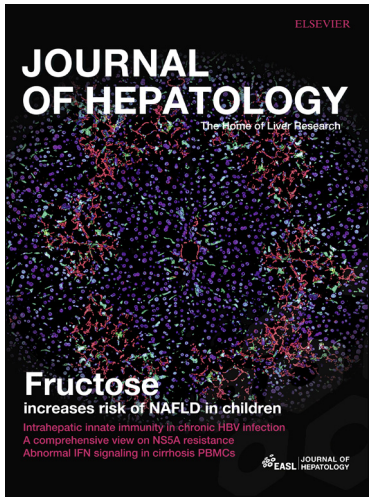
Accepted Manuscript



A combined model based on spleen stiffness measurement and Baveno VI criteria to rule out high risk varices in advanced chronic liver disease

Antonio Colecchia, Federico Ravaioli, Giovanni Marasco, Agostino Colli, Elton Dajti, Annarita Di Biase, Maria Letizia Bacchi Reggiani, Annalisa Berzigotti, Massimo Pinzani, Davide Festi

PII: S0168-8278(18)32034-8
DOI: https://doi.org/10.1016/j.jhep.2018.04.023
Reference: JHEPAT 6957

To appear in: *Journal of Hepatology*

Received Date: 8 August 2017
Revised Date: 14 April 2018
Accepted Date: 18 April 2018

Please cite this article as: Colecchia, A., Ravaioli, F., Marasco, G., Colli, A., Dajti, E., Biase, A.D., Reggiani, M.L.B., Berzigotti, A., Pinzani, M., Festi, D., A combined model based on spleen stiffness measurement and Baveno VI criteria to rule out high risk varices in advanced chronic liver disease, *Journal of Hepatology* (2018), doi: https://doi.org/10.1016/j.jhep.2018.04.023

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

A combined model based on spleen stiffness measurement and Baveno VI criteria to rule out high risk varices in advanced chronic liver disease

Antonio Colecchia[1,5,†,*], Federico Ravaioli[1,†], Giovanni Marasco[1], Agostino Colli[6], Elton Dajti[1], Annarita Di Biase [3], Maria Letizia Bacchi Reggiani[2], Annalisa Berzigotti[4], Massimo Pinzani[7], Davide Festi[1]

[1]Department of Medical and Surgical Sciences (DIMEC), University of Bologna, Italy; [2]Department of Experimental, diagnostic and specialty Medicine (DIMES), University of Bologna, Italy; [3]Department of Pediatrics, University of Modena, Modena, Italy; [4]Hepatology, Inselspital, University Clinic of Visceral Surgery and Medicine (UVCM), University of Bern, Switzerland; [5]UOC. Gastroenterologia, Azienda Ospedaliera Universitaria Integrata, Verona, Italy; [6] Department of Internal Medicine, General Hospital, Lecco, Italy; [7]Department of Hepatology, Royal Free Hospital NHS Trust; Institute for Liver and Digestive Health, University College London, London, Great Britain.

[†]Antonio Colecchia and Federico Ravaioli contributed to this manuscript equally.

Keywords: oesophageal varices; portal hypertension; liver stiffness measurement; high risk varices; oesophagogastroduodenoscopy spared.

*Corresponding authors: Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy; UOC. Gastroenterologia, Azienda Ospedaliera Universitaria Integrata, Verona, Italy. *Email address*: antonio.colecchia@aovr.veneto.it (A. Colecchia).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

List of Abbreviations: cACLD, compensated advanced chronic liver disease; CLD, chronic liver disease; CSPH, clinically significant portal hypertension; C-P, Child-Turcotte-Pugh; EGD, esophagogastroduodenoscopy; EV, oesophageal varices; GEV, gastroesophageal varices; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; HVPG, hepatic venous pressure gradient; INR, international normalized ratio; kPa, kilopascals; LS, liver stiffness; MELD, Model for End-Stage Liver Disease; NITs, non-invasive tests; NPV, negative predictive value; OD, odds ratios; PH, portal hypertension; PPV, positive predictive value; SS, spleen stiffness; SVR, sustained virological response; TE, transient elastography.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Background & Aims: Recently, Baveno VI guidelines suggested that esophagogastroduodenoscopy (EGD) can be avoided in patients with cACLD who have a liver stiffness measurement (LSM) < 20 kPa and platelet count >150,000/mm^3. We aimed to: assess the performance of spleen stiffness measurement (SSM) in ruling out patients with high-risk varices (HRV); validate Baveno VI criteria in a large population and assess how the sequential use of Baveno VI criteria and SSM could safely avoid the need for endoscopy.

Methods: We retrospectively analysed 498 cACLD patients who had undergone LSM/SSM by transient elastography (TE) (Fibroscan®), platelet count and EGDs from 2012 to 2016 referred to our tertiary centre. The new combined model was validated internally by a split-validation method, and externally in a prospective multicentre cohort of 115 patients.

Results: SSM, LSM, platelet count and Child-Pugh-B were independent predictors of HRV. Applying the newly identified SSM cut-off (≤46 kPa) or Baveno VI criteria, 35.8% and 21.7% of patients in the internal validation cohort could have avoided EGD, with HRV being missed in only 2 % in both cases. The combination of SSM with Baveno VI criteria would have led to additionally avoiding 22.5% of EGDs, reaching a final value of 43.8% spared EGDs, with <5% missed HRV. Results were confirmed in the prospective external validation cohort, as the combined Baveno VI/SSM≤46 Model would have safely spared (0 HRV missed) 37.4% of EGDs, compared to 16.5% avoiding Baveno VI Criteria only.

Conclusions: A non-invasive prediction model combining SSM with Baveno VI criteria may be useful to rule out HRV and could make it possible to avoid a significantly larger number of unnecessary EGDs compared to Baveno VI criteria only.

Lay summary:

Spleen stiffness measurement (SSM) assessed by TE, the most widely used electrographic technique, is a non-invasive technique that can help the physician to better stratify the degree of portal hypertension and the risk of oesophageal varices in patients with cACLD. Performing SSM

ACCEPTED MANUSCRIPT

together with LSM during the same examination is simple and fast and this sequential model can identify a greater number of patients that can safely avoid EGD, which is an invasive and expensive examination.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Introduction [H1]

The increase of portal pressure above the threshold of clinically significant portal hypertension (CSPH, Hepatic Vein Pressure Gradient (HVPG) ≥10 mmHg) is a landmark in the natural history of advanced chronic liver disease. Patients with CSPH are at risk of developing gastroesophageal varices (EV) and clinical decompensation (variceal bleeding, ascites, jaundice, encephalopathy), which mark the transition from a compensated stage to a stage of the disease (decompensated) characterized by a much higher mortality [1]. In particular, variceal bleeding is still associated with 10-15% mortality despite the advances in its treatment [2].

Since the risk of variceal bleeding mostly depends upon the size of EV and can be reduced with appropriate medical or endoscopic treatment in patients with high risk EV (HRV), endoscopic screening of EV is currently the standard of care in patients with cirrhosis [3].

Recommendations on the use of screening endoscopy (EGD) for the detection of HRV in all cirrhotic patients have been issued since the first Baveno consensus workshop in 1992 [4]. However, a large proportion of cirrhotic patients do not present HRV, thus making endoscopy a non-ideal screening test that is, on the other hand, associated with significant costs and patient discomfort [5]. Accordingly, in the last decade increased attention has been dedicated to identify sufficiently accurate non-invasive tests (NITs) able to rule in and rule out CSPH and HRV and thus to reduce or avoid the use of invasive methods such as HVPG measurement and EGD [6,7].

The recent 2015 Baveno VI consensus workshop [8] highlighted the diagnostic role of NITs such as liver stiffness measurement (LSM) in defining the presence of CSPH and HRV. In particular, the so-called Baveno VI criteria useful in excluding HRV were defined as follows: patients with LSM <20 kPa (assessed by transient elastography, TE) and a platelet count (PLT) >150,000/mm3 were considered very unlikely to have high-risk varices (<5%), and endoscopy could consequently be safely avoided. The above criteria can also be applied for longitudinal follow up, prompting screening endoscopy if LSM increases or PLT decreases [8]. To date, several papers have provided validation of these criteria [7,9–14] confirming that Baveno VI criteria correctly identify 98-100% of patients who could safely avoid endoscopy. These criteria have recently also

ACCEPTED MANUSCRIPT

been adopted by the American Association for the Study of the Liver [3] with a suggestion to use NIT stratification according to Baveno VI criteria in all patients with a new diagnosis of cirrhosis.

A possible limitation of the Baveno VI criteria [8] is related to a substantially low number of spared endoscopies (15-25%) [9,10]. Although different authors have tried to modify LSM and PLT cut-off (25 kPa or 100.000-120.000/mm^3), the rate of avoided EGDs did not exceed 30% [11,15]. Most patients not presenting HRV therefore still require endoscopy according to the Baveno criteria and it is becoming evident that more accurate non-invasive methodologies are needed to improve risk stratification [16].

In recent years, spleen stiffness measurement (SSM), carried out with various techniques [TE, Acoustic Radiation Force Impulse elastography (ARFI), 2D-shear wave elastography (SSI) and Magnetic Resonance Elastography (MRE)], has been proposed by several authors [17,18] as an accurate diagnostic tool to assess CSPH and to predict EV presence or absence. In a study performed by our group, SSM was more closely related to the severity of PH as compared to LSM [19], thus allowing a more accurate prediction of clinical decompensation, as documented in a subsequent study [20]. Furthermore, it has recently been observed that also by using 2D-SWE the sequential use of LSM and SSM leads to an improvement of the non-invasive diagnosis of CSPH as compared to LSM alone [21,22] and also in combination with Lok-score to diagnose HRV [23].

