Received: 8 July 2022 | Accepted: 17 August 2022

DOI: 10.1002/hep4.2091

ORIGINAL ARTICLE

Noninvasive predictors of clinically significant portal hypertension in NASH cirrhosis: Validation of ANTICIPATE models and development of a lab-based model

Anahita Rabiee[1] | **Yanhong Deng**[2] | **Maria Ciarleglio**[2] | **Jean L. Chan**[3] | **Monica Pons**[4] | **Joan Genesca**[4,5] | **Guadalupe Garcia-Tsao**[1,6]

[1]Section of Digestive Diseases, Yale School of Medicine, New Haven, Connecticut, USA

[2]Department of Biostatistics, School of Public Health, Yale University, New Haven, Connecticut, USA

[3]Conatus Pharmaceuticals at the time of study conduct, San Diego, California, USA

[4]Liver Unit, Department of Internal Medicine, Hospital Universitari Vall d'Hebron, Vall d'Hebron Institut de Recerca (VHIR), Vall d'Hebron Barcelona Hospital Campus, Universitat Autònoma de Barcelona, Barcelona, Spain

[5]Centro de Investigacion Biomédica en Red de Enfermedades Hepáticas y Digestivas, Instituto de Salud Carlos III, Madrid, Spain

[6]VA-CT Healthcare System, New Haven, Connecticut, USA

Correspondence
Guadalupe Garcia-Tsao, VA Connecticut Healthcare System, 950 Campbell Avenue, West Haven Campus, West Haven, CT 06516, USA.
Email: guadalupe.garcia-tsao@yale.edu

Funding information
Yale Liver Center, Grant/Award Number: P30 DK34989

Abstract

Clinically significant portal hypertension (CSPH), defined as hepatic venous pressure gradient (HVPG) ≥ 10 mm Hg, identifies patients with compensated cirrhosis at a high risk of decompensation. However, HVPG is an invasive and nuanced method. The ANTICIPATE models, which include liver stiffness measurements by transient elastography (TE) and platelet count ± body mass index, are robust noninvasive surrogates of CSPH but required external validation in patients with nonalcoholic steatohepatitis (NASH) cirrhosis. Additionally, TE is not widely available worldwide. The aims of the study were: (1) to externally validate the ANTICIPATE models using baseline data from patients with compensated NASH cirrhosis screened/enrolled in a multicenter international randomized controlled trial; and (2) to develop and externally validate a model using only laboratory values. Regarding aim 1, both ANTICIPATE models showed good calibration and discrimination (area under the curve [AUC] > 0.8) in our cohort (n = 222). Regarding aim 2, a new lab-based model using the Fibrosis-4 index (FIB-4 [age, aspartate aminotransferase, alanine aminotransferase, platelet count]) plus serum albumin was developed. The discrimination in the training cohort (n = 309) was good (AUC of 0.78 [95% confidence interval [CI]:0.72–0.83]). It was then externally validated in a separate cohort of 245 patients with compensated NASH cirrhosis (AUC of 0.8 [95% CI: 0.75–0.86]). Given the difference in the prevalence of CSPH between training (74%) and validation (39%) cohorts, the model required an update of the baseline risk to achieve a good calibration. The updated model was named FIB4+. In conclusion, both ANTICIPATE models performed well in predicting the presence of CSPH in NASH cirrhosis. A model using FIB-4 plus albumin (FIB4+) can be used to predict CSPH where TE is not available.

INTRODUCTION

Portal hypertension is the main driver of cirrhosis decompensation, which is defined as the development of ascites, variceal hemorrhage, or hepatic encephalopathy. Portal pressure is determined by the hepatic venous pressure gradient (HVPG) with a normal HVPG ranging between 3 and 5 mm Hg. A study derived from the "timolol" randomized controlled trial (RCT), which included 213 patients with compensated cirrhosis, showed that an HVPG ≥ 10 mm Hg was the main predictor of development of cirrhosis decompensation.[1] Therefore, this threshold HVPG now defines the entity known as clinically significant portal hypertension (CSPH). In addition to its prognostic value, the stratification of patients with cirrhosis into those with and without CSPH has important clinical implications, as it has been shown that, in patients with cirrhosis and CSPH, decompensation can be prevented by the use of nonselective beta blockers (NSBBs).[2] However, obtaining the HVPG requires an invasive procedure and is not readily available in daily clinical practice. Recently, liver stiffness measurements (LSMs) obtained by transient elastography (TE), a noninvasive test, have been found to be useful in predicting the presence of CSPH.[3–6]

In a multicenter study by Abraldes et al. (ANTICIPATE study), among several noninvasive markers, a combination of LSM and platelet count (ANTICIPATE model) was found to be predictive of CSPH, among other outcomes. This study included patients with LSM ≥ 10 kPa, which defines compensated advanced chronic liver disease (cACLD). Most patients had hepatitis C (66%), and only a minority (5%) had nonalcoholic steatohepatitis (NASH).[6] Subsequently a multicenter study by Pons et al. aimed to externally validate the ANTICIPATE model in patients with cACLD of all etiologies.[7] In this study, 248 patients (29.7%) had NASH cirrhosis, and the ANTICIPATE model did not perform as well in patients with NASH who were obese. Therefore, a model was developed incorporating body mass index (BMI) in addition to LSM and platelet count (ANTICIPATE-NASH model), which performed well in patients with NASH cirrhosis, both obese and nonobese.[7] Our study had two objectives. The first was to externally validate the predictive value of the ANTICIPATE and ANTICIPATE-NASH models in determining the presence of CSPH in a cohort of patients with compensated NASH cirrhosis included in an international multicenter RCT. Because TE is not available in many centers worldwide, our second objective was to determine whether a model excluding LSM and including easily obtainable laboratory indices/values could also predict CSPH in these patients.

METHODS

Participants

Baseline data of patients screened/enrolled in a RCT of emricasan versus placebo in patients with NASH cirrhosis were included in the current study.[8] The trial enrolled patients from 59 sites in the United States and Europe. The key inclusion criterion for the main trial was NASH cirrhosis with an HVPG ≥ 12 mm Hg, and 263 patients were randomized in the trial, 188 of whom had compensated cirrhosis.[8] However, for the current study, we also included data from 121 screened patients with compensated cirrhosis who could not be included in the RCT because screening HVPG was < 12 mm Hg.

The sponsoring company, Conatus, kindly provided a deidentified database of baseline characteristics of patients included/screened in the trial, including BMI, laboratory values, LSM (using TE) and HVPG values that constituted the basis for the current study. For the first aim of the study, the following patients were excluded: (1) those without LSMs, or with poor-quality LSM defined as LSM with interquartile range > 30% or success rate < 60% on 10 measurements (similar to the criteria used by ANTICIPATE and ANTICIPATE-NASH studies); (2) those with poor-quality HVPG tracing (from an ordinal scale of excellent/good/fair/poor per the central reader); and (3) those without an available platelet count. For the second aim of the study, patients with an incomplete set of laboratory values at baseline were excluded.

