 but ≤ −1.39, and ALBI > −1.39, respectively NA, not available; NAFLD, non-alcoholic fatty liver disease; PBC, primary biliary cirrhosis, PSC, primary sclerosing cholangitis. Continuous variables presented as median and interquartile range. Categorical variables presented as numbers and percentages.

Table 2

Table 2: Summary of median decompensation-free survival, percentage decompensation at 3/5 years and hazard ratios of the different categories

Cohort	Category	N	Median decompensation-free survival (95% CI), years	% decompensation at 3-years	% decompensation at 5-years	Hazard ratio (95% CI)	p-value
UK	**Low risk group**	103	Not reached	9.32 (4.94, 17.20)	15.70 (9.24, 25.97)	1	
	High risk group	18	3.49 (1.14, .)	48.44 (27.51, 74.43)	74.22 (45.47, 95.17)	7.10 (3.07, 16.42)	<0.0001
Egypt	**Low risk group**	63	Not reached	16.15 (9.03, 27.96)	16.15 (9.03, 27.96)	1	
	High risk group	15	2.60 (1.10, .)	50.77 (28.07, 78.22)	50.77 (28.07, 78.22)	5.10 (2.07, 12.59)	<0.0001
Ireland	**Low risk group**	86	Not reached	5.02 (1.91, 12.86)	6.41 (2.71, 14.78)	1	
	High risk group	19	2.65 (0.75, .)	55.09 (34.45, 78.07)	55.09 (34.45, 78.07)	12.54 (4.25, 36.93)	<0.0001

Table 3

Table 3: Comparing ALBI-FIB4 model to other scores

Cohort	Model	Harrell's C index (95% C.I.*)	p-value** (C-index)	Gönen & Heller's K (95% C.I.*)	p-value** (K)	Royston & Sauerbrei's R^2D (95% C.I.)	p-value** (R^2D)	Models compared (for IDI/NRI)	Integrated discrimination improvement (IDI) (95% C.I.##)	p-value## (IDI)	Net reclassification improvement (NRI) (95% C.I.##)	p-value## (NRI)
UK	ALBI-FIB4	0.805 (0.718, 0.873)	Ref	0.684 (0.624, 0.734)	Ref	0.422 (0.220, 0.589)	Ref					
	ALBI	0.754 (0.629, 0.830)	0.078	0.691 (0.622, 0.740)	0.722	0.369 (0.164, 0.576)	0.469	ALBI-FIB4 vs ALBI	0.046 (-0.047, 0.140)	0.270	0.232 (-0.236, 0.560)	0.362
	FIB4	0.775 (0.680, 0.855)	0.402	0.617 (0.572, 0.686)	0.005	0.381 (0.201, 0.567)	0.561	ALBI-FIB4 vs FIB4	0.066 (-0.016, 0.169)	0.132	0.205 (-0.172, 0.543)	0.292
	MELD	0.699 (0.583, 0.795)	0.058	0.569 (0.511, 0.633)	0.004	0.177 (0.014, 0.361)	0.026	ALBI-FIB4 vs MELD	0.189 (0.055, 0.329)	0.002	0.443 (0.023, 0.706)	0.038
	Child-Pugh score	0.701 (0.594, 0.797)	0.046	0.635 (0.569, 0.682)	0.106	0.313 (0.067, 0.531)	0.351	ALBI-FIB4 vs CP-score	0.104 (-0.028, 0.220)	0.108	0.307 (-0.042, 0.628)	0.082
Egypt & Ireland	ALBI-FIB4	0.776 (0.676, 0.853)	Ref	0.693 (0.653, 0.732)	Ref	0.481 (0.304, 0.659)	Ref					
	ALBI	0.743 (0.631, 0.839)	0.189	0.732 (0.670, 0.783)	0.076	0.440 (0.253, 0.619)	0.495	ALBI-FIB4 vs ALBI	-0.004 (-0.106, 0.078)	0.953	-0.200 (-0.403, 0.326)	0.547
	FIB4	0.697 (0.589, 0.786)	0.045	0.636 (0.596, 0.675)	0.001	0.267 (0.081, 0.474)	0.011	ALBI-FIB4 vs FIB4	0.069 (0.016, 0.123)	0.006	0.359 (0.167, 0.561)	<0.0001
	MELD	0.703 (0.601, 0.794)	0.025	0.591 (0.550, 0.650)	0.001	0.230 (0.084, 0.375)	0.002	ALBI-FIB4 vs MELD	0.153 (0.047, 0.253)	0.008	0.386 (0.079, 0.643)	0.024
	Child-Pugh score	0.692 (0.603, 0.775)	0.031	0.641 (0.601, 0.680)	0.019	0.395 (0.206, 0.589)	0.312	ALBI-FIB4 vs CP-score	0.057 (-0.033, 0.129)	0.182	0.214 (-0.191, 0.467)	0.194

*Bias-corrected bootstrap confidence intervals (C.I.) estimated from 1000 samples. **p-value estimated from 1000 bootstrap samples of the difference.
##Confidence intervals and p-values were estimated via perturbation-resampling (1000 iterations).
IDI and NRI calculated at 5 year time-point.

Revised Manuscript in Word or RTF (no changes marked)

Click here to access/download
Revised Manuscript in Word or RTF (no changes marked)
ALBI paper revision_jan 2019_Final_clean version.docx