The present study, performed in a cohort of patients with compensated advanced chronic liver disease (cACLD), aimed to: a) assess the performance of SSM in ruling out HRV and avoiding EGD; b) validate Baveno VI criteria in a large clinical practice setting of tertiary referral centres, and c) assess whether the combined use of Baveno VI criteria and SSM can increase the number of safely avoided EGD.

ACCEPTED MANUSCRIPT

Patients and methods [H1]

Patients [H2]

Patients with chronic liver diseases referred to the Department of Medical and Surgical Sciences of the University of Bologna (Italy) for liver disease severity staging were consecutive enrolled from 2012 to 2016. The data collected, included LSM, SSM, laboratory tests and EGDs, were retrospectively analysed. From September 2017 to February 2018, patients with chronic liver diseases referred to our three centres (Bologna, Italy; Verona, Italy; Bern, Switzerland) were prospectively evaluated in order to externally validate the new combined model.

Inclusion criteria [H2]

The inclusion criteria of the retrospective cohort were: cACLD of any aetiology with LSM ≥10 kPa according to Baveno VI criteria [8], availability of non-invasive tests (laboratory tests and SSM) and EGDs performed within 3 months of TE assessment. In the prospective cohort, all patients with LSM>10 kPa were included. If no EGD was available within three months prior to TE evaluation, they were subsequently evaluated with upper endoscopy.

Exclusion Criteria [H2]

Exclusion criteria were the presence of decompensation events (ascites and hepatic encephalopathy), a previous episode of variceal bleeding and/or EV banding legation, current use of non-selective beta-blockers, portal vein thrombosis, transjugular intrahepatic portosystemic shunt (TIPS) and non-cirrhotic portal hypertension. In the retrospective cohort, a further exclusion criterion was ongoing or ending <6 months before the TE session antiviral treatment; while in the prospective cohort direct-acting antiviral (DAA) treatment always was excluded.

Liver and spleen stiffness measurements [H2]

ACCEPTED MANUSCRIPT

LSM values were assessed by expert operators, using the FibroScan® apparatus with “M” probe (Echosens, Paris, France) after overnight fasting and after a complete abdominal ultrasound (US) examination. LSM values were obtained as previously reported [19] and the reliability criteria considered were according to the last EFSUMB Guidelines and Recommendations on the Clinical Use of Ultrasound Elastography [24]; At least 10 measurements with an IQR/M ≤30 % were used as reliability criteria. SSM was assessed on the same day as LSM assessment, with the same probe utilized to perform LSM using the FibroScan® apparatus, as previously described [19,20].

Upper GI endoscopic examination [H2]

A standard endoscopic examination was performed by third-level hospital expert operators. The endoscopic findings were recorded and graded as follows: grade I: varices were flattened by insufflation; grade II: varices non-confluent and protruding in the lumen despite insufflation; grade III: confluent varices were not flattened by insufflation. The presence of red signs was also recorded in all patients. According to the criteria proposed at the Baveno VI Consensus Conference, HRV were defined as EV grade 1 with red signs or grade 2 or higher [8].

Statistical analysis [H2]

Statistical analyses were performed using Stata/SE (Version 14.2; Stata Corp, Texas, United States of America) for Windows; continuous variables are expressed as median and values of 25% and 75% percentiles; categorical data are expressed as numbers (percentages). For multivariate model development, we used split-sample method [25]. The derivation dataset consisted of 54 randomly selected cases (HRV) and 204 ~~129~~ randomly selected controls (no HRV) from the original datasets of 100 cases and 398 controls; consequently, the internal validation dataset includes 46 cases and 194 ~~123~~ controls. For group comparisons of categorical and continuous variables, chi square test or Mann-Whitney test and McNemar test were used, as appropriate. Univariate and multivariable logistic regression analyses were performed on the

ACCEPTED MANUSCRIPT

derivation dataset to evaluate the detection of HRV in cACLD. Demographic, clinical, biochemical and elastometric variables were considered as risk variables. LSM, PLTs and SSM were evaluated as continuous and dichotomous data: Baveno VI cut-off criteria [8] were adopted for LSM and PLTs whereas, in order to rule-out patients with HRV, the value corresponding to the negative likelihood ratio (LR-) closest to 0.05 was chosen for SSM cut-off. Multivariable logistic regression analyses were conducted on variables that reached $p<0.1$ at univariate analysis. The final multivariate logistic regression model was built from the set of candidate variables by removing predictors based on p values, in a stepwise manner. Multicollinearity amongst variables was detected by means of Spearman correlation test. For each multivariable logistic regression, the model discrimination and calibration were reported together with AIC (Akaike information criterion) and BIC (Bayesian information criterion) measures for comparing maximum likelihood models. Model discrimination was assessed calculating the Area under the Receiver Operator Characteristic (AUROC) curve (or c-statistic), whereas model calibration was determined by Hosmer-Lemeshow (H-L) technique and the goodness of fit of the models was assessed by AIC. We made an a priori decision to consider only models with AUROC c-statistic ≥ 0.75 adequate. To compare effects across nested logistic regression models the likelihood ratio (LR) test was used; to test the equality of the area under the curves, the DeLong test was performed [26]. The final multivariable logistic models were fitted on an internal and external validation dataset and on the entire retrospective population. In the multivariate model of the external validation cohort, the inter-center variability was taken into account. In order to determine the relative importance of independent variables in the multivariable logistic regression, a “Dominance Analysis” was performed [27]. Based on the multivariate logistic regression model with LSM, PLTs and SSM as continuous variables, a nomogram was drawn up. Sensitivity, specificity, likelihood ratio (LR) + and LR- were calculated at each cut-off and also using an intention to diagnose principle (ITD) [28], including non-evaluable results (SSM) either as “false negative” or as “false positive” according to the results of the reference standard (EGDs). All p values refer to two-tailed tests of significance. $P<0.05$ was considered significant.

ACCEPTED MANUSCRIPT

Ethics [H2]

This study was conducted in compliance with the Declaration of Helsinki and approved by the local institutional review board [Institutional Ethics Committee of Sant'Orsola-Malpighi University Hospital (Bologna, Italy)]; protocol number 2080/2017 (retrospective study) and 4163/2017 (prospective study). All patients provided written informed consent for participation in retrospective analysis.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Results [H1]

Patient characteristics [H2]

Over the study period, 643 ACLD patients fulfilled the inclusion criteria; 24 (3.7%) patients with non-cirrhotic portal hypertension, 10 (1.6%) with portal/mesenteric/splenic vein thrombosis, 18 (2.8%) current users of non-selective beta-blockers, 10 (1.6%) with previous EV banding ligation and 83 (12.9%) with hepatic decompensated disease were excluded. A total of 498 patients were finally included in the study cohort (**Figure 1**); of these, 26 (5.2%) patients had an inadequate SSM.

Two hundred and ninety-one (58.4%) out of 498 patients were male, the predominant aetiology of the liver disease was related to HCV (n= 424, 85.1%). Three hundred and fifty patients were Child Pugh (CP) A (70.3%), 148 were CP-B (29.7%) and the median Model for End-Stage Liver Disease (MELD) value was 8 (~~IQR~~ 7-10). LSM median was 21.3 kPa (~~IQR~~ 15-30.3) and median SSM was 54.3 kPa (~~IQR~~ 40 – 67.8). Two hundred and fifty-two (50.6%) out of 498 patients had EV ~~varices~~ and 100 of these had HRV (20.1%). According to the split-sample method – at baseline, the characteristics of the patients assigned to the derivation set (n=258) and to the internal validation set (n=240) were not statistically different. (**Table 1)**

Predictors of high risk varices [H2]

High risk varices (HRV) were observed in 100 patients (20.1%) of the whole population, of whom 54 (20,8%) were in the derivation cohort and 46 (19%) in the internal validation cohort. In the derivation cohort **(Table 2**), CP-B, MELD score, ALT levels, higher SSM values – higher LSM values and lower platelet count were correlated with the presence of HRV and considered in the multivariate models.

AUROC curves for LSM, platelets and SSM for the detection of HRV were 0.768 (95%CI= 0.697-0.837), 0.717 (95%CI= 0.664-0.784) and 0.847 (95%CI= 0.795-0.898), respectively (S**upplementary material Figure 1**).

ACCEPTED MANUSCRIPT

With the aim of identifying, by SSM ROC curves, the most accurate SSM cut-off to rule out patients with HRV [corresponding to a low probability (<5%) of HRV presence], a cut-off ≤ 46 kPa was chosen.