External validation cohort consisted of 245 patients with compensated NASH cACLD and a LSM ≥ 10 kPa enrolled in the Pons et al. study.[7] This was an international cohort of patients from different centers across Europe and Canada. Two-thirds of the cohort consisted of patients with biopsy-proven NASH enrolled prospectively between 2016 and 2018. The remaining cohort consisted of retrospective data from previously published databases.

Predicted outcome

The outcome of interest was CSPH at screening/baseline defined as an HVPG ≥ 10 mm Hg. Per the Emricasan protocol, HVPG was performed by experienced teams following standard guidelines.[8] Tracings were read independently by a senior experienced investigator (GGT).

Models

We first externally validated the ANTICIPATE and ANTICIPATE-NASH models.

The ANTICIPATE model includes LSM and platelet count to predict the presence of CSPH in patients with compensated cirrhosis[6] and uses the following formula: Logit = −6.3320165 + 3.1001014 × ln(LSM) – 0.020481545 × Platelet count, with LSM expressed in kilopascals and with platelet count values > 150 introduced as 150.

The ANTICIPATE-NASH model includes LSM, platelet count, and BMI to predict the presence of CSPH in patients with compensated NASH cirrhosis.[7] It uses the following formula: Logit = −3.9529402 + 2.2835809 × log(LSM) – 0.033777725 × BMI – 0.014490895 × Platelet count, with LSM expressed in kPa.

We then developed a model using two markers that had been shown to predict either CSPH or decompensation in patients with cirrhosis (mostly hepatitis C virus or alcohol),[1,9–11] which consisted of serum albumin and Fibrosis-4 index (FIB-4).[12] Serum albumin has been found to be predictive of decompensation in two separate cohorts.[1,9] FIB-4,[12] which includes age, aspartate aminotransferase (AST), alanine aminotransferase (ALT), and platelet count (using the following formula: (Age [year] × AST [U/L]) / (platelet count [10^9/L]) × (ALT [U/L])$^{1/2}$) has been shown to predict CSPH[11] and decompensation.[10] Our team also validated FIB-4 for prediction of CSPH in two separate cohorts using individual patient data from two previously published studies[3,13] (validation data available upon request).

Variables

LSM was measured by TE (FibroScan) using either M or XL probes. Screening demographic and laboratory data collected for the study were used. We defined obesity as BMI ≥ 30.

Statistical analysis

For the first aim, we assessed the performance of the ANTICIPATE and the ANTICIPATE-NASH models by evaluating the model's calibration and discrimination in our cohort using the methodology described by Moons et al.[14] Calibration is the agreement between observed and predicted risks.[14] Calibration was assessed by (1) "calibration-in-the-large," which compares the average predicted risk with the overall event rate (i.e., when the average predicted risk is higher than the overall event rate, the model overestimates risk in general; and when it is lower, the model underestimates the risk); (2) calibration plot, which plots the observed and predicted risks across deciles of predicted risk, and a measure of goodness-of-fit assessed using the Hosmer and Lemeshow test (a p value > 0.05 or failure to reject the null hypothesis of adequate fit indicates insufficient evidence to conclude lack of fit or poor calibration); and (3) calibration intercept and slope, which was estimated by regressing the binary outcome on the log odds (fitted logit) of the calibration model. The calibration slope has a target value of one, and the calibration intercept has a target value of zero.[15]

Discrimination refers to the model's ability to differentiate between patients with different outcomes.[14] The discrimination was assessed via receiver operating characteristic (ROC) curves and their corresponding C-statistics. The two models were first validated in the entire cohort and then in the subset of obese patients with NASH.

We then developed a logistic regression model in which the outcome was CSPH and the predictors were predetermined and included FIB-4 and albumin. The model was internally validated and corrected for optimism by bootstrapping.[16] When bootstrapping, 500 data sets of the same size were generated by random selection with replacement from the main data set. The discrimination of the model was evaluated by ROC curve analysis, and the reported C-statistic. Calibration was evaluated by plotting the agreement of predicted and observed probabilities. We then assessed the diagnostic performance of our model in the validation cohort of 245 patients from the Pons et al. study. Diagnostic accuracy of the model was evaluated by using the Brier score, which is the mean difference between predicted and observed probabilities. The score ranges from 0 to 1, with smaller scores indicating higher accuracy. Finally, a nomogram was developed based on the final corrected logistic regression model. p values ≤ 0.05 were considered statistically significant. Data were analyzed using SAS version 9.4 (SAS Institute) and R Statistical Software (R Core Team [2020], R: A language and environment for statistical computing; R Foundation for Statistical Computing) using the rms (Harrell FE Jr. rms: Regression Modeling Strategies: R Package Version 5.1–2. 2018; https://CRAN.R-project.org/package5rms) and ggplot2 (Wickham H. ggplot2: Elegant Graphics for Data Analysis, 2016; Springer-Verlag) packages.

RESULTS

AIM 1: External validation of ANTICIPATE and ANTICIPATE-NASH models

A total of 222 patients were included in the analysis (Figure 1); 165 of 222 (74%) had CSPH and 171 of 222 (77%) were obese. CSPH was present in 129 of 171 (75%) obese patients and in 36 of 51 (71%) nonobese patients. Baseline demographics of patients are summarized in Table 1. The prevalence of CSPH was somewhat higher in our cohort (74%) compared with the ANTICIPATE study (66%),[6] and much higher than that observed in patients with NASH in the ANTICIPATE-NASH study (39%)[7] (Table S1).

ANTICIPATE model

The predicted probability of CSPH using the model in our population was 171 of 222 (77%), which was very close to the observed proportion of CSPH cases (164 of 222 [74%]); therefore the model showed good calibration in-the-large. The calibration intercept was not significantly different from 0 (−0.0353, 95% confidence interval [CI] −0.4760 to 0.4053), and the calibration slope was not significantly different from 1 (0.7713, 95% CI 0.5351 and 1.0075), all indicative of a good calibration. The *p* value for the goodness-of-fit test (Hosmer Lemeshow test) was 0.26, indicative of absence of poor calibration. The model discrimination was evaluated by ROC curves. The area under the curve (AUC) was 0.8213 (95% CI 0.7575–0.8850), indicating good discrimination (Figure 2).

