



Including Ratio of Platelets to Liver Stiffness Improves Accuracy of Screening for Esophageal Varices That Require Treatment

Arthur Berger, Federico Ravaioli, Oana Farcau, Davide Festi, Horia Stefanescu, François Buisson, Nathalie Ganne-Carriè, Annalisa Berzigotti, Salvatore Petta, Sylvie Sacher Huvelin, et al.

► To cite this version:

Arthur Berger, Federico Ravaioli, Oana Farcau, Davide Festi, Horia Stefanescu, et al.. Including Ratio of Platelets to Liver Stiffness Improves Accuracy of Screening for Esophageal Varices That Require Treatment. *Clinical Gastroenterology and Hepatology*, 2021, 19 (4), pp.777-787.e17. 10.1016/j.cgh.2020.06.022 . hal-03258178

HAL Id: hal-03258178

<https://u-paris.hal.science/hal-03258178>

Submitted on 22 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Including Ratio of Platelets to Liver Stiffness Improves Accuracy of Screening for Esophageal Varices That Require Treatment

Short Title: Personalized esophageal varice screening

Arthur Berger MD^{*1}, Federico Ravaioli MD^{*2}, Oana Farcau MD³, Davide Festi MD², Horia Stefanescu MD³, François Buisson MD¹, Pierre Nahon MD⁴, Christophe Bureau MD⁵, Nathalie Ganne-Carrié MD⁴, Annalisa Berzigotti MD⁶, Victor de Ledinghen MD⁷, Salvatore Petta MD⁸, Paul Calès MD¹, and multicenter groups: VO-VCO, ANRS CO12 CIRVIR, M116 and VEB6.⁹ Angers, Bologna, Cluj, Bondy, Toulouse, Bern, Bordeaux.

* The authors whose names are in bold share first authorship.

Authors' institutions:

¹ Hepato-Gastroenterology Department, University Hospital, Angers, France; HIFIH laboratory, UPRES 3859, UNIV Angers, France

² Gastroenterology and Hepatology Unit, Department of Medical and Surgical Sciences (DIMEC), University of Bologna, Bologna, Italy

³ Liver Unit, Regional Institute of Gastroenterology and Hepatology, Cluj-Napoca, Romania

⁴ Hepato-Gastroenterology Department, Paris Seine-Saint-Denis University Hospitals, APHP, Jean Verdier Hospital, Bondy; University Paris 13, Bobigny; INSERM UMR 1162, Paris, France

⁵ Hepato-Gastroenterology Department, Purpan University Hospital, Toulouse, France

University of Bern, Switzerland

⁷ Hepatology Department, Haut-Lévêque Hospital, Bordeaux, France

⁸ Department of Gastroenterology and Hepatology, PROMISE, University of Palermo, Palermo, Italy

⁹ All authors and centers, merged in the VEB6 group for the present study, are listed at the end of the main text.

Correspondence:

Paul Calès, Hepatology, CHU, 49933 Angers Cedex 09, France. Tel: (33) 2 41 35 34 10, Fax: (33) 2 41 35 41 19, E-mail: paul.cales@univ-angers.fr
<http://orcid.org/0000-0003-4866-5274>

Grant Support:

Institutional support from participating centers, PHRC and ANRS (FRENSH: France Recherche Nord & Sud Sida-HIV Hépatites) who had no other role in the present study.

Abbreviations:

B6C: Baveno VI criteria

cACLD: compensated advanced chronic liver disease

CLD: chronic liver disease

EB6C: expanded Baveno VI criteria

EGD: esophagogastroduodenoscopy

EV: esophageal varices

INR: international normalized ratio

LSM: liver stiffness measurement

MELD: model for end-stage liver disease

NAFLD: non-alcoholic fatty liver disease

NIT: non-invasive test

NPV: negative predictive value

PLER: platelet/liver elastometry ratio

PLEASE: platelet/liver elastometry ratio adjusted on etiology, sex, INR

PPV: positive predictive value

VCTE: vibration-controlled transient elastometry

VNT: varices needing treatment

Disclosures:

Arthur Berger: consultant for Lilly until December 2018

Federico Ravaioli, Oana Farcau, Horia Stefanescu, François Buisson, Davide Festi, Annalisa

Berzigotti: no disclosures

Pierre Nahon: honoraria from Abbvie, Bayer, Bristol-Myers Squibb, Gilead and Ipsen;

consultant for Abbvie, Bayer, Bristol-Myers Squibb and Ipsen

Christophe Bureau: honoraria from Gore, Gilead, Abbvie

Nathalie Ganne-Carrié: grant from Echosens for a monocenter clinical study (M123);

honoraria from Abbvie, Bayer, BMS and Gilead

Victor de Ledinghen: lectures: AbbVie, Gilead, Intercept Pharma, Echosens, Supersonic

Imagine; consultant for Abbvie, Gilead, Pfizer

Salvatore Petta: advisor and/or speaker for Abbvie, Gilead and Intercept

Paul Calès: consultant for Echosens until January 2019, Guarantor of the article

Author contributions:

Arthur Berger: study supervision, critical revision of the manuscript, data collection, verification of statistics

Federico Ravaoli: study supervision, critical revision of the manuscript, data collection

Oana Farcau: critical revision of the manuscript, data collection

Davide Festi: study supervision, critical revision of the manuscript, data collection

Horia Stefanescu: critical revision of the manuscript, data collection

François Buisson: critical revision of the manuscript, data collection, literature review

Pierre Nahon: critical revision of the manuscript, data collection

Christophe Bureau: critical revision of the manuscript, data collection

Nathalie Ganne-Carriè: critical revision of the manuscript, data collection

Annalisa Berzigotti: critical revision of the manuscript, data collection

Victor de Ledinghen: critical revision of the manuscript, data collection

Salvatore Petta: critical revision of the manuscript, data collection

Paul Calès: study concept and design, analysis and interpretation of data, drafting of the manuscript, data collection, statistics, guarantor of the article (including final responsibility for the decision to submit for publication)

VEB6 group: data collection

Item count:

3992 words (main text <4000 including the figure and table legends, and references), number of figures: 3, number of tables: 3, number of references: 17, number of supplementary files

spaces), short title: 40 characters (<45)

Authors' e-mails:

arthur.berger@chu-bordeaux.fr

f.ravaioli@unibo.it

oana.farcu@gmail.com

davide.festi@unibo.it

horia.stefanescu@irgh.ro

francois.buisson.nantes@gmail.com

pierre.nahon@aphp.fr

bureau.c@chu-toulouse.fr

nathalie.ganne@aphp.fr

Annalisa.Berzigotti@insel.ch

victor.deledinghen@chu-bordeaux.fr

salvatore.petta@unipa.it

paul.cales@univ-angers.fr

Background & Aims: Based on platelets and liver stiffness measurements, the Baveno VI criteria (B6C), the expanded B6C (EB6C), and the ANTICIPATE score can be used to rule out varices needing treatment (VNT) in patients with compensated chronic liver disease. We aimed to improve these tests by including data on ratio of platelets to liver stiffness.

Methods: In a retrospective analysis of data from 10 study populations, collected from 2004 through 2018, we randomly assigned data from 2368 patients with chronic liver disease of different etiologies to a derivation population (n=1579; 15.1% with VNT, 50.2% with viral hepatitis, 28.9% with nonalcoholic fatty liver disease, 28.9% with alcohol-associated liver disease, 20.8% with model for end-stage liver disease scores of 9.5 ± 3.0 , and 93.0% with liver stiffness measurements ≥ 10 kPa) or a validation population (n=789). Test results were compared with results from a sequential algorithm (VariScreen). VariScreen incorporated data on platelets or liver stiffness measurements and then ratio of platelets to liver stiffness measurement, adjusted for etiology, patient sex, and international normalized ratio.