At multivariate analysis of the derivation cohort, in model 1, CP-B, LSM <20 kPa, PLT > 150.000/mm3 and SSM≤ 46kPa were significantly associated with HRV (**Model 1 in Table 2).** To evaluate the contribution of SSM in predicting HVR, the multivariate analysis was performed to construct a second model including only CP-B, LSM <20 kPa, PLT > 150,000/mm3 **(Model 2 in Table 2)**. LR Test between model 1 vs. 2 was $\chi 2$ = 20.89 (p-value < 0.00001). Model 1 (with SSM) showed better discrimination (AUROC = 0.787) and calibration (Hosmer-Lemeshow p-value = 0.342) parameters than model 2 (Table 2); according to the dominance analysis, SSM was the most important independent predictor for HRV (**Supplemental Material Table 4**). When these models were applied on the internal validation cohort and on the whole population, good calibration and discrimination of the models were confirmed **(Table 3).** Furthermore, an individual estimation for HRV risk by nomogram, according to the multivariate logistic regression model on the derivation cohort population, was also performed **Figure 2 and Supplemental Material Table 1.**

Diagnostic performance of different NITs in ruling out HRV

In the internal validation cohort, the performance of SSM (≤ 46kPa) in ruling out HRV showed a sensitivity of 97.8%, a specificity of 43.8%, NPV of 98.8% – and an LR- of 0.05. SSM was significantly more accurate (AUROC comparison, De Long test p<0.0001) than Baveno VI (sensitivity 97.8%, specificity 26.3%, NPV 98.1%, LR- 0.08). The combined model Baveno VI/SSM≤46 was significantly more accurate than SSM or Baveno VI alone (AUROC comparison, p=0.048 and p=0.0001, respectively), showing sensitivity 95.7%, specificity 53.1%, NPV 98.1%, and LR- 0.08.

Operator characteristics of the diagnostic tests and criteria under examination in the internal validation cohort (**Table 4)** as well as in the derivation and in the entire retrospective population cohorts (**Supplemental Material table 2)** are reported.

ACCEPTED MANUSCRIPT

SSM-based models and Baveno VI criteria in avoiding EGDs [H2]

In the 240 evaluated patients belonging to the internal validation cohort, we tested the following models to exclude HRV, SSM (≤46 kPa) alone, Baveno VI criteria alone, and a new combined model Baveno VI/SSM≤46. - SSM (≤46 kPa) alone demonstrated that endoscopy was avoidable in 86 – patients (35.8%). Within the group of patients with SSM ≤46 kPa, 15 – patients presented EV (6.3%), of whom only one had HRV (1.2%). Instead, fifty-two (52) out of the 240 patients fulfilled the Baveno VI criteria and consequently could have avoided endoscopy (21.7%); among them, only one of the 5 patients with EV (2.1%) was misclassified as a non-HRV patient (1.9%). One hundred and fifty-four (154) patients with SSM >46 kPa and 188 patients who did not meet Baveno VI criteria would have undergone an endoscopy examination. Ninety-eight percent of patients with HRV had undergone endoscopy according to SSM≤46 kPa and Baveno VI criteria. Comparing SSM ≤46 kPa and Baveno VI criteria, the proportion of avoided EGDs (35.8% vs 21.7%) was significantly improved ($p<0.0008$) and the rates of misclassified HRV were 1.2% and 1.9%, respectively.

By applying the SSM cut-off (≤46 kPa) for the detection of HRV in the internal validation cohort among the patients who did not fulfil Baveno VI criteria (n= 188), another 53 patients would have avoided endoscopy, reaching a total of 105 avoided endoscopies (43.8%). Among the patients identified with the new combined model Baveno VI/SSM≤46 and who could have avoided endoscopy, 19 patients presented EV (7.9%), 2 of whom had HRV (1.9%). On the other hand, more than 98% of patients with HRV had properly undergone endoscopy examination. A graphic representation of the above results in all cohorts considered is shown in **Supplementary material figure 2.** The combination of these two diagnostic tests in a simple, sequential algorithm made it possible to avoid a further 22.1% of unnecessary EGDs if compared with Baveno VI criteria alone, thus reaching a final value of 43.8% of avoided endoscopies in the internal validation cohort. (**Figure 3 A**). Noteworthy, among 52 patients avoiding endoscopy according to Baveno VI criteria, 33 had an SSM ≤46 kPa and 19 had an SSM >46 kPa. Moreover, (**Supplemental material Table 2)** the rates of avoided EGDs of the evaluated diagnostic tests ~~were~~ did not differ significantly, both when Baveno VI criteria and

ACCEPTED MANUSCRIPT

SSM cut-off ≤46 were used alone or sequentially in the combined model, in the derivation cohort (19.4%, 32.2%, 40.7%, respectively) and in the entire retrospective population (20.3%, 33.9%, 42.2%, respectively) (**Supplemental material figure 3)**. The proportion of avoided endoscopies by using the new combined model Baveno VI/SSM≤46 kPa was significantly higher than that of Baveno VI criteria (p=0.0001) and SSM ≤46 kPa testing alone (p=0.009).

Validation of SSM-based models in Prospective External Validation Cohort [H2]

Two hundred and sixteen (216) patients were prospectively considered for enrolment in the external validation cohort. Following the exclusion criteria, a total of 115 cACLD patients were included in the final prospective cohort, of whom 57 (49.6%) were enrolled from the Bologna centre, 32 (27.8%) from Verona, and 26 (22.6%) from Bern. SSM was unsuccessful in 15 (13%) patients (**Figure 1B**).

General characteristics of the external validation cohort are reported in **Table 1**. HRV prevalence was 13%. Aetiology and aminotransferases ~~transaminase~~ values were the only variables that differed significantly compared to the entire retrospective cohort. The accuracy of the multivariate Model 1 in the prediction of HRV in the external validation cohort is described in **Table 3**.

The performance of the various NITs at ruling-out HRV is reported in **Table 4**. Applying Baveno VI criteria, 19 (16,5%) would have avoided EGD screening; none of them presented HRV. Once again, by applying the SSM ≤46 cut-off, HRV was ruled out in an additional of 24 patients, resulting in 43 (37.4%) sparing patient an EGD (no HRV missed) through the new combined model Baveno VI/SSM≤46 (**Figure 3B)**.

Discussion [H1]

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Liver cirrhosis is generally characterized by a long phase of compensated disease; however, when decompensation occurs, the survival of these patients drastically decreases[1]. It therefore becomes very important to achieve proper risk stratification in patients with compensated liver cirrhosis, and this should be particularly focused on identifying and staging portal hypertension and its main complications [1,4,29–31]. Since gold-standard methods for PH and varices are invasive and expensive, there is increasing interest in the development of rapid and accurate NITs [6], LSM by TE in particular [6,8,20].

The results of the present study show that not only CP-B and Baveno VI criteria variables (LSM and PLT) but also mainly SSM performed by TE represents an accurate, non-invasive diagnostic tool in patients with cACLD for excluding the presence of high-risk oesophageal varices. In fact, SSM proved the most relevant independent predictor for HRV by multivariate analysis: the model including not only CP-B and Baveno VI variables but also SSM seems more appropriate to rule out patients with HRV and shows a significantly higher accuracy than a model without SSM **(Table 2-3)**. Accordingly, a highly usable tool represented by nomograms for the prediction of the HRV individual absolute risk was drawn up (**Figure 2**). Furthermore, the sequential use of Baveno VI criteria and SSM in the combined model Baveno VI/SSM≤46 allows an important significant reduction in the number of unnecessary endoscopies when compared to Baveno VI criteria.

In the present study, SSM was the most accurate diagnostic tool in ruling out HRV patients (SSM AUROC= 0.847). Even though SSM ≤46 kPa and Baveno VI criteria showed the same very high sensitivity (97.8%) with high NPV (98.1 vs. 98.8%) (**Table 4**), in the internal validation cohort as well as in the entire retrospective population, SSM had better specificity (26.3% vs. 43.8%).

Using SSM as a single diagnostic test to rule out HRV with a single cut-off ≤46 kPa, the proportion of avoided endoscopies was 35.8%, significantly ($p<0.01$) higher than obtained by applying the Baveno VI criteria (21.7%). Importantly, even if one HRV would have been missed by using Baveno VI criteria or the SSM model (≤46 kPa) alone, the number of HRV correctly detected was ≥95%. The observed high diagnostic

ACCEPTED MANUSCRIPT

accuracy of SSM for HRV is not surprising, since we [19,20], others [32,33] and two meta-analyses [18,34] have already documented its excellent performance in the assessment of PH severity and in predicting the presence of oesophageal varices and HRV in cirrhosis.

Up to now, only one large study [9] has validated Baveno VI criteria in 310 patients (23% with EV and 5% with HRV) with cACLD. In this study, Baveno VI criteria were satisfied in 33% of the patients and HRV were missed in 2%. The Baveno VI criteria were less fulfilled in our study (19.4% - 20.9% - 16.5%) derivation, and internal and external validation set respectively), but they maintained an excellent NPV (HRV missed 0% - 2.2% - 2%). The differences in the results might be explained by a higher proportion of varices in our studied population (50.6% EV and 20.1% HRV), likely as a consequence of more severe liver disease (a higher proportion of CP-B and a higher baseline LSM as compared to Maurice's study [9]). Nonetheless, the proportion of EV and HRV in our population is in line with what was previously observed in compensated cirrhosis [3,35–37], and the proportion of spared endoscopies is similar to that recently reported by a Spanish study in patients with mostly viral aetiology cirrhosis [38].

Several studies [9–14,39] and one meta-analysis [15] have validated Baveno VI criteria as safe for screening for HRV in cACLD patients, with an overall missed HRV of <5%. The proportion of spared EGDs with Baveno VI criteria is however relatively low (15-25%), as also confirmed by our data (17-22%). It should be considered that HRV prevalence in these different cohorts could have influenced the diagnostic performance of NITs. Many authors have proposed different cut-offs and models in order to increase the number of EGDs that could be avoided. The performance of these models in our entire retrospective cohort is shown in **Supplemental Material Table 3.** It appears that more permissive criteria would increase the number of avoided EGDs. However, this would be with lower rate than using our SSM-based models and most importantly, with a rate of HRV missed above the accepted threshold of 5%.