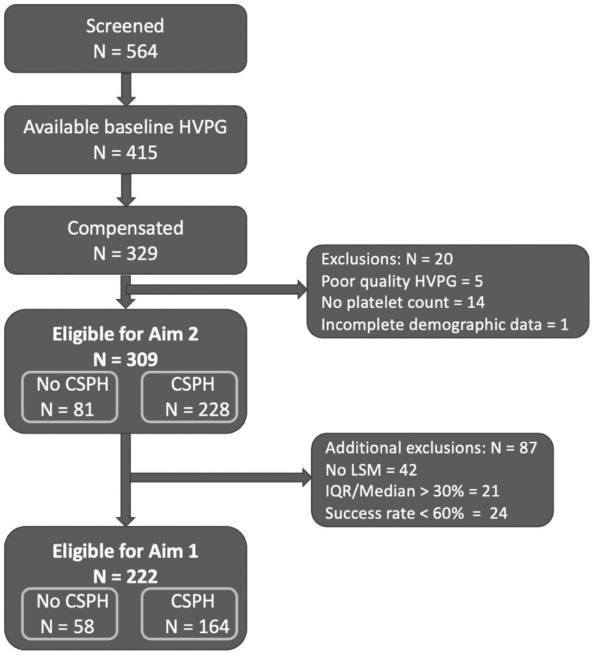

FIGURE 1 Study flow chart. Abbreviations: CSPH, clinically significant portal hypertension; HVPG, hepatic venous pressure gradient; IQR, interquartile range; LSM, liver stiffness measurement.

Given the performance of the model, an update did not seem necessary; however, with a small change to the baseline log odds (adding −0.2781 to the intercept), the calibration in-the-large was improved to predict CSPH in 164 of 222 (74%), which was the same as observed (164 of 222). It is important to note, however, that calibration in-the-large does not necessarily mean that the model was able to predict the exact same patients who had the outcome, and although the numbers are identical, the model could be identifying a patient who does not have CSPH as having CSPH or vice versa.

We also checked the calibration of the model in the obese subpopulation (n = 171; Figure S1) and in three different BMI categories (nonobese [BMI < 30; n = 51], obesity class I [BMI between 30 and 35; n = 86], and obesity classes II and III [BMI ≥ 35; n = 85]). The results are summarized in Figure S2.

TABLE 1 Baseline characteristics of the cohort for Aim 1: Validation of ANTICIPATE models

Median (IQR) or n (%)[a]	No CSPH (n = 58)	CSPH (n = 164)	Total (n = 222)	*p* value
Age (years)	60 (54–68)	62 (56–68)	62 (56–68)	0.41
Gender (female)	29 (50%)	94 (57%)	123 (55%)	0.34
Race (Caucasian)[b]	49 (85%)	146 (90%)	195 (88%)	0.09
Type 2 diabetes	36 (62%)	123 (75%)	159 (72%)	0.06
NSBB use	6 (10%)	49 (30%)	55 (25%)	0.003
Obese (BMI ≥ 30)	43 (74%)	128 (78%)	171 (77%)	0.54
BMI (kg/m^2)	32 (30–36)	34 (31–38)	33 (30–38)	0.15
Albumin (g/L)	4.4 (4.0–4.5)	4.1 (3.8–4.4)	4.1 (3.8–4.4)	0.001
ALT (U/L)	44 (26–51)	33 (25–45)	35 (25–48)	0.013
Creatinine (mg/dl)	0.7 (0.6–0.9)	0.6 (0.5–0.8)	0.7 (0.5–0.8)	0.013
Platelets (K/mm^3)	144 (110–179)	95 (71–125)	104 (75–143)	<0.001
INR[b]	1.1 (1.0–1.2)	1.1 (1.1–1.2)	1.1 (1.1–1.2)	<0.001
AST (U/L)	41 (30–57)	44 (33–57)	43 (32–57)	0.30
MELD-Na score	7 (6–8)	8 (7–10)	8 (7–9)	<0.001
CTP score[b]	5 (5–5)	5 (5–5)	5 (5–5)	0.1606
LSM (kPa)	21 (12–27)	34 (23–48)	28 (21–42)	<0.001

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; CSPH, clinically significant portal hypertension; CTP, Child-Turcotte-Pugh score; INR, international normalized ratio; MELD-Na, Model for End-Stage Liver Disease–Sodium; NSBB, nonselective beta blocker.

[a]Continuous variables are summarized using median (IQR), and categorical variables are summarized using the number of patients (percent of total).

[b]Missing values: Race not available for 1 patient in the no-CSPH group; INR was not available in 2 patients in the CSPH group; CTP was not available in 1 patient in the no-CSPH group and 4 patients in the CSPH group.

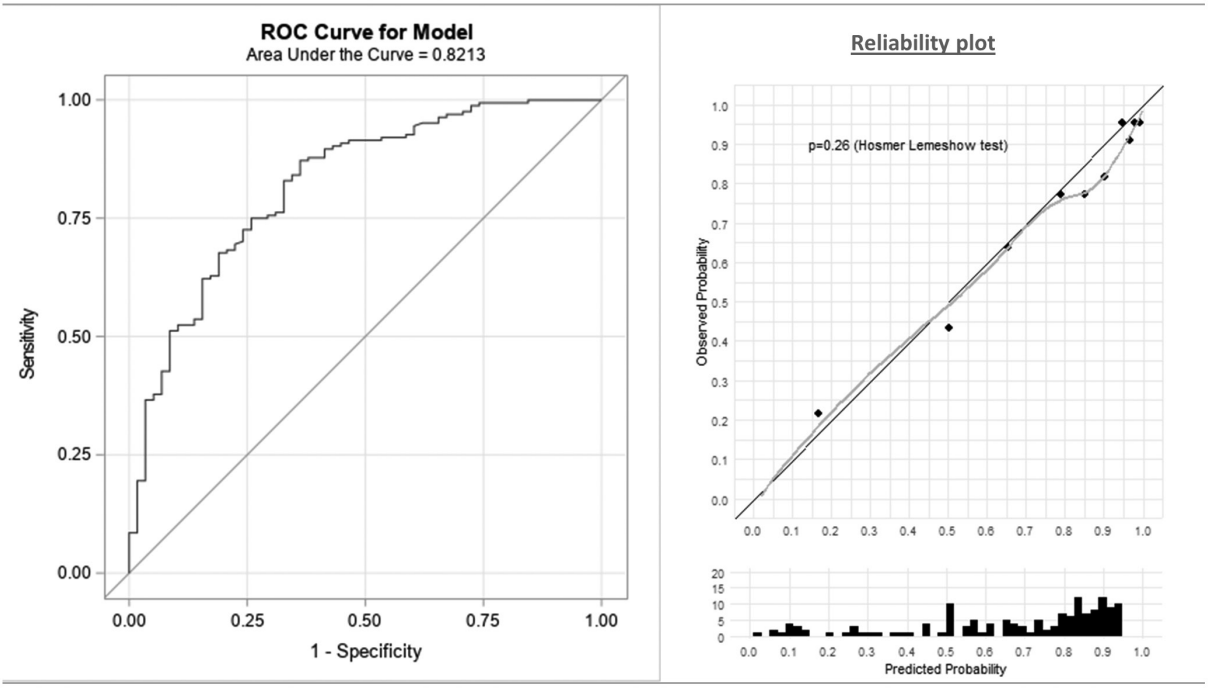

FIGURE 2 ANTICIPATE model performance in the entire cohort (n = 222). Left panel: Receiver operating characteristic (ROC) curves showing the discriminative value of the ANTICIPATE model. The 45-degree line is the line of no discrimination. Right panel: Calibration plots. The 45-degree line denotes the perfect agreement between predicted and observed risk. The gray line approximates the agreement between predicted and observed risks across subgroups of participants ranked by increasing predicted risks. The rug histogram shows the distribution of patients' risks in the cohort.