Results: In the derivation population, endoscopies were spared for 23.9% of patients using the B6C (VNTs missed in 2.9%), 24.3% of patients using the ANTICIPATE score (VNTs missed in 4.6%), 34.5% of patients using VariScreen (VNTs missed in 2.9%), and 41.9% of patients using the EB6C (VNTs missed in 10.9%). Differences in spared endoscopy rates were significant ($P \leq .001$) except for B6C vs ANTICIPATE and in missed VNT only for EB6C vs the others ($P \leq .009$). VariScreen was the only safe test regardless of sex or etiology (missed VNT $\leq 5\%$). Moreover, VariScreen secured screening without missed VNT in patients with model for end-stage liver disease scores > 10 . This overall strategy performed better than a

safety did not differ significantly among populations.

Conclusions: In a retrospective study of data from 2368 patients with chronic liver disease, we found that the B6C are safe whereas the EB6C are unsafe, based on missed VNT. The VariScreen algorithm performed well in patients with chronic liver disease of any etiology or severity—it is the only test that safely rules out VNTs and can be used in clinical practice.

Words: 363 ($n \leq 260$)

Keywords:

Portal hypertension; non-invasive diagnosis; Baveno VI criteria; cirrhosis; MELD

Non-invasive tests (NITs) have been widely validated as accurate predictors of liver fibrosis or portal hypertension complications such as esophageal varices (EVs). In 2015, the original Baveno VI criteria (B6C), based on platelets and LSM, made it possible to rule out varices needing treatment (VNT) ¹. The B6C's aim is to avoid unnecessary esophagogastroduodenoscopy (endoscopy hereafter), the rate of which may be exaggerated by liver fibrosis NITs ². Although largely validated, the clinical impact of the B6C was judged modest, providing a spared endoscopy rate of only about 20% ³. Therefore, several authors proposed improvements. For example, the expanded B6C (EB6C), with more permissive cutoffs, doubled the spared endoscopy rate ⁴. These NITs are applied in chronic liver disease (CLD) according to severity, with a lower limit defined by LSM ≥ 10 kPa and an upper limit by the lack of complications defining compensated advanced CLD (cACLD) according to Baveno VI consensus ¹.

The current situation raises a two-tiered question: which current proposal offers the best combination of platelets and LSM, and can that currently-best proposal be improved? Another issue is to define the population where these NITs can be applied. Thus, in the work reported here, our two main objectives were i) to compare the performance of different NITs employing platelets and LSM to rule out VNT in a large CLD population; and ii) to build and validate a new algorithm to improve VNT screening safety and performance. Our secondary objectives were to evaluate the impact of the clinical background (mainly CLD severity and etiology) on these tests. Thus, we compared an overall strategy (NITs performed in all patients) and a selective strategy (NITs performed in cACLD and endoscopy in all other patients) to screen VNT. Therefore, for this study, we included patients with CLD regardless of main disease etiology and liver function severity.

Participants

In this retro-prospective study, data from patients with CLD were collected from centers participating in studies wherein VNT was usually the main outcome and VCTE the measurement outcome. The protocol conformed to the Declaration of Helsinki and received approval from the ethics review boards of all participating centers. All patients gave informed consent. The included patients were retrospectively randomized in derivation (2/3) and validation (1/3) populations with stratification on VNT and etiology.

Patients included in previously-recorded CLD subpopulations of any main etiology (alcoholic, non-alcoholic fatty liver disease (NAFLD), hepatitis B or C virus) were eligible for inclusion if they had undergone an endoscopy to determine EV size. Platelets count, LSM by VCTE (using the M probe), known EV stage and a maximum delay of six months between endoscopy and LSM or platelets were the four minimum inclusion data. Interventional treatment for portal hypertension complications (TIPS, band ligation or sclerotherapy of EV) was an exclusion criterion. Total patient recruitment counted 4132 patients across 10 subpopulations and 47 centers (details in Supplemental Material). Finally, 2368 patients were included: 1579 in the derivation and 789 in the validation populations (Figure 1). All authors had access to the study data and had reviewed and approved the final manuscript.