ACCEPTED MANUSCRIPT

Since LSM is characterized by a poor correlation with the prediction of CSPH when employed for the detection of HVPG values >12 mmHg, probably due to the increasing relevance of extra-hepatic factors in conditioning the progression of PH [40], SSM can better evaluate PH progression beyond this level, representing a more useful NIT in EV detection [19,34,41].

In accordance with these physiological hypotheses, we made a sequential algorithm **(Figure 4)** which firstly evaluated HRV risk by Baveno VI criteria and subsequently according to the SSM cut-off (≤46kPa)

Interestingly, we observed that the sequential use of Baveno VI criteria and SSM in the new combined model Baveno VI/SSM≤46 allows further improvement of the results obtained by both strategies used alone **(Table 4 and Supplemental Material Table 2)**. The application of this new sequential algorithm made it possible to increase the number of avoided unneeded endoscopies from 102 to 210, with a final proportion of 42.2% in the entire retrospective cohort, thus doubling the proportion obtained by Baveno VI criteria alone (**Figure 3**). It should be noted that the risk of missing HRV by applying our combined model was acceptable (<5%), according to Baveno VI consensus [8].

In addition to the internal validation by the split-sample method, our combined model Baveno VI/SSM≤46 kPa was also validated externally in a prospective multicentre cohort of 115 cACLD patients with a HRV prevalence of 13% **(Table 1)**. The combined model maintained a superior performance in ruling out HRV, safely sparing 37.4% of EGDs (0 HRV missed), in comparison to 16.5% of EGDs that would have safely been avoided by applying only the Baveno VI Criteria **(Figure 3 and Table 4)**.

In our opinion, SSM≤46 kPa is clearly able to stratify patients for HRV risk better than Baveno VI criteria does. In addition, the combined use of LSM and SSM as an initial HRV screening strategy in the cACLD population is a more ~~an~~ attractive strategy that might avoid misclassification of patients at high risk, despite relatively low LSM or normal PLT values.

Our study has some limitations. One of the main problems of SSM is the rate of unsuccessful examination (15-20%) [42]. Thus, in order to reduce this rate, we performed an US-guided examination, as previously described [19,20], improving the feasibility of SSM (87-94%). In this

ACCEPTED MANUSCRIPT

view, a dedicated SSM-TE device with a built-in US probe for spleen detecting would be desirable. In addition, inclusion of unsuccessful SSM did not negatively influence the performance of models by the ITD approach as reported in the Results section.

Moreover, the rate of spared EGDs did not improve significantly adding Baveno VI criteria to the model including SSM≤46 kPa alone in the prospective cohort, differently to what previously shown in the larger retrospective cohort. In this view, sample size could play a role in the ability to detect a significant improvement combining the two models. Moreover, we believe that LSM is demanded for the non-invasive assessment of cACLD patients, and therefore a key point in the newly proposed sequential algorithm.

Another limitation of the study was the possible influence of any antiviral treatment on the elastography measurement [24]. However, since the availability, and consequently the efficacy, of antiviral treatments were different in the two studied cohorts (mostly interferon in the retrospective cohort, and DAAs in the prospective cohort), the antiviral treatment with DAAs was considered as an exclusion criterion.

Finally, this study did not include patients in whom LSM and SSM were assessed by other ultrasound elastosonographic techniques. Nevertheless, TE is still considered the best validated and most available elastography technique [24], and most centres use TE as the only technique to assess liver and spleen stiffness. Our study has the major strength of including a prospective external multicentre validation cohort of patient who confirmed the results of the retrospective cohort. We also enrolled a large number of well characterized, compensated patients. Another important strength of our study is that we propose both a new, simple sequential algorithm based on binary, easy to remember cut-offs **(Figure 4)**, and also a readily available tool, namely nomograms, based on the continuous modelling of risk of HRV by SSM (**Figure 2**). The latter makes it possible to avoid losing potentially relevant information and aims at individualizing care, thus potentially leading to better decisions in specific individual situations [7].

ACCEPTED MANUSCRIPT

In conclusion, our study indicates that SSM combined with Baveno VI criteria in a simple sequential algorithm is an accurate and non-invasive test for ruling out HRV and safely avoiding a much larger proportion of unnecessary endoscopies then Baveno VI criteria alone. At the same time SSM is also a strong independent predictor of the presence of HRV.

Financial support [H1]

The authors received no financial support to produce this manuscript.

Conflict of interest [H1]

The authors declare no conflicts of interest that pertain to this work.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions [H1]

AC, FR, GM: collected data, analysed data, wrote the manuscript, approved the final manuscript. AC, ED, DBA, MLBR, AB, MP: analysed data and contributed to the drafting and final approval of the manuscript. AC, DF: provided overall oversight of the study, analysed data and contributed to the drafting and final approval of the manuscript.

References

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

[1] 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;44:217–31. doi:10.1016/j.jhep.2005.10.013.

[2] Garcia-Tsao G, Sanyal AJ, Grace ND, Carey WD, Practice Guidelines Committee of American Association for Study of Liver Diseases, Practice Parameters Committee of American College of Gastroenterology. Prevention and management of gastroesophageal varices and variceal hemorrhage in cirrhosis. Am J Gastroenterol 2007;102:2086–102. doi:10.1111/j.1572-0241.2007.01481.x.

[3] Garcia-Tsao G, Abraldes JG, Berzigotti A, Bosch J. Portal hypertensive bleeding in cirrhosis: Risk stratification, diagnosis, and management: 2016 practice guidance by the American Association for the study of liver diseases. Hepatology 2017;65:310–35. doi:10.1002/hep.28906.

[4] de Franchis R, Pascal JP, Ancona E, Burroughs AK, Henderson M, Fleig W, et al. Definitions, methodology and therapeutic strategies in portal hypertension. A Consensus Development Workshop, Baveno, Lake Maggiore, Italy, April 5 and 6, 1990. J Hepatol 1992;15:256–61.

[5] Berzigotti A, Bosch J, Boyer TD. Use of noninvasive markers of portal hypertension and timing of screening endoscopy for gastroesophageal varices in patients with chronic liver disease. Hepatology 2014;59:729–31. doi:10.1002/hep.26652.

[6] EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. J Hepatol 2015;63:237–64. doi:10.1016/j.jhep.2015.04.006.

[7] 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. doi:10.1002/hep.28824.

[8] de Franchis R, Baveno VI Faculty. Expanding consensus in portal hypertension: Report of the Baveno VI Consensus Workshop: Stratifying risk and individualizing care for portal hypertension. J Hepatol 2015;63:743–52. doi:10.1016/j.jhep.2015.05.022.

[9] Maurice JB, Brodkin E, Arnold F, Navaratnam A, Paine H, Khawar S, et al. Validation of the Baveno VI criteria to identify low risk cirrhotic

ACCEPTED MANUSCRIPT

patients not requiring endoscopic surveillance for varices. J Hepatol 2016;65:899–905. doi:10.1016/j.jhep.2016.06.021.

[10] Jangouk P, Turco L, De Oliveira A, Schepis F, Villa E, Garcia-Tsao G. Validating, deconstructing and refining Baveno criteria for ruling out high-risk varices in patients with compensated cirrhosis. Liver Int 2017;37:1177–83. doi:10.1111/liv.13379.

[11] Llop E, Lopez M, de la Revilla J, Fernandez N, Trapero M, Hernandez M, et al. Validation of noninvasive methods to predict the presence of gastroesophageal varices in a cohort of patients with compensated advanced chronic liver disease. J Gastroenterol Hepatol 2017;32:1867–72. doi:10.1111/jgh.13781.

[12] Silva MJ, Bernardes C, Pinto J, Loureiro R, Duarte P, Mendes M, et al. Baveno VI Recommendation on Avoidance of Screening Endoscopy in Cirrhotic Patients: Are We There Yet. GE Port J Gastroenterol 2016:79–83. doi:10.1159/000452693.

[13] Sousa M, Fernandes S, Proença L, Silva AP, Leite S, Silva J, et al. The Baveno VI criteria for predicting esophageal varices: validation in real life practice. Rev Española Enfermedades Dig 2017;109. doi:10.17235/reed.2017.5052/2017.

[14] Calès P, Sacher-Huvelin S, Valla D, Bureau C, Olivier A, Oberti F, et al. Large oesophageal varice screening by a sequential algorithm using a cirrhosis blood test and optionally capsule endoscopy. Liver Int 2018;38:84–93. doi:10.1111/liv.13497.

[15] Marot A, Trépo E, Doerig C, Schoepfer A, Moreno C, Deltenre P. Liver stiffness and platelet count for identifying patients with compensated liver disease at low risk of variceal bleeding. Liver Int 2017;37:707–16. doi:10.1111/liv.13318.

[16] Castera L, Pinzani M, Bosch J. Non invasive evaluation of portal hypertension using transient elastography. J Hepatol 2012;56:696–703. doi:10.1016/j.jhep.2011.07.005.

[17] Colecchia A, Marasco G, Taddia M, Montrone L, Eusebi LH, Mandolesi D, et al. Liver and spleen stiffness and other noninvasive methods to assess portal hypertension in cirrhotic patients: a review of the literature. Eur J Gastroenterol Hepatol 2015;27:992–1001. doi:10.1097/MEG.0000000000000393.

ACCEPTED MANUSCRIPT

[18] Singh S, Eaton JE, Murad MH, Tanaka H, Iijima H, Talwalkar JA. Accuracy of Spleen Stiffness Measurement in Detection of Esophageal Varices in Patients With Chronic Liver Disease: Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol 2014;12:935–945.e4. doi:10.1016/j.cgh.2013.09.013.

[19] 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. doi:10.1053/j.gastro.2012.05.035.