ANTICIPATE-NASH model

Using the ANTICIPATE-NASH model in our population, the predicted probability of CSPH was 151 of 222 (68%), as opposed to the observed proportion of CSPH, which was 164 of 222 (74%); therefore, the model slightly underestimated the presence of CSPH. The calibration intercept was 0.3735 (95% CI −0.00496 to 0.7520), and the calibration slope was 1.0301 (95% CI 0.7127 and 1.3475), which indicates good calibration. The *p* value from the goodness-of-fit test (Hosmer Lemeshow test) was 0.51, which is indicative of absence of poor calibration. The model also showed good discrimination as evaluated by ROC curves. The AUC was 0.8182 (95% CI 0.7563–0.8802) (Figure S3).

Given the performance of the model, an update did not seem necessary; however, with a small change to the baseline log-odds (adding +0.0131 to the intercept), the calibration in-the-large was improved to predict 164 of 222 (74%).

We then applied the ANTICIPATE-NASH model to only the obese subpopulation (n = 171). The predicted probability of CSPH was 115 of 171 (67%) compared with the observed proportion of CSPH, which was 128 of 171 (75%); therefore, the calibration in-the-large shows that the model slightly underestimated the probability of CSPH. The calibration intercept was 0.5039 (95% CI 0.0824–0.9253) and calibration slope was 0.9942 (95% CI 0.6462–1.3423), which indicate good calibration. The *p* value for the goodness-of-fit test (Hosmer Lemeshow test) was 0.21, which is indicative of absence of poor calibration. The model's discrimination was good with an AUC of 0.8174 (95% CI 0.7445–0.8903) (Figure 3).

AIM 2: Development of a model using only laboratory values

Using prespecified parameters including FIB-4 (age [years] × AST [U/L])/(platelet count [10^9/L]) × (ALT [U/L]$^{1/2}$) and serum albumin, which are easily available using lab tests, we developed a model to predict CSPH (called “FIB4+”). Because this lab-based model does not use LSM, we were able to include all 309 patients screened/enrolled in the trial, even those without LSM or those with poor-quality LSM (Figure 1). Baseline demographics of the study participants for Aim 2 are summarized in Table 2. The prevalence of CSPH was 74%.

In the lab-based model, albumin did not reach statistical significance, however we chose to retain it, given its clinical importance in the prediction of decompensation.[1,9] The model showed good discrimination with an AUC of 0.78 (95% CI 0.7193–0.8343) in the training cohort (Figure S4). We then internally validated the model by the bootstrapping method (500 replications). (The model formula and the bootstrapping validation are shown in Figure S4). We additionally externally validated

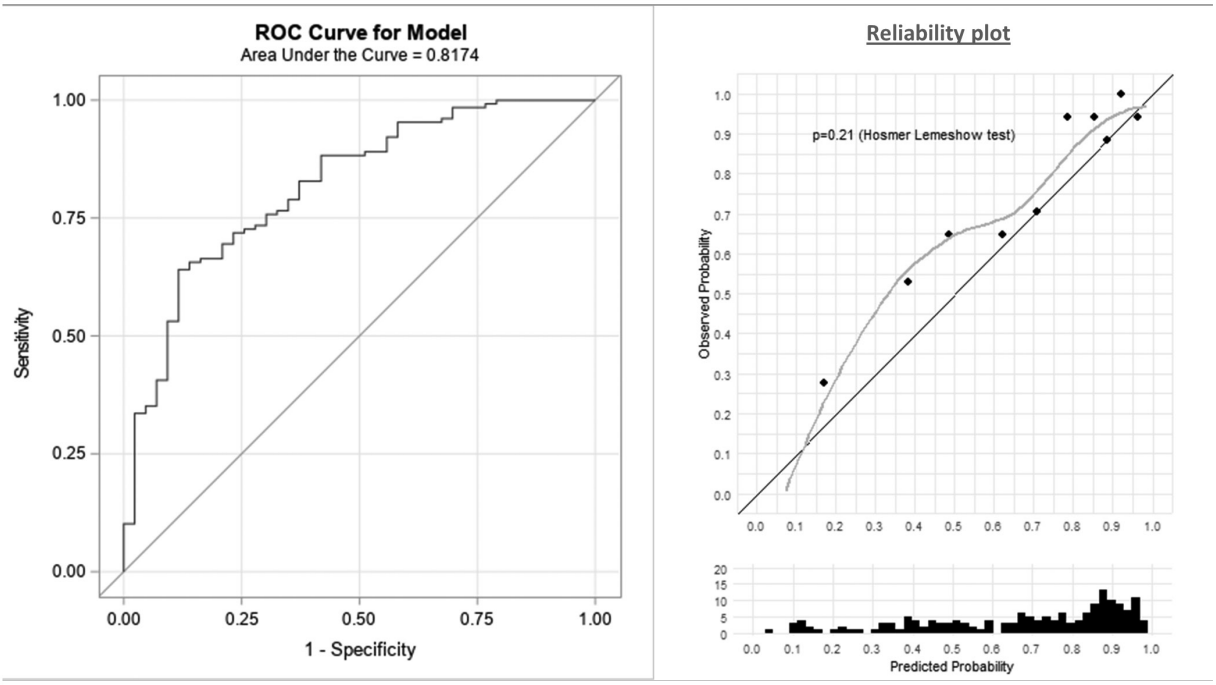

FIGURE 3 ANTICIPATE–NASH performance in the obese subpopulation (n = 171). Left panel: Receiver operating characteristic (ROC) curves showing the discriminative value of the ANTICIPATE-NASH model. The 45-degree line is the line of no discrimination. Right panel: Calibration plots. The 45-degree line denotes the perfect agreement between predicted and observed risk. The gray line approximates the agreement between predicted and observed risks across subgroups of participants ranked by increasing predicted risks. The rug histogram shows the distribution of patients' risks in the cohort.