Methods

Data collection

The main laboratory data were liver function tests, blood cell count and serum creatinine. The model for end-stage liver disease (MELD) score included bilirubin, international normalized ratio (INR) and creatinine.

Endoscopy - A standard endoscopy was performed by an experienced operator, and grades of EV were recorded.

LSM - All LSMs were performed by experienced operators using VCTE, specifically M probe-equipped Fibroscan devices (Echosens, Paris, France). Patients with LSM using the XL probe only were not included because this group was too small to be separately evaluated (Figure 1). LSM reliability was defined according to the criteria of Boursier et al ⁵.

Definitions

Objectives - The main objective was to evaluate outcome measurements in current and new NITs. The secondary objectives were to evaluate outcome measurements of all NITs as a function of clinical background (age, sex, CLD etiology and severity), and to propose a sequential algorithm based on simple initial clinical rules and with limited test calculations.

Outcome - The main outcome was VNT, defined as large EV (grade 2 or 3, i.e., a diameter ≥ 5 mm).

Outcome measurements - The main outcome measurements for the NITs were the spared endoscopy rate as the performance descriptor and the missed VNT rate as the safety descriptor. The spared endoscopy rate was calculated as the ratio between the number of patients with a low VNT prevalence according to the NIT and the total number of patients. This corresponded to the negative test rate and was dependent on VNT prevalence. The missed VNT rate was calculated with the number of patients with missed VNT as the numerator and the number of patients with VNT as the denominator (other definitions are

0% in CLD with poor liver function (severe CLD hereafter, MELD score ≥ 10 : see Supplemental Material). The negative likelihood ratio (LR-) was a secondary outcome measurement. LR- is the ratio of 1 - sensitivity for VNT (i.e., missed VNT rate) divided by specificity for missed VNT, and therefore theoretically independent of VNT prevalence.

Previous tests evaluated

Three published tests were compared: the B6C, the EB6C and the ANTICIPATE score ⁷. Details are provided in the Supplemental Material.

Construction of new algorithm

Principles - In terms of their construction and application, NITs are either “selective”, i.e., intended for CLDs of a specific severity, or “overall”, i.e., unrestricted. These limits affect cutoff values. Here, we developed the two types but, considering their results, chose to report new selective NITs only in the Supplemental Material. Obviously, the ideal overall NIT must have a high spared endoscopy rate, but it must also provide a missed VNT rate that both falls under the $\leq 5\%$ threshold globally and decreases with liver severity. Indeed, the risk and the severity of EV bleeding markedly increase as a function of liver severity ⁸.

New tests - Currently, in the B6C, platelets and LSM are used with their own cutoffs, producing a cutoff pair. However, this latter varies in the literature. The *platelets/liver elastometry ratio* (PLER) has the strong advantage of circumventing the cutoff pair choice dilemma, since it confers a single cutoff for a fixed rate of missed VNT. We observed that sex, age, etiology and INR or MELD score were independent significant predictors of VNT in addition to PLER. Moreover, PLER significantly interacted with etiology, sex and INR. Therefore, we constructed a score for VNT diagnosis based on these five variables, called

in the Supplemental Material. Cutoffs of the PLER or PLEASE scores were determined to provide binary tests indicating a VNT rate $\leq 5\%$ or $> 5\%$.

Finally, we developed a sequential algorithm, named VariScreen^{PLI} for *varice screening by platelets, LSM and INR*, which includes three optional steps to reduce calculations whenever possible (Figure 2). VariScreen^{PLI}, called VariScreen hereafter, is different from our previous algorithm, VariScreen^{CE}, based on CirrhoMeter (blood test) and esophageal capsule endoscopy⁹. First, PLER needed to be calculated only in patients at risk of VNT, i.e., those having platelets ≤ 402 G/l and even LSM ≥ 9 kPa. Second, when PLER was ≥ 17 , the VNT risk was 0% and endoscopy unnecessary. Third, when PLER was < 6.2 , the VNT prevalence was $> 5\%$ in all clinical settings and endoscopy was mandatory. The grey zone between PLER scores ≥ 6.2 and < 17 required PLEASE calculation. The three new NITs are available as a free calculator (<http://forge.info.univ-angers.fr/~gh/wstat/pler-please-variscreen.php>).