[20] Colecchia A, Colli A, Casazza G, Mandolesi D, Schiumerini R, Reggiani LB, et al. Spleen stiffness measurement can predict clinical complications in compensated HCV-related cirrhosis: A prospective study. J Hepatol 2014;60:1158–64. doi:10.1016/j.jhep.2014.02.024.

[21] Jansen C, Bogs C, Verlinden W, Thiele M, Moller P, Gortzen J, et al. Shear-wave elastography of the liver and spleen identifies clinically significant portal hypertension: A prospective multicentre study. Liver Int 2017;37:396–405. doi:10.1111/liv.13243.

[22] Jansen C, Bogs C, Krag A, Francque S, Trebicka J. Sequential shear-wave elastography of liver and spleen rules out clinically significant portal hypertension in compensated advanced chronic liver disease. Gut 2017;66:558–9. doi:10.1136/gutjnl-2016-311955.

[23] Stefanescu H, Radu C, Procopet B, Lupsor-Platon M, Habic A, Tantau M, et al. Non-invasive menage a trois for the prediction of high-risk varices: stepwise algorithm using lok score, liver and spleen stiffness. Liver Int 2015;35:317–25. doi:10.1111/liv.12687.

[24] Dietrich CF, Bamber J, Berzigotti A, Bota S, Cantisani V, Castera L, et al. EFSUMB Guidelines and Recommendations on the Clinical Use of Liver Ultrasound Elastography, Update 2017 (Long Version). Ultraschall Der Medizin 2017;38:e16–47. doi:10.1055/s-0043-103952.

[25] Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54:774–81.

[26] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves:

ACCEPTED MANUSCRIPT

a nonparametric approach. Biometrics 1988;44:837–45.

[27] Budescu D. Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. Psychol Bull 1993;114:542–51. doi:http://dx.doi.org/10.1037/0033-2909.114.3.542.

[28] Schuetz GM, Schlattmann P, Dewey M. Use of 3x2 tables with an intention to diagnose approach to assess clinical performance of diagnostic tests: meta-analytical evaluation of coronary CT angiography studies. BMJ 2012;345:e6717–e6717. doi:10.1136/bmj.e6717.

[29] 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. doi:10.1053/j.gastro.2007.05.024.

[30] D'Amico G, Pasta L, Morabito A, D'Amico M, Caltagirone M, Malizia G, et al. Competing risks and prognostic stages of cirrhosis: a 25-year inception cohort study of 494 patients. Aliment Pharmacol Ther 2014;39:1180–93. doi:10.1111/apt.12721.

[31] Garcia-Tsao G, Groszmann RJ, Fisher RL, Conn HO, Atterbury CE, Glickman M. Portal pressure, presence of gastroesophageal varices and variceal bleeding. Hepatology 1985;5:419–24.

[32] Takuma Y, Nouso K, Morimoto Y, Tomokuni J, Sahara A, Toshikuni N, et al. Measurement of spleen stiffness by acoustic radiation force impulse imaging identifies cirrhotic patients with esophageal varices. Gastroenterology 2013;144:92–101.e2. doi:10.1053/j.gastro.2012.09.049.

[33] Stefanescu H, Grigorescu M, Lupsor M, Procopet B, Maniu A, Badea R. Spleen stiffness measurement using Fibroscan for the noninvasive assessment of esophageal varices in liver cirrhosis patients. J Gastroenterol Hepatol 2011;26:164–70. doi:10.1111/j.1440-1746.2010.06325.x.

[34] Ma X, Wang L, Wu H, Feng Y, Han X, Bu H, et al. Spleen Stiffness Is Superior to Liver Stiffness for Predicting Esophageal Varices in Chronic Liver Disease: A Meta-Analysis. PLoS One 2016;11:e0165786. doi:10.1371/journal.pone.0165786.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

[35] Kovalak M, Lake J, Mattek N, Eisen G, Lieberman D, Zaman A. Endoscopic screening for varices in cirrhotic patients: data from a national endoscopic database. Gastrointest Endosc 2007;65:82–8. doi:10.1016/j.gie.2006.08.023.

[36] Pagliaro L, D'Amico G, Pasta L, Politi F, Vizzini G TM, Al. E. Portal hypertension in cirrhosis: natural history. Bosch J, Groszmann RJ, Eds Portal Hypertens Pathophysiol Treat 1994;Oxford, UK:72–92.

[37] Merli M, Nicolini G, Angeloni S, Rinaldi V, De Santis A, Merkel C, et al. Incidence and natural history of small esophageal varices in cirrhotic patients. J Hepatol 2003;38:266–72.

[38] Augustin S, Pons M, Genesca J. Validating the Baveno VI recommendations for screening varices. J Hepatol 2017;66:459–60. doi:10.1016/j.jhep.2016.09.027.

[39] Augustin S, Pons M, Maurice JB, Bureau C, Stefanescu H, Ney M, et al. Expanding the Baveno VI criteria for the screening of varices in patients with compensated advanced chronic liver disease. Hepatology 2017;0:1–9. doi:10.1002/hep.29363.

[40] Vizzutti F, Arena U, Romanelli RG, Rega L, Foschi M, Colagrande S, et al. Liver stiffness measurement predicts severe portal hypertension in patients with HCV-related cirrhosis. Hepatology 2007;45:1290–7. doi:10.1002/hep.21665.

[41] Castéra L, García-Tsao G. When the spleen gets tough, the varices get going. Gastroenterology 2013;144:19–22. doi:10.1053/j.gastro.2012.11.015.

[42] Berzigotti A. Non invasive evaluation of portal hypertension using ultrasound elastography. J Hepatol 2017. doi:10.1016/j.jhep.2017.02.003

Tables and Figure legends:

Fig. 1. The study design flow chart of the enrolled patients; Internal Retrospective Cohort (A) and External Prospective Cohort (B).

ACCEPTED MANUSCRIPT

Fig. 2. Nomogram to predict the presence of High Risk Varices (HRV) by Platelet count, Liver and Spleen Stiffness Measurements. To calculate the probability of HRV, trace a vertical line from the predictor to the risk axis. [In nomograms with one variable, to calculate the probability of HRV, trace a vertical line from the predictor to the risk axis. For the nomogram with two variables (LS and PLT), trace a vertical line from each of the predictors' axis to the first line ("points"). Add the total points, and trace a vertical line from the "total points" axis to the risk axis to calculate the risk of HRV]

Fig. 3. Rate of endoscopies avoided by each of the non-invasive models in the (Blue) Internal Retrospective Validation Cohort (A) and (Green) External Prospective Validation Cohort (B). [Darker sections represent the quote of EGDs spared by models; In the new combined model Baveno VI/SSM ≤46 kPa, the rate of avoided endoscopies is reported also in the column graphs as the sum of the patients that fulfil only the Baveno VI Criteria (SSM>46kPa), patients with SSM ≤46 kPa (and outside Baveno VI Criteria) and patients that would have avoided endoscopy according to both models (in overlapping area).]

Fig. 4. The new combined model Baveno VI/SSM ≤46 kPa. [Algorithm-based screening for High Risk Varices in patients with suspicion of cACLD according to the new combined model Baveno VI/SSM ≤46 kPa.]

Table 1. Baseline characteristics of the patient population. Qualitative data were expressed as number and percentages (%); quantitative data were expressed as median (25% - 75% quantiles)

*patients with SSM available: 472 (entire population); 242 (derivation cohort); 230 (internal validation cohort); 100 (prospective external validation cohort)

§Significantly different variables between entire retrospective and prospective validation cohort

ACCEPTED MANUSCRIPT

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; Body Mass Index, BMI; High risk varices, HRV; International normalized ratio, INR; Liver Stiffness Measurement, LSM; Low risk varices, LRV; Model for End-Stage Liver Disease MELD; Spleen Stiffness Measurement SSM;

Table 2. Univariate and multivariate association of each potential predictor for high risk varices in the derivation cohort.

Qualitative data were expressed as number and perceptual (%); quantitative data were expressed as median (25% - 75% quantiles)

#Multivariate models were performed only in patients with SSM available

*Multivariate Model 1 variables considered were Child-Pugh B, MELD score, ALT, LSM<20kPa, Platelet > 150.000/mm3 and SSM ≤46 kPa

°Multivariate Model 2 variables considered were Child-Pugh B, MELD score, ALT, LSM<20kPa, Platelet > 150.000/mm3

Akaike information criterion, AIC; Area under the receiver operator characteristic curve, AUROC; Alanine aminotransferase, ALT; Aspartate aminotransferase, AST; Body Mass Index, BMI; Bayesian information criterion, BIC; Confidence Interval, CI; High Risk Varices, Hosmer–Lemeshow goodness-of-fit test, H-L; HRV; Liver Stiffness Measurement, LSM; Model for End-Stage Liver Disease MELD; Odds Ratio, OD; Spleen Stiffness Measurement SSM

Table 3. Multivariate association of each potential predictor for High Risk Varices in the Internal and External validation cohorts and in the entire population.

[#]Multivariate models were performed only in patients with SSM available

[§]Multivariate model take into account inter-center variability

***Multivariate Model 1** variables considered ~~included~~ Child-Pugh B, MELD score, ALT, LSM<20kPa, Platelet > 150.000/mm^3 and SSM ≤46 kPa

°Multivariate Model 2 variables considered ~~included~~ Child-Pugh B, MELD score, ALT, LSM<20kPa, Platelet > 150.000/mm^3

Akaike information criterion, AIC; Bayesian information criterion, Under the receiver operator characteristic curve, AUROC; BIC; Confidence Interval, CI; High Risk Varices, Hosmer–Lemeshow goodness-of-fit test, H-L; HRV; Liver Stiffness Measurement, LSM; Odds Ratio, OD; Spleen Stiffness Measurement SSM.