TABLE 2 Baseline characteristics of the cohort for Aim 2: Development of FIB4+ model

Median (IQR) or n (%)[a]	No CSPH (n = 81)	CSPH (n = 228)	Total (n = 309)	*p* value
Age (years)	60 (53–67)	62 (56–68)	62 (56–68)	0.32
Gender (female)	43 (53%)	128 (56%)	171 (55%)	0.63
Race (Caucasian)[b]	72 (89%)	208 (92%)	280 (91%)	0.09
Type 2 diabetes	52 (64%)	176 (77%)	228 (74%)	0.022
NSBB use	11 (14%)	71 (31%)	82 (27%)	0.002
Obese (BMI≥30)	61 (75%)	178 (78%)	239 (77%)	0.61
BMI (kg/m^2)	32 (30–37)	34 (31–39)	34 (30–39)	0.06
Albumin (g/L)	4.4 (4.1–4.5)	4.0 (3.7–4.4)	4.1 (3.8–4.4)	<0.001
ALT (U/L)	43 (24–51)	33 (24–46)	35 (24–48)	0.048
Creatinine (mg/dl)	0.7 (0.6–0.8)	0.6 (0.5–0.8)	0.6 (0.5–0.8)	0.035
Platelets (K/mm^3)	141 (104–179)	94 (71–124)	102 (75–140)	<0.001
INR[b]	1.1 (1–1.1)	1.1 (1.1–1.2)	1.1 (1–1.2)	<0.001
AST (U/L)	39 (27–56)	43 (32–56)	41 (31–56)	0.040
MELD-Na score	7 (6–8)	8 (7–10)	8 (7–9)	<0.001
CTP score[b]	5 (5–5)	5 (5–5)	5 (5–5)	0.0126
LSM (kPa)[b]	21 (12–28)	32 (22–48)	28 (19–41)	<0.001

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; CSPH, clinically significant portal hypertension; CTP, Child-Turcotte-Pugh score; INR, international normalized ratio; MELD-Na, Model for End-Stage Liver Disease–Sodium; NSBB, nonselective beta blocker.

[a]Continuous variables are summarized using median (IQR), and categorical variables are summarized using the number of patients (percent of total).

[b]Missing values: Race not available for 1 patient in the no-CSPH group; LSM was not available for 14 patients in the no-CSPH and in 27 patients in the CSPH group; INR was not available in 4 patients in the CSPH group; CTP was not available in 1 patient in the no-CSPH group and 4 patients in the CSPH group.

our model in 245 patients with NASH cACLD from the Pons et al. group. The model showed good discrimination with AUC of 0.8039 (95% CI 0.7480–0.8599) in the validation cohort (Figure 4). The model's initial reliability plot in the validation cohort grossly overestimated the presence of CSPH (predicted 63% vs. observed: 39%);

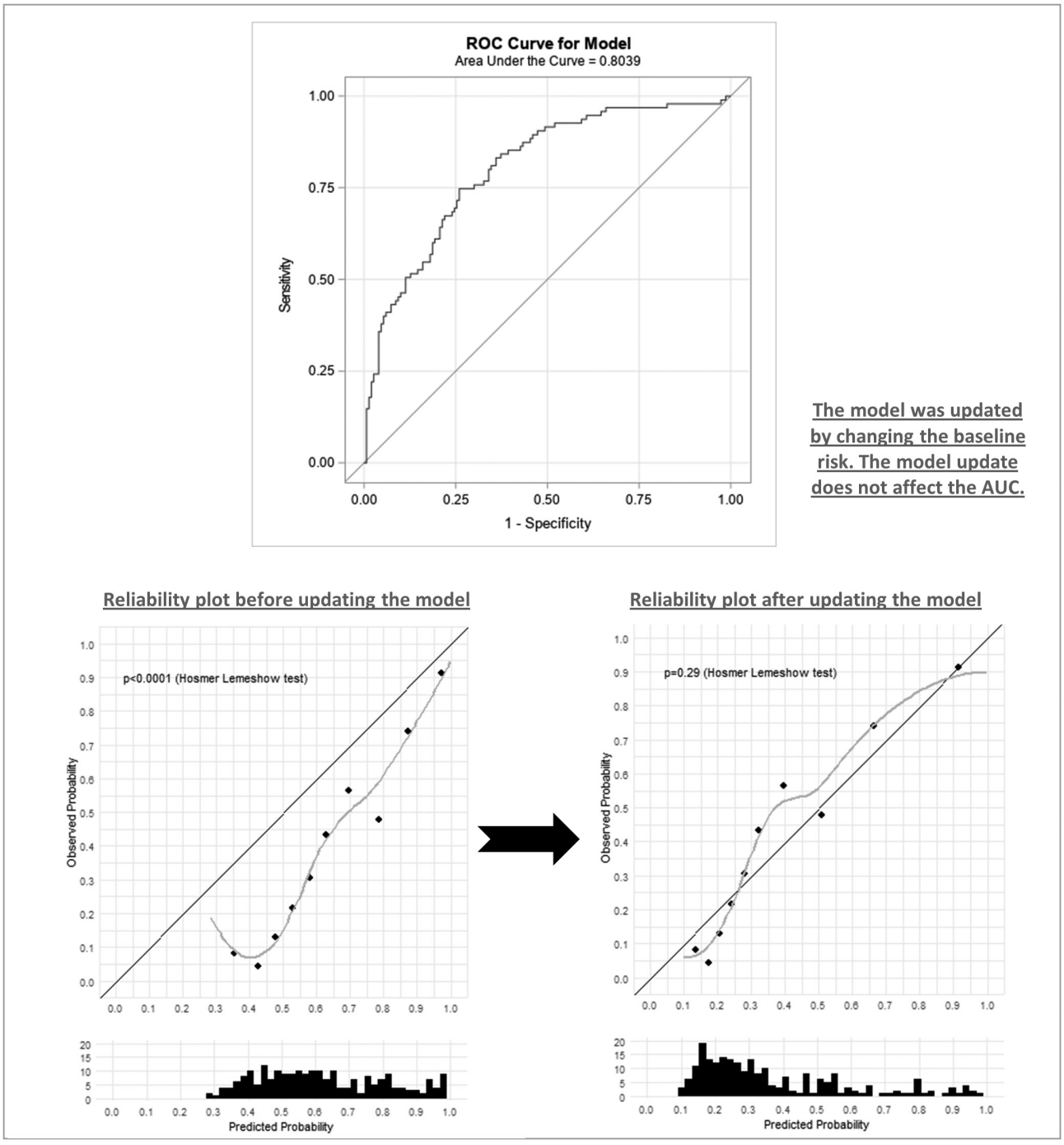

FIGURE 4 FIB4+ model performance in the validation cohort (n = 245). Top panel: Receiver operating characteristic (ROC) curves showing the discriminative value of the FIB4+ model. The 45-degree line is the line of no discrimination. Bottom-left panel: Calibration plots before updating the model. The 45-degree line denotes the perfect agreement between predicted and observed risk. The gray line approximates the agreement between predicted and observed risks across subgroups of participants ranked by increasing predicted risks. The rug histogram shows the distribution of patients' risks in the validation cohort. Bottom-right panel: Calibration plots after updating the model. Abbreviation: AUC, area under the curve.