Statistics

Quantitative variables were expressed as mean \pm standard deviation and compared using the Student t test or analysis of variance (ANOVA). Qualitative variables were expressed as proportions and compared using the Chi² test or Fisher test when unpaired and the Cochran or McNemar test when paired. Data were reported according to STARD and Liver FibroSTARD¹⁰ statements and analyzed on a partial intention-to-diagnose basis. Thus, all data were analyzed irrespective of VCTE reliability criteria⁵ (except in one NAFLD subpopulation¹¹) but missing data were not replaced and patients with unsuccessful examinations (LSM and endoscopy) were not included. Scores including independent predictors were determined by binary logistic regression. Sample size calculation is quoted in the Supplemental Material.

USA).

RESULTS

Patient characteristics

Derivation population - Nearly two-thirds of the 1579 patients were men (Table 1). Viral CLD was the most frequent etiology (50%), followed by NAFLD (29%) and alcoholic CLD (21%). LSMs ≥ 10 kPa were observed in 93.0% of patients. VNT prevalence was 15.1%. Severe fibrosis (F3 or F4 Metavir) estimated by LSM¹² was observed in 90% of patients. Patients with VNT were more frequently male or of alcoholic etiology and had more severe CLD.

Validation population - None of the characteristics of the 789 patients were significantly different from those of the derivation population (Table 1).

Test characteristics

Performance and safety

Derivation population - All NITs had a missed VNT rate <5% except for the EB6C where it was 10.9% (95% CI: 7.0-15.3) and furthermore significantly higher than the rates of all other NITs (Table 2). The spared endoscopy rate was significantly different between tests ($p < 0.001$, except B6C vs ANTICIPATE: $p = 0.751$): B6C: 23.9% (21.8-26.2), ANTICIPATE: 24.3% (21.9-26.4), VariScreen: 34.5% (31.9-36.8) and EB6C: 41.9% (39.1-44.3).

Validation population - Again, all the NITs had a safe missed VNT rate except for the EB6C where it was 11.7% (6.4-17.9) and furthermore significantly higher than the rates of all other

($p < 0.001$, except B6C vs ANTICIPATE: $p = 0.320$) in the same increasing order: B6C: 23.7% (20.7-26.6), ANTICIPATE: 25.0% (21.9-27.9), VariScreen: 36.2% (32.7-39.4) and EB6C: 44.4% (40.8-47.8). Test performance and safety were not significantly different between populations.

Other characteristics

Diagnostic indices are provided in Table 3. Briefly, the only test with LR- ≤ 0.1 in both populations was VariScreen while EB6C was distinguished by a far higher LR- (0.23) compared to the other tests. The patient rates at each step of the VariScreen algorithm in the derivation population are detailed in Figure S1 in the Supplemental Material.

Sensitivity analysis

NIT performance and safety

The impact of confounding factors (details in the Supplemental Material) on NIT performance and safety is reported in Table S1.

NIT secureness

Securing NITs mainly requires a detailed analysis of their safety in high MELD scores but also consideration as to whether the drop in performance in high MELD scores is worthwhile. Therefore, the impact of the MELD score is detailed here as a function of MELD deciles. The spared endoscopy rate decreased linearly as a function of MELD score in all NITs (Figure 3A). The number of missed VNT remained relatively stable and low (one to two) across MELD deciles in all NITs except EB6C. This latter again showed an unsecured profile, with

performance was low in the high MELD scores. However, their interest does merit analysis through the post-test probability gain. The LR- are presented as a function of MELD deciles in Figure 3C. LR- curves as a function of cumulatively decreasing MELD deciles showed that only two NITs, B6C and VariScreen, gave excellent results (LR- always <0.1) with a perfect LR- ($=0$) in the last three MELD deciles, i.e., the most severe CLD. Finally, VariScreen and B6C were the only secured test with no missed VNT in MELD score ≥ 10 , (Table S2).