ACCEPTED MANUSCRIPT

Table 4. Operating characteristics for different non-invasive models in ruling-out High-Risk Varices in Internal and External validation cohorts.

Prevalence of HRV was 19.2% in the validation cohort and 13% in prospective external validation cohort.

* Unfeasible SSM were considered as False Positives (SSM >46 kPa), according to ITD principle.

Abbreviations: 95% Confidence Interval (95% CI); Like-hood ratio (LR); Liver Stiffness Measurement, LSM; Negative Predictive Values (NPV); Positive Predictive Values (PPV); Spleen Stiffness Measurement SSM

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Variable	Entire Retrospective Population (n=498)	Derivation Cohort (n= 258)	Internal Validation Cohort (n=240)	Prospective External Validation Cohort (n=115)
Age (years)	60 (50-70)	60.5 (51 - 69)	60 (50 - 72)	58 (53 - 67)
Male	291 (58.4)	160 (62)	131 (55)	78 (67.8)
BMI (kg/m2)	25.9 (23.9-28)	25.8 (24 - 28)	26 (23.8 – 28.6)	26.4 (23.9 – 30.2)
Aetiology[§]				
Hepatitis C	424 (85.1)	221 (85.7)	203 (84)	47 (40.9)
Hepatitis B	29 (5.9)	16 (6.2)	14 (6)	5 (4.4)
Alcohol	35 (7)	16 (6.2)	19 (8)	19 (16.5)
NAFLD/NASH				19 (16.5)
Autoimmune				3 (2.6)
Miscellaneous	10 (2)	5 (1.9)	4 (2)	22 (19.1)
Child Pugh Score[§]				
A	350 (70.3)	173 (67.1)	177 (73.8)	100 (86.9)
B	148 (29.7)	85 (32.9)	63 (26.2)	15 (13.1)
MELD Score	8 (7-10)	8 (7 - 10)	8 (7 - 10)	7 (7 - 10)
Esophageal Varices	252 (50.6)	129 (50)	123 (51)	60 (52.2)
LRV	152 (30.5)	75 (29.1)	77 (32)	45 (39.1)
HRV	100 (20.1)	54 (20.9)	46 (19)	15 (13)
Laboratory results				

ACCEPTED MANUSCRIPT

Platelets (cells x 10^9/L)	110 (79-115)	111 (79 – 155)	120 (81 – 173)	113 (80 - 156)
ALT (U/L) [§]	52 (31-88)	52 (32 - 89)	53 (31 - 88)	32 (22 - 50)
AST (U/L) [§]	54 (35-84)	55 (37 - 83)	53 (34 - 86)	39 (30 - 57)
Bilirubin (mg/dL)	0.84 (0.63-1.28)	0.86 (0.67 – 1.22)	0.82 (0.58 – 1.29)	0.9 (0.54 – 1.36)
Albumin (g/dL)	3.8 (3.4-4.1)	3.8 (3.5 – 4.1)	3.8 (3.3 – 4.1)	3.8 (3.6 – 4.1)
Creatinine (mg/dL)	0.8 (0.70-0.94)	0.81 (0.7 – 0.93)	0.79 (0.68 – 0.97)	0.85 (0.73 – 0.92)
INR	1.14 (1.07-1.24)	1.14 (1.07 – 1.24)	1.13 (1.05 – 1.22)	1.17 (1.09 – 1.22)
Non-invasive tests				
LSM (kPa)	21.3 (15-30.3)	21.5 (14.5 – 30.3)	20.6 (15 – 28.1)	22.3 (15.1 – 35.8)
SSM (kPa)*	54 (40 – 67.3)	55 (40.9 – 69.1)	53.2 (39.7 – 65.2)	60.4 (37.4 - 75)

Table 1: Baseline characteristics of the patient population

Qualitative data were expressed as number and perceptual (%); quantitative data were expressed as median (25% - 75% quantiles)

*patients with SSM available: 472 (entire population); 242 (derivation cohort); 230 (internal validation cohort); 100 (prospective external validation cohort)

[§]Significantly different variables between entire retrospective and prospective validation cohort

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; Body Mass Index, BMI; High risk varices, HRV; International normalized ratio, INR; Liver Stiffness Measurement, LSM; Low risk varices, LRV; Model for End-Stage Liver Disease MELD; Spleen Stiffness Measurement SSM;

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Derivation cohort (n=258)[#]

	Univariate analysis				Multivariate analysis			
					Model 1*		Model 2°	
Variable	**Absence of HRV (n=204)**	**HRV (n=54)**	**OR [95% CI]**	**p-value**	**OR [95% CI]**	**p-value**	**OR [95% CI]**	**p-value**
Age (years)	61 (59 - 69)	60 (53 - 69)	1.014 (0.988 – 1.042)	0.298				
BMI (kg/m2)	25.7 (24 - 28)	26.3 (25 - 28.1)	1.041 (0.888 – 1.222)	0.619				
Aetiology			1.380 (0.779 – 2.613)	0.322				
Hepatitis C	174 (85)	47 (87)						
Hepatitis B	14 (7)	2 (3.5)						
Alcohol	13 (7)	3 (6)						
Miscellaneous	3 (1)	2 (3.5)						
Child Pugh Score B	56 (66)	29 (34)	3.066 (1.654 – 5.692)	**< 0.0001**	2.728 (1.367 – 5.447)	**0.004**	3.036 (1.560 – 5.908)	**0.001**
MELD Score	8 (7 - 10)	9 (8 - 11)	1.121 (0.985 – 1.275)	0.082				
ALT (U/L)	56 (32 - 94)	46 (31 - 63)	0.994 (0.987 – 1.001)	0.075				
AST (U/L)	54 (37 - 84)	56 (38 - 75)	0.998 (0.991 – 1.005)	0.560				
LSM (kPa)	18.3 (14 –	31.7 (24.2 –	1.068 (1.042 –	**< 0.0001**				

ACCEPTED MANUSCRIPT

	26.3)	38.6)	1.096)					
LSM <20 kPa (yes)	107 (91%)	10 (9%)	0.206 (0.098 – 0.432)	**< 0.0001**	0.812 (0.343 – 0.975)	**0.063**	0.347 (0.159 – 0.761)	**0.008**
Platelets (cells x 10^9/L)	125 (88 - 163)	84 (59 - 112)	0.986 (0.979 – 0.993)	**< 0.0001**				
Platelet > 150.000/mm^3 (yes)	67 (94%)	4 (6%)	0.164 (0.057 – 0.472)	**0.001**	0.437 (0.135 – 0.794)	**0.047**	0.306 (0.101 – 0.928)	**0.036**
SSM (kPa)	49.7 (38.6 – 64.7)	72.8 (64.4 - 75)	1.108 (1.072 – 1.145)	**< 0.0001**				
SSM ≤46 kPa (yes)	86 (99%)	1 (1%)	0.026 (0.004 – 0.191)	**< 0.0001**	0.038 (0.005 – 0.313)	**<0.0001**		
					Model 1: ○ AUROC = 0.787 ○ H-L= 0.342 ○ AIC= 211.4 ○ BIC= 228.9		**Model 2:** ○ AUROC = 0. 732 ○ H-L= 0.348 ○ AIC= 230.3 ○ BIC= 244.3	
					LR Test between model 1 vs. 2, χ^2 = 20.89; **p-value < 0.00001** **DeLong Test p-value = 0.0024**			

Table 2: Univariate and multivariate association of each potential predictor and High-Risk Varices in derivation cohort.

Qualitative data were expressed as number and perceptual (%); quantitative data were expressed as median (25% - 75% quantiles)

[#]Multivariate models were performed only in patients with SSM available

***Multivariate Model 1** variables considered were Child-Pugh B, MELD score, ALT, LSM<20kPa, Platelet > 150.000/mm^3 and SSM ≤46 kPa

°**Multivariate Model 2** variables considered were Child-Pugh B, MELD score, ALT, LSM<20kPa, Platelet > 150.000/mm^3

ACCEPTED MANUSCRIPT

Abbreviations: Akaike information criterion, AIC; Area under the receiver operator characteristic curve, AUROC; Alanine aminotransferase, ALT; Aspartate aminotransferase, AST; Body Mass Index, BMI; Bayesian information criterion, BIC; Confidence Interval, CI; High Risk Varices, Hosmer–Lemeshow goodness-of-fit test, H-L; HRV; Liver Stiffness Measurement, LSM; Model for End-Stage Liver Disease MELD; Odds Ratio, OD; Spleen Stiffness Measurement SSM

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

		Multivariate analysis[#]				
		Internal Validation Cohort (n=230)		Retrospective Entire Population (n=472)		External Validation Cohort (n=115)
		Model 1*	**Model 2°**	**Model 1***	**Model 2°**	**Model 1*[§]**
Discrimination	**AUROC**	0.804	0.747	0.789	0.741	0.831
Calibration	**Hosmer – Lemeshow *p*- value**	0.935	0.758	0.591	0.449	0.387
Goodness of fit	**AIC**	188.3	207.5	393.4	433.9	76.17
AUROC Comparison		**Delong Test p-value = 0.0005**		**Delong Test p-value =0.0001**		

Table 3: Multivariate association of each potential predictor and High Risk Varices in validation cohorts and entire population.