however, the overestimation was consistent across all risk categories (Figure 4). Therefore, the lab-based model was updated by adding −1.2621 to the baseline risk to account for the difference in the prevalence of CSPH (74% vs. 39% in the derivation and validation cohorts, respectively), resulting in a good calibration. The calibration in-the-large in the updated lab-based model was 95 of 248 (38%), very close to the observed 97 of 248 (39%). The calibration intercept after the update was −0.0212 (95% CI −0.3584 to 0.3161), and calibration slope (before and after the update) was 0.9610 (95% CI 0.6465–1.2754), which indicates good

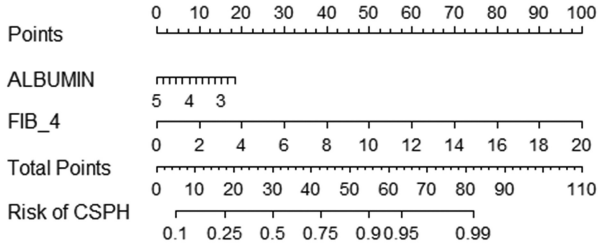

FIGURE 5 FIB4+ model nomogram

calibration. The *p* value for the Hosmer Lemeshow goodness-of-fit test was 0.29, which also indicates no evidence of poor calibration. The Brier score for the updated FIB4+ model in the validation cohort was 0.169, which indicates that the model provides accurate predictions for the presence of CSPH. The final calibrated FIB4+ model uses the following formula: log odds (CSPH) = 0.7207 – (0.6729 × albumin) + (0.4408 × FIB-4). We then developed a nomogram using the updated FIB4+ model, which estimates the risk of having CSPH based on each patient's individual data (Figure 5).

DISCUSSION

Noninvasive methods for detection of CSPH in patients with cirrhosis are critical, as the paradigm is shifting toward starting treatment with NSBBs once CSPH is detected, to prevent decompensation.[17,18]

Aim 1: Validation of ANTICIPATE models

In the validation study by Pons et al., although the ANTICIPATE model had excellent performance in patients with viral, alcohol, and nonobese NASH etiologies, when it was applied to the obese NASH population it overestimated the presence of CSPH. NASH population has been challenging for many reasons. It has been underrepresented in all previous studies, and there are specific aspects of this population, such as technical difficulty with performing elastography[19] and possible increase in liver stiffness with severe steatosis and steatohepatitis, which makes accurate assessment of LSM challenging.[20]

Our study was able to externally validate both the ANTICIPATE and ANTICIPATE-NASH models. The discrimination of both models was >0.8 in the entire cohort and in the obese subpopulation. In terms of calibration, the ANTICIPATE model slightly overestimated the presence of CSPH (calibration in the large: expected event rate 77% vs. 74% observed event rate), but the ANTICIPATE-NASH slightly underestimated the presence of CSPH in our cohort (calibration in the large: expected event rate 68% vs. 74% observed event rate). There are two potential explanations for the difference in performance of the ANTICIPATE models in our cohort as compared with the Pons et al. cohort. First, this could be related to a higher prevalence of CSPH in our cohort, which was 74% compared with 66% in the ANTICIPATE study,[6] and 39% in the NASH population of the ANTICIPATE-NASH study.[7] This higher prevalence is likely due to the fact that patient enrollment in our cohort was skewed toward patients with a higher likelihood of having CSPH (to meet the trial eligibility) before undergoing HVPG measurement. Nevertheless, in prior series of patients with compensated cirrhosis from multiple etiologies, the prevalence of CSPH has been reported to range from 63% to 73%.[3,21–25] Second, it could be that when the population with NASH presents with very advanced disease (compensated cirrhosis with high HVPG), the correlation of LSM and HVPG behaves more like other etiologies of cirrhosis.

Although the nomogram from the original studies provides a more accurate assessment of each individual patient's risk of having CSPH, Baveno consensus has proposed simplified prediction cutoffs to rule in CSPH in patients with viral or alcohol-related etiologies for a greater applicability in clinical practice. These cutoffs include LSM ≥ 25 kPa, LSM 20–25 kPa with platelet count ≤ 150K, and LSM 15–20 kPa with platelet count ≤ 110K.[18,26] However, Pons et al. cautioned against the use of a single cutoff of LSM ≥ 25 kPa in patients with NASH, particularly in patients with obesity, as it significantly overestimated the presence of CSPH in their population. In their study, only 63% of patients with obese NASH with LSM ≥ 25 actually had CSPH, whereas in patients with cirrhosis of viral or alcohol-related etiologies, more than 90% of patients with this cutoff had CSPH. In our cohort, CSPH was present in 114 of 132 (86%) of patients with LSM ≥ 25 in the entire cohort and in 91 of 105 (87%) of the obese subpopulation. Among patients with LSM 20–25 kPa and platelet count of ≤150K, CSPH was present in 24 of 32 (75%), and among patients with LSM 15–20 kPa and platelet count ≤110K, CSPH was present in 12 of 17 (71%) of patients. Nevertheless, and similar to the Pons et al. study, the performance of a single cutoff of LSM ≥ 25 kPa was fairly poor in our cohort of patients with NASH with an AUC of only 0.69, sensitivity of 70%, specificity of 69%, positive predictive value (PPV) of 86%, and negative predictive value (NPV) of 44% (Figure S5). Interestingly, despite the high prevalence of CSPH in our cohort, which would significantly increase the PPV of a test, the PPV was only 86%.