Clinical application

PLER lent itself to a simple clinical rule: when platelets (expressed in the usual G/l units) are ≥ 10 times (in viral or alcoholic CLD) or ≥ 15 times (in NAFLD) the LSM value (expressed in kPa with VCTE), VNT prevalence is below $\approx 5\%$ and endoscopy can be avoided (details in Supplemental Material). Although VariScreen performed roughly the same as the PLEASE test alone (Table S3), it showed greater clinical relevance: first, it avoided calculations, e.g., PLER was avoided in 5.6% (5.5-6.7%) and PLEASE in 64.5% (62.1-67.0%) of patients in the derivation population. Second, VariScreen was safer than the PLEASE test as the missed VNT rate was 0% in MELD score ≥ 10 in both populations. Finally, VariScreen offered simplicity via an easy-to-use calculator based on successive, optional, simple queries on patient characteristics (<http://forge.info.univ-angers.fr/~gh/wstat/pler-please-variscreen.php>). The calculator includes a reliability analysis between tests that optimized performance: 35.7% vs 34.5% in the derivation population, $p < 0.001$.

Main results - We found the B6C to be robust and safely applicable regardless of etiology and liver severity. However, the missed VNT rate was borderline in NAFLD (7.4%) and the spared endoscopy rate modest in alcoholic CLD (16.0%). In contrast, despite their high spared endoscopy rate, the EB6C were disqualified because of a high missed VNT rate (11-12%) and a lack of robustness with disparate performance across etiologies and sexes. Our findings on the EB6C's missed VNT rate reflect those of previous studies ⁶. In this work, for the first time, we also validated the ANTICIPATE score as a test to spare endoscopy. It offered good performance and robustness, apart the same limits mentioned above for the B6C: a borderline missed VNT rate in NAFLD (7.4%) and a low spared endoscopy rate in alcoholic CLD (17.2%). Finally, among the safe NITs presented here to spare endoscopy, VariScreen offered the best performance. Our derivation and validation populations had similar characteristics and displayed similar NIT results, aspects that reinforce the external validity of our results.

Originalities - First, our population was the largest to date with a large spectrum in liver function and etiologies. Therefore, the robustness of NITs could be reported according to the three main influencing factors: sex, CLD etiology and severity. We found that the influences of etiology and sex were only moderate but that of severity was marked. Indeed, VNT prevalence, platelets and LSM are strongly related to liver severity ¹³. This results in the reduction of the performance of NITs in severe CLD. Accordingly, the Baveno VI statements suggest applying B6C in cACLD. In comparison to previous studies, we were more able to examine the applicability of NITs according to liver severity. For the B6C, we found an unsatisfactory spared endoscopy rate in severe CLD (8% in the third MELD tertile), confirming the Baveno VI statement that these criteria underperform in severe CLD.

problem. Indeed, VariScreen provided secureness (0% missed VNT in MELD ≥ 10), robustness and better performance in severe CLD (13% spared endoscopy in the third MELD tertile). This figure may seem low, but it reflects the most improved post-test probability, with a perfect LR- of 0 in the highest MELD deciles (Figure 3C). Here, the LR- reflects the effectiveness of the endoscopy-sparing effect of the NIT ¹⁴.

Finally, we had enough patients to evaluate the new NITs in NAFLD (29%) and alcoholic CLD (21%) and moreover our present study is the first to validate the B6C in the latter. Our results suggest that several NITs can be used in alcoholic CLD, albeit with a significantly lower spared endoscopy rate (<20% except for VariScreen at 25%). Concerning NAFLD, the spared endoscopy rates of all NITs were highly satisfactory but the missed VNT rate was safe only for VariScreen. In the present study, we lowered the LSM cutoff for VNT presence from 10 to 9 kPa. This increased endoscopy rate by 0.5% but reduced missed VNT by 1.4%, which significantly impacted the safety limit (details in Supplemental Material).

Limits - Our study comprised numerous centers, which necessarily induced variability.