[#]Multivariate models were performed only in patients with SSM available

[§]Multivariate model take into account inter-center variability

***Multivariate Model 1** variables considered ~~included~~ Child-Pugh B, MELD score, ALT, LSM<20kPa, Platelet > 150.000/mm^3 and SSM ≤46 kPa

°Multivariate Model 2 variables considered ~~included~~ Child-Pugh B, MELD score, ALT, LSM<20kPa, Platelet > 150.000/mm^3

Abbreviations: Akaike information criterion, AIC; Bayesian information criterion, Under the receiver operator characteristic curve, AUROC; BIC; Confidence Interval, CI; High Risk Varices, Hosmer–Lemeshow goodness-of-fit test, H-L; HRV; Liver Stiffness Measurement, LSM; Odds Ratio, OD; Spleen

ACCEPTED MANUSCRIPT

Stiffness Measurement SSM.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Patients	Internal Validation Cohort (n=240)	External Validation cohort (n= 115)

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

	LSM <20 kPa + PLT>150.000/mm^3 (n= 52)	SSM ≤ 46 kPa* (n= 86)	Combined Model Baveno VI/SSM≤ 46 kPa* (n= 105)	LSM <20 kPa + PLT>150.000/mm^3 (n= 19)	SSM ≤ 46 kPa* (n= 35)	Combined Model Baveno VI/SSM≤ 46 kPa* (n= 43)
True positive	45	45	44	15	15	15
False positive	143	109	91	81	65	57
True negative	1	1	2	0	0	0
False negative	51	85	103	19	35	43
EDGs spared (%)	21.7	35.8	43.8	16.5	30.4	37.4
HRV missed (HRV missed/HRV total, %)	2.2	2.2	4.3	0	0	0
Operating characteristics						
Sensitivity (95% CI)	0.978 (0.885 – 0.999)	0.978 (0.885 – 0.999)	0.957 (0.852 – 0.995)	1 (0.753 – 1)	1 (0.753 – 1)	1 (0.753 – 1)
Specificity (95% CI)	0.263 (0.202 – 0.331)	0.438 (0.367 – 0.511)	0.531 (0.458 – 0.603)	0.190 (0.118 – 0.281)	0.350 (0.257 – 0.452)	0.430 (0.331 – 0.533)
PPV (95% CI)	0.239 (0.223 – 0.257)	0.292 (0.266 – 0.320)	0.326 (0.291 – 0.362)	0.156 (0.144 – 0.169)	0.188 (0.167 – 0.210)	0.208 (0.182 – 0.238)
NPV (95% CI)	0.981 (0.879 – 0.997)	0.988 (0.924 – 0.998)	0.981 (0.930 – 0.995)	1	1	1
LR-positive (95% CI)	1.33 (1.21– 1.46)	1.74 (1.53 –	2.04 (1.73 –	1.23 (1.12 – 1.36)	1.54 (1.33 –	1.75 (1.48 –

ACCEPTED MANUSCRIPT

		1.99)	2.40)		1.78)	2.08)
LR-negative (95% CI)	0.08 (0.01 – 0.58)	0.05 (0.01-0.35)	0.08 (0.02 – 0.32)	0	0	0

Table 4: Operating characteristics for different non-invasive tests in ruling-out High-Risk Varices in internal validation and prospective external validation cohorts

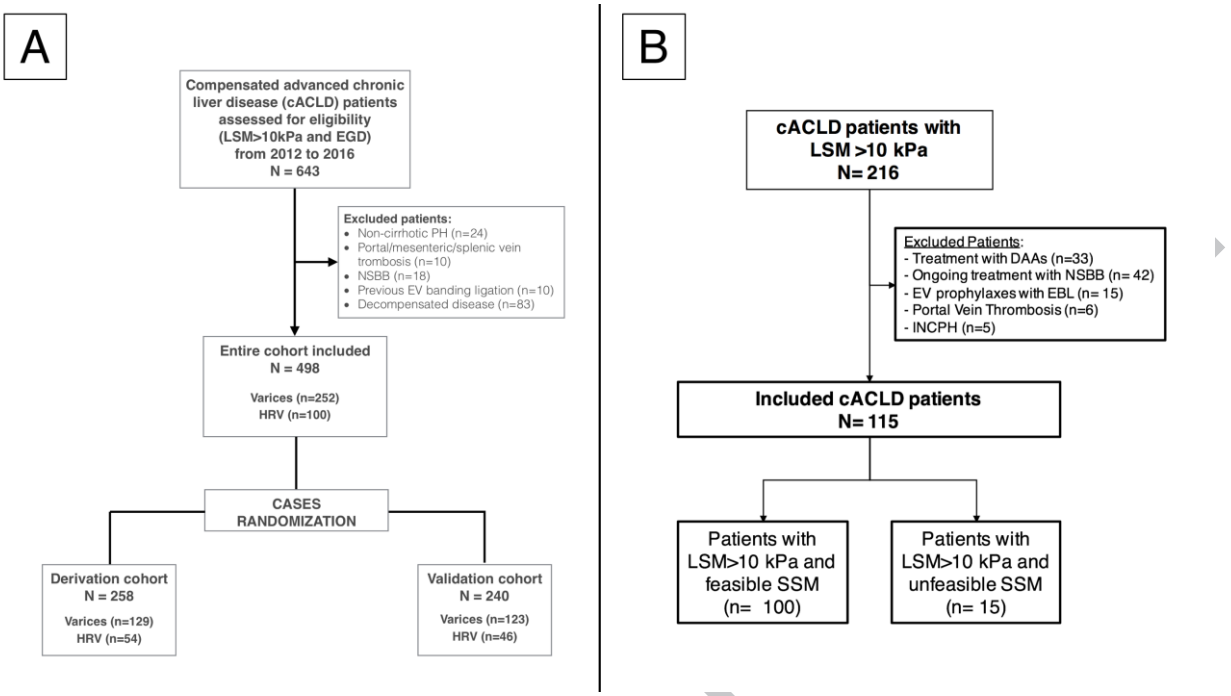
Prevalence of HRV was 19.2% in the validation cohort and 13% in prospective external validation cohort.

* Unfeasible SSM were considered as False Positives (SSM >46 kPa), according to ITD principle.

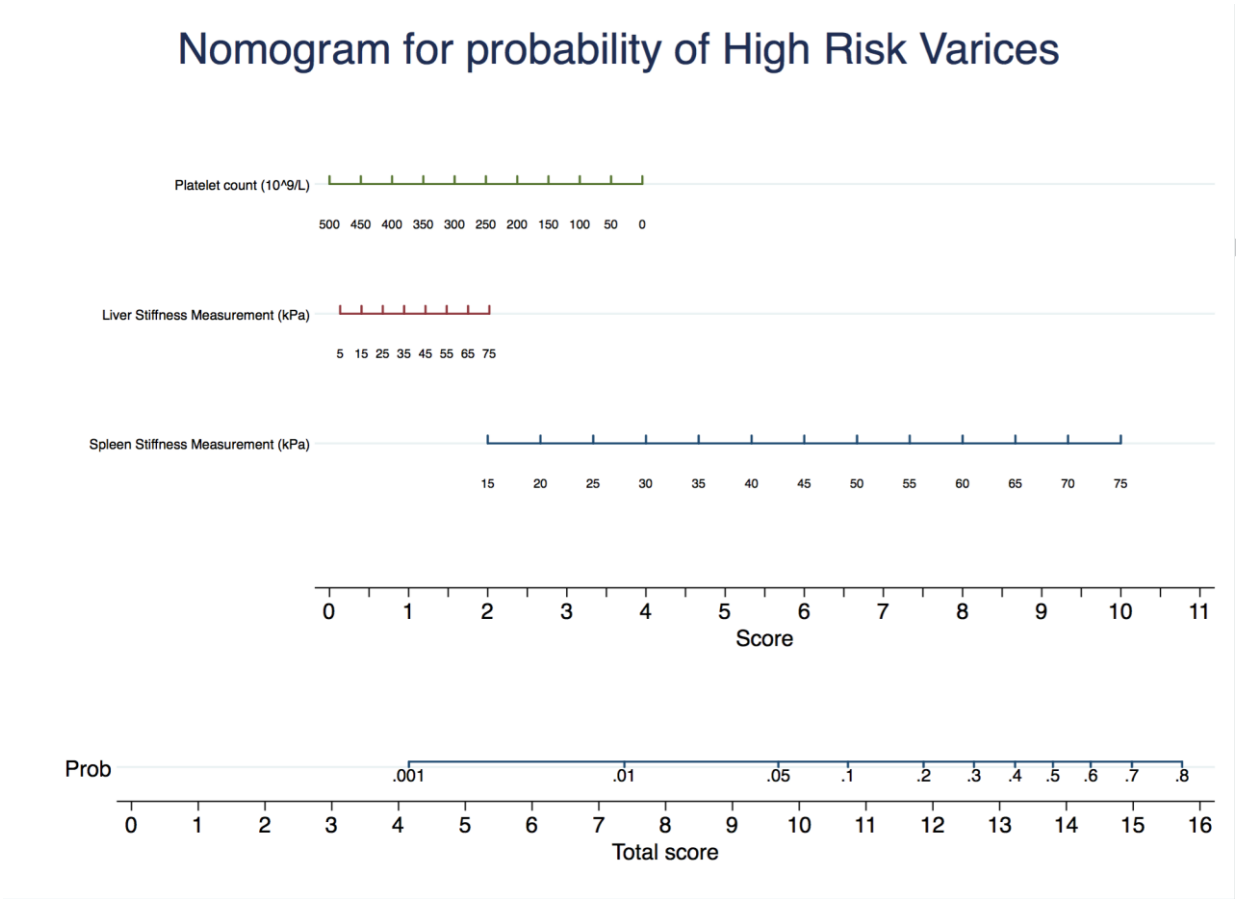
Abbreviations: 95% Confidence Interval (95% CI); Like-hood ratio (LR); Liver Stiffness Measurement, LSM; Negative Predictive Values (NPV); Positive Predictive Values (PPV); Spleen Stiffness Measurement SSM