In summary, per the first aim of our study, we were able to validate both the ANTICIPATE and ANTICIPATE-NASH models in our cohort of NASH cirrhosis, with and without obesity. Although we performed a minor update to the intercept to fit our cohort, we still recommend using the original ANTICIPATE and ANTICIPATE-NASH formula/nomogram, as these cohorts are more representative of patients seen in

clinical practice compared with those selected for a randomized controlled trial.[6,7]

Aim 2: Development of FIB4+ model

Because many centers in the United States and worldwide do not have the resources to acquire a TE device to perform LSMs, we considered it of great importance to develop a model that would forego LSMs and would use readily available laboratory test results. Similar to LSM, the FIB-4 has been used to determine the presence of advanced liver fibrosis/cirrhosis in patients with chronic liver disease. FIB-4 has also been shown to be a strong predictor of presence of varices (an indicator of CSPH) in patients with compensated cirrhosis.[27] This is not surprising, as, in its formula it includes AST (a marker of endothelial dysfunction) and platelet count, which according to the ANTICIPATE study was the second-most important predictor of CSPH (after LSM) and could solely predict CSPH.[28] FIB-4 has been used previously in patients with other etiologies of cirrhosis for prediction of CSPH.[11] Additionally, FIB-4 has been used to predict clinical outcomes and is an important predictor of decompensation in patients with compensated cirrhosis.[10]

In fact, in an as-yet-unpublished study in which our group analyzed data obtained from the timolol study, in which the prevalence of CSPH was 63% (139 of 219)[13] and validated it in a second cohort,[3] we found that FIB-4 had a very good discriminative ability and calibration for prediction of CSPH. We added serum albumin because, after HVPG, it was the most important predictor of the development of decompensation in patients with compensated cirrhosis.[1,9]

Although in the training cohort the addition of albumin to the model did not change its discriminatory ability, in the validation cohort the model including albumin performed better (AUC = 0.794 [FIB-4 alone] vs. 0.8039 [FIB4+]). Another reason for keeping the albumin in the final model is its potential for prediction of outcomes.[9] Because predicting CSPH is a surrogate in the prediction of cirrhosis decompensation, we think that maintaining albumin in the model will strengthen its ability to predict outcomes, although this needs to be further investigated in future prospective studies.

While the discrimination of our model was 0.8 in the validation cohort, calibration was not good, and the model overestimated the presence of CSPH in all risk categories (predicted risks were systematically too high but remained proportionally accurate). By adjusting the baseline risk of our original prediction model to that of individuals included in the validation cohort, the calibration was easily improved without changing its discriminatory ability. It is important to mention that no other updates to the model (such as change in the coefficient of each variable or addition of a new variable) were needed to improve its calibration. Because, as mentioned previously, the validation cohort (Pons et al. cohort) is more representative of patients with NASH seen in the outpatient clinics compared with those in our training cohort (preselected for a RCT), we recommend the use of the updated model, which we call the FIB4+ model in clinical practice. Based on our FIB4+ nomogram (Figure 5), patients with FIB-4 ≥ 6 and albumin < 3.6 have a risk of CSPH > 75%. To facilitate this calculation, we have created an online calculator that includes the five variables of the FIB4+ model (http://www.fib4plus.com).

Strengths and limitations of our study

The strengths of our study are the inclusion of a large number of patients with NASH cirrhosis, which is currently the most common etiology of chronic liver disease,[29] with a significant proportion of patients with CSPH, which allowed for the validation of previous models. The development of an inexpensive model that can be applied in any center and the external validation of this model in a large sample of patients from a very diverse setting is also a major strength.[30] The fact that our FIB4+ model performed well in a separate cohort from different institutions/centers and with very different baseline characteristics suggests that our model can be generalizable to other populations of patients with compensated NASH cirrhosis. Another major strength of our study is evaluation of the performance characteristics of previous models and our own model, not only by discrimination (ROC curves and their corresponding C-statistics) but also by detailed evaluation of different aspects of calibration (calibration in the large or the calibration intercept and the calibration slope).

Our study has several limitations. First, the data used for our study were collected as part of screening for enrollment of patients in a RCT. The reason for higher prevalence of CSPH in our cohort is that enrollment was skewed toward patients with higher likelihood of having CSPH, before undergoing HVPG measurement. Given the slow recruitment (a large proportion of patients undergoing HVPG measurement did not meet the criteria for study inclusion [HVPG ≥ 12]), the study added another inclusion criterion, which was a platelet count ≤ 125K or LSM ≥ 20 kPa during screening. This change in protocol was implemented halfway through the study and occurred at different times at each study center due to the internal review board approval process; therefore, we could not identify the patients enrolled before or after this change in protocol. This skewed the population toward patients with higher pretest probability of CSPH. Although this could affect the PPV and NPV,

the baseline prevalence does not affect sensitivity and specificity, and thus the overall discrimination of the model. Furthermore, our model still showed acceptable performance characteristics in a separate validation cohort with a significantly lower prevalence of CSPH. Second, the patients could have been on NSBBs, which could have artificially decreased the portal pressure gradient. Although NSBBs were held 24 h before HVPG measurement, this could have still decreased the portal pressures and misclassified them as not having CSPH, while in fact their HVPG would have been ≥10 if measured while not being on NSBBs. Finally, the updated FIB4+ model requires further validation in a separate cohort of patients with NASH cirrhosis to demonstrate generalizability.

Conclusions

Our study showed that both ANTICIPATE and ANTICIPATE-NASH models perform very well in patients with NASH cirrhosis, with good calibration and discrimination. The ANTICIPATE-NASH model also performed well in the subset of patients with obesity. We were also able to develop a simple model using FIB-4 and albumin to predict CSPH in patients with compensated NASH cirrhosis and externally validate and update the model in a separate validation cohort (Pons et al. cohort). The use of our final updated model, FIB4+, will be particularly useful for centers in the world that lack TE to perform LSM.

ACKNOWLEDGMENT

The authors thank the following investigators for kindly sharing their data from the ANTICIPATE-NASH study: Salvador Augustin, Bernhard Scheiner, Maeva Guillaume, Matteo Rosselli, Susana G. Rodrigues, Horia Stefanescu, Mang M. Ma, Mattias Mandorfer, Mayka Mergeay-Fabre, Bogdan Procopet, Philipp Schwabl, Arnulf Ferlitsch, Georg Semmler, Annalisa Berzigotti, Emmanuel Tsochatzis, Christophe Bureau, Thomas Reiberger, Jaime Bosch, and Juan G. Abraldes.

FUNDING INFORMATION

Supported by the Yale Liver Center, National Institutes of Health (P30 DK34989).

CONFLICT OF INTEREST

Nothing to report.