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

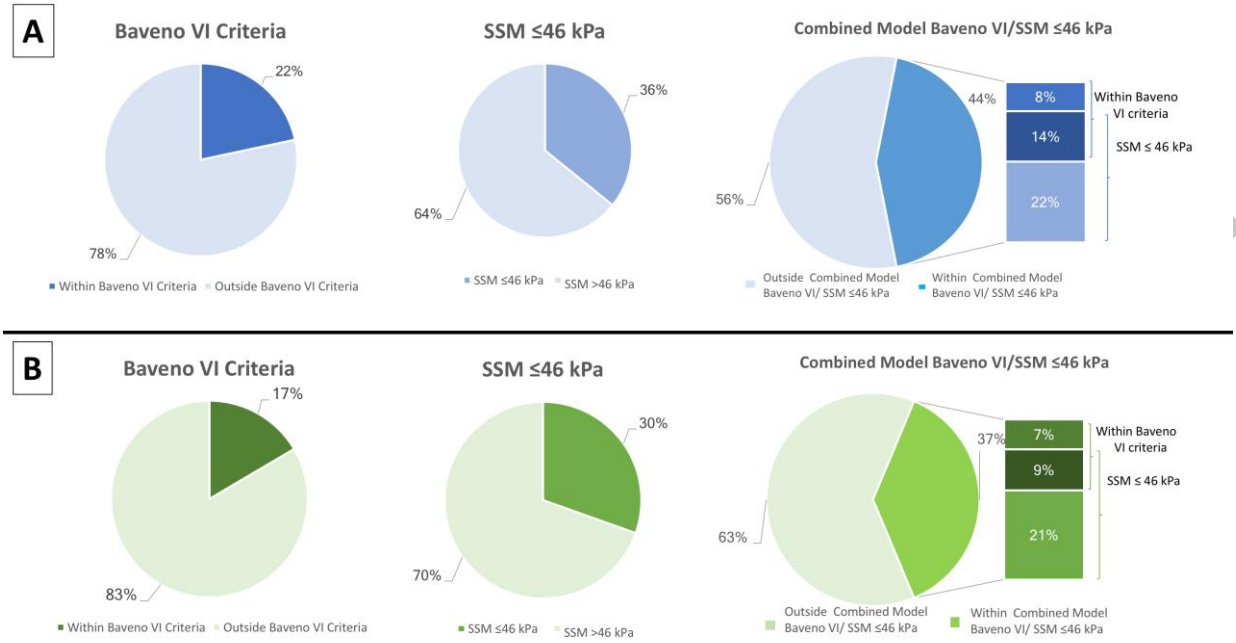


ACCEPTED MANUSCRIPT

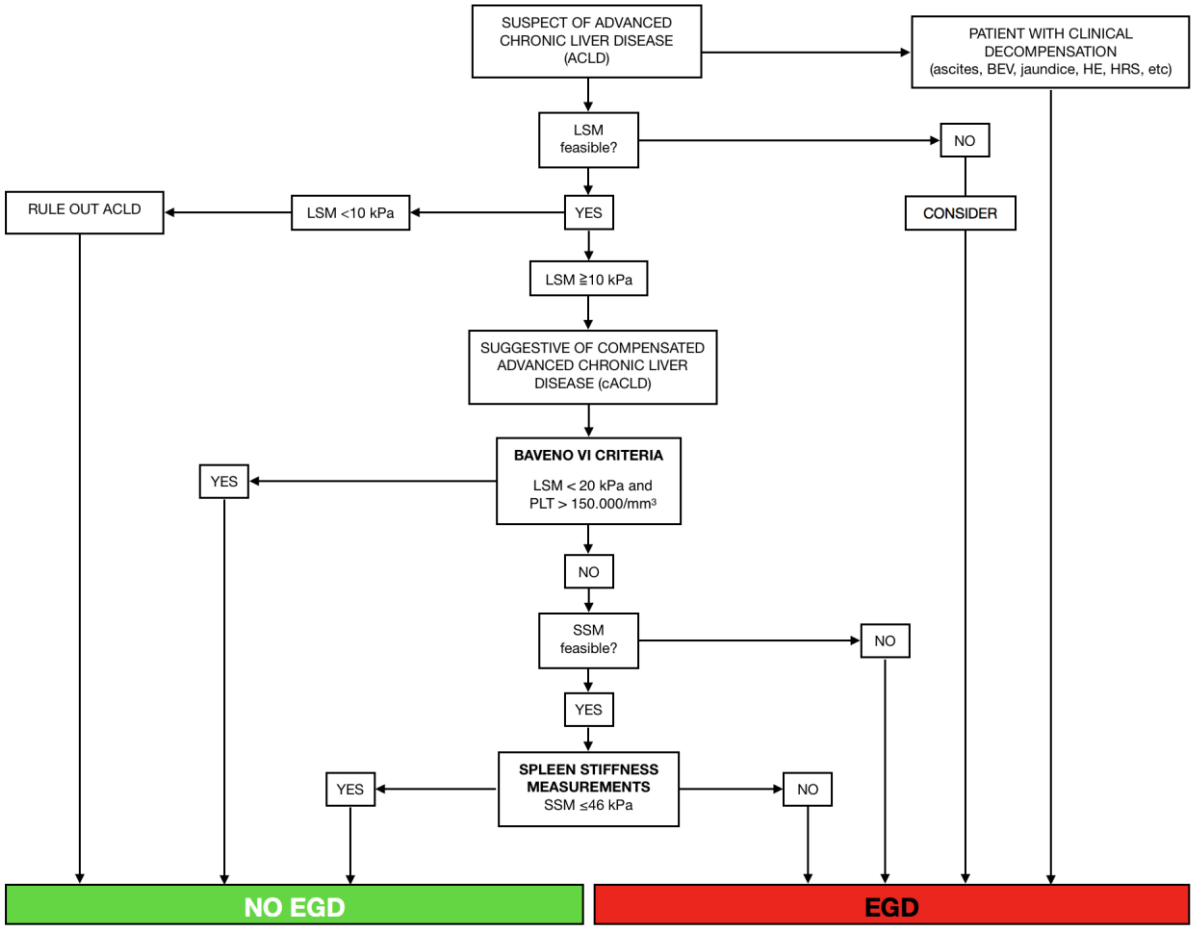


ACCEPTED

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT



ACCEPTED

ACCEPTED MANUSCRIPT

Ref: JHEPAT-D-17-01577

Title: A spleen stiffness measurement-based model for the recognition of high risk varices: Baveno VI criteria and beyond

Article Type: Original Article

Journal of Hepatology

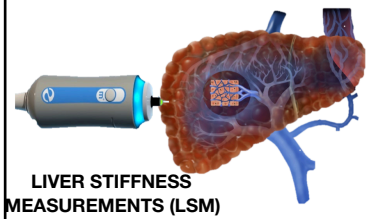
Highlights

- The Baveno VI criteria (LSM < 20 kPa and platelet count > 150,000) aim to identify patients at low risk (<5%) of having EV requiring treatment (HRV) and therefore could safely avoid endoscopy (EGD) in cACLD patients.
- Spleen Stiffness Measurement (SSM) is the most useful non-invasive test able to evaluate portal hypertension degree, presence of EV and probability of clinical decompensation.
- SSM proved an independent risk factor for HRV with LSM, platelet count and Child-Pugh B.
- SSM cut-off (≤46 kPa) shows a better specificity, maintaining the same high sensitivity as Baveno VI in ruling out patients with HRV.
- A new combined Baveno VI/SSM≤46 model can safely avoid more EGDs than Baveno VI or SSM≤46 alone.

ACCEPTED MANUSCRIPT

SUGGESTIVE OF COMPENSATED ADVANCED CHRONIC LIVER DISEASE (cACLD)

BAVENO VI CRITERIA



LIVER STIFFNESS MEASUREMENTS (LSM)

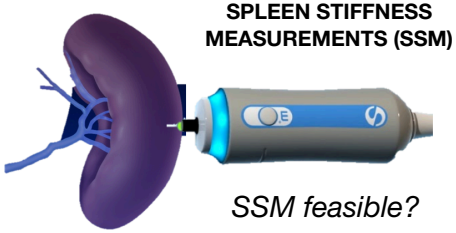
LSM < 20 kPa and
PLT > 150.000/mm^3

PLATELETS COUNT (PLT)

YES

NO

SPLEEN STIFFNESS MEASUREMENTS (SSM)

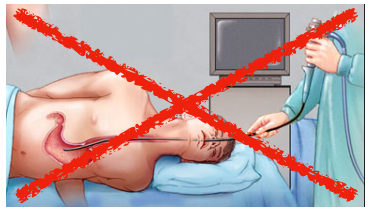
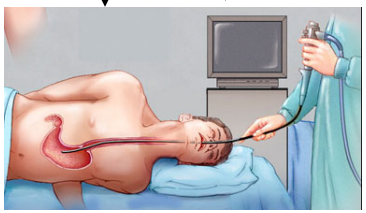


SSM feasible?

NO

YES

SSM ≤46 kPa

YES

NO