ORCID

Anahita Rabiee https://orcid.org/0000-0002-8833-3592
Joan Genesca https://orcid.org/0000-0002-0831-8422

REFERENCES

1. Ripoll C, Groszmann R, Garcia-Tsao G, Grace N, Burroughs A, Planas R, et al. Hepatic venous pressure gradient predicts clinical decompensation in patients with compensated cirrhosis. Gastroenterology. 2007;133:481–8.
2. Villanueva C, Albillos A, Genescà J, Garcia-Pagan JC, Calleja JL, Aracil C, et al. β blockers to prevent decompensation of cirrhosis in patients with clinically significant portal hypertension (PREDESCI): a randomised, double-blind, placebo-controlled, multicentre trial. Lancet. 2019;393:1597–608.
3. Berzigotti A, Seijo S, Arena U, Abraldes JG, Vizzutti F, García–Pagán JC, et al. Elastography, spleen size, and platelet count identify portal hypertension in patients with compensated cirrhosis. Gastroenterology. 2013;144:102–11.e1.
4. Bureau C, Metivier S, Peron JM, Selves J, Robic MA, Gourraud PA, et al. Transient elastography accurately predicts presence of significant portal hypertension in patients with chronic liver disease: investigating the performance of transient elastography. Aliment Pharmacol Ther. 2008;27:1261–8.
5. Lemoine M, Katsahian S, Ziol M, Nahon P, Ganne-Carrie N, Kazemi F, et al. Liver stiffness measurement as a predictive tool of clinically significant portal hypertension in patients with compensated hepatitis C virus or alcohol-related cirrhosis. Aliment Pharmacol Ther. 2008;28:1102–10.
6. Abraldes JG, Bureau C, Stefanescu H, Augustin S, Ney M, Blasco H, et al. Noninvasive tools and risk of clinically significant portal hypertension and varices in compensated cirrhosis: the "Anticipate" study. Hepatology. 2016;64:2173–84.
7. Pons M, Augustin S, Scheiner B, Guillaume M, Rosselli M, Rodrigues SG, et al. Noninvasive diagnosis of portal hypertension in patients with compensated advanced chronic liver disease. Am J Gastroenterol. 2021;116:723–32.
8. Garcia-Tsao G, Bosch J, Kayali Z, Harrison SA, Abdelmalek MF, Lawitz E, et al. Randomized placebo-controlled trial of emricasan for non-alcoholic steatohepatitis-related cirrhosis with severe portal hypertension. J Hepatol. 2020;72:885–95.
9. Ripoll C, Bari K, Garcia-Tsao G. Serum albumin can identify patients with compensated cirrhosis with a good prognosis. J Clin Gastroenterol. 2015;49:613–9.
10. Guha IN, Harris R, Berhane S, Dillon A, Coffey L, James MW, et al. Validation of a model for identification of patients with compensated cirrhosis at high risk of decompensation. Clin Gastroenterol Hepatol. 2019;17:2330–8.e1.
11. Procopet B, Cristea VM, Robic MA, Grigorescu M, Agachi PS, Metivier S, et al. Serum tests, liver stiffness and artificial neural networks for diagnosing cirrhosis and portal hypertension. Dig Liver Dis. 2015;47:411–6.
12. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology. 2006;43:1317–25.
13. Groszmann RJ, Garcia-Tsao G, Bosch J, Grace ND, Burroughs AK, Planas R, et al. Beta-blockers to prevent gastroesophageal varices in patients with cirrhosis. N Engl J Med. 2005;353:2254–61.
14. Moons KGM, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012;98:691–8.
15. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, on behalf of Topic Group "Evaluating diagnostic tests and prediction models" of the STRATOS initiative. Calibration: the Achilles heel of predictive analytics. BMC Med. 2019;17:230.
16. Steyerberg EW, Harrell FE Jr, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models. J Clin Epidemiol. 2001;54:774–81.

17. Garcia-Tsao G, Abraldes JG. Nonselective beta-blockers in compensated cirrhosis: preventing variceal hemorrhage or preventing decompensation? Gastroenterology. 2021;161:770–3.
18. de Franchis R, Bosch J, Garcia-Tsao G, Reiberger T, Ripoll C, Abraldes JG, et al. Baveno VII—renewing consensus in portal hypertension. J Hepatol. 2022;76:959–74.
19. Castéra L, Foucher J, Bernard PH, Carvalho F, Allaix D, Merrouche W, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. Hepatology. 2010;51:828–35.
20. Petta S, Maida M, Macaluso FS, di Marco V, Cammà C, Cabibi D, et al. The severity of steatosis influences liver stiffness measurement in patients with nonalcoholic fatty liver disease: steatohepatitis/metabolic liver disease. Hepatology. 2015;62:1101–10.
21. Villanueva C, Albillos A, Genescà J, Abraldes JG, Calleja JL, Aracil C, et al. Development of hyperdynamic circulation and response to β-blockers in compensated cirrhosis with portal hypertension: liver failure/cirrhosis/portal hypertension. Hepatology. 2016;63:197–206.
22. Reiberger T, Ferlitsch A, Payer BA, Pinter M, Schwabl P, Stift J, et al. Noninvasive screening for liver fibrosis and portal hypertension by transient elastography—a large single center experience. Wien Klin Wochenschr. 2012;124:395–402.
23. Augustin S, Millán L, González A, Martell M, Gelabert A, Segarra A, et al. Detection of early portal hypertension with routine data and liver stiffness in patients with asymptomatic liver disease: a prospective study. J Hepatol. 2014;60:561–9.
24. Colecchia A, Montrone L, Scaioli E, Bacchi–Reggiani ML, Colli A, Casazza G, et al. Measurement of spleen stiffness to evaluate portal hypertension and the presence of esophageal varices in patients with HCV-related cirrhosis. Gastroenterology. 2012;143:646–54.
25. Kitson MT, Roberts SK, Colman JC, Paul E, Button P, Kemp W. Liver stiffness and the prediction of clinically significant portal hypertension and portal hypertensive complications. Scand J Gastroenterol. 2015;50:462–9.
26. de Franchis R. Expanding consensus in portal hypertension. J Hepatol. 2015;63:743–52.
27. Deng H, Qi X, Guo X. Diagnostic accuracy of APRI, AAR, FIB-4, FI, King, Lok, Forns, and fibroindex scores in predicting the presence of esophageal varices in liver cirrhosis: a systematic review and meta-analysis. Medicine. 2015;94:e1795.
28. Abraldes JG, Garcia-Tsao G. Simple clinical tools to predict decompensation in patients with compensated cirrhosis: an unmet need. Clin Gastroenterol Hepatol. 2019;17:2179–81.
29. Moon AM, Singal AG, Tapper EB. Contemporary epidemiology of chronic liver disease and cirrhosis. Clin Gastroenterol Hepatol. 2020;18:2650–66.
30. Vergouwe Y, Steyerberg EW, Eijkemans MJC, Habbema JDF. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. J Clin Epidemiol. 2005;58:475–83.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Rabiee A, Deng Y, Ciarleglio M, Chan JL, Pons M, Genesca J, et al. Noninvasive predictors of clinically significant portal hypertension in NASH cirrhosis: Validation of ANTICIPATE models and development of a lab-based model. Hepatol Commun. 2022;6:3324–3334. https://doi.org/10.1002/hep4.2091