However, the performances of the published tests observed here were very similar to those of previous studies ^{2, 3, 15}, which validates our data record. Moreover, there were no significant differences between the 10 subpopulations for the main outcomes. Our study was retrospective in design, but data were prospectively recorded and VNT was the main outcome regarding data collection (except in the prospective CIRVIR and NAFLD-B6C studies). The Baveno VI statements define VNT as large EV and grade one EV with red signs ¹. Here, we did not include grade one EV with red signs for two reasons. First, many data had been collected before the Baveno VI statements were published. Second, significant differences in the prevalence of grade one EV with red signs was observed between centers ¹⁶. The issue of grade one EV with red signs merits a fully prospective study. We decided to analyze the

upper limit for B6C application). Indeed, the quantitative MELD score has several advantages over qualitative variables reflecting complications. It results in less recording variability and treatment influence, and furthermore eases interaction testing and cutoff determination.

Within this line, treatments were not evaluated due to their multiplicity and variable recordings. However, the B6C have been validated in viral cirrhosis with sustained virological response or under non-selective beta-blockers ¹⁷. Moreover, normal transaminases, considered as a surrogate marker of inactive CLD, did not impact the NIT hierarchy in the present study. Concerning NAFLD, we note that XL probe LSMs must not be used in NITs for VNT constructed with M probe LSMs (details in Supplemental Material): the performance of the VCTE XL probe needs further study. The VariScreen formula seems somewhat complex since there are cutoffs for platelets and/or LSMs specific to each etiology and sex, but they only reflect interactions by confounding factors that must be considered according to recent recommendations on diagnostic scores ¹⁰.

The most important putative limit to consider is the clinical application of NITs for VNT as a function of liver severity. VariScreen provided the best safety profile (no missed VNT) in severe CLD (MELD score ≥ 10) regardless of etiology. Finally, we consider that a secured strategy, i.e., one without missed VNT in severe CLD, can be applied despite a low spared endoscopy rate for two reasons. Firstly, it is very safe as concerns risky VNT. Secondly, it also increases safety in clinical use since there is no MELD limit to inadvertently neglect (a possibility with unsafe tests in real life). The 5.0% missed VNT rate of VariScreen was borderline in the validation population. Therefore, independent validation is needed (see requirements in Supplemental Material).

Clinical application - PLER can be used at bedside: when the platelet count (G/l) is ≥ 10 -fold the LSM value (kPa with VCTE) in viral and alcoholic CLD, or ≥ 15 -fold the LSM value in

250 G/l and LSM at 25 kPa would not need endoscopy. For a better-performing diagnosis, we recommend the VariScreen algorithm, which did not require a calculator in 65% of the patients in the present population. Moreover, as its initial diagnostic tools are commonly employed in the clinic, this algorithm can be applied to the general CLD population in whom the sparing effect will be amplified. Our calculator is freely available for use both in clinical practice (tailored to the patient) and in research (for database analyses). According to the influence of delays between endoscopy and LSM or blood biology on main outcome measurements, a negative NIT should be repeated every six months, corresponding to the usual cirrhosis surveillance interval. Finally, an overall strategy, where NITs are applied to all patients offers better performance than a selective strategy where NITs are applied only in compensated patients (details in Supplemental Material).

Conclusion - In the present study, we comprehensively show that the Baveno VI criteria are robust whatever the etiology or severity of CLD (with a small safety limit in NAFLD), but that the expanded Baveno VI criteria are unsafe due to a high missed VNT rate. We present an easy-to-use and universally-deployable score (PLER) with very simple rules (platelets \geq ten-fold LSM = no endoscopy in viral and alcoholic CLD). We prefer however the sequential VariScreen algorithm for its applicability in all main-etiology CLDs and especially its secureness in severe CLD. VariScreen was the only test for ruling out VNT that was safe for wide use in clinical practice. This overall strategy performed better than a selective strategy restricted to compensated CLD.

1. de Franchis R, Faculty BV. 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.
2. Pateu E, Oberti F, Cales P. The noninvasive diagnosis of esophageal varices and its application in clinical practice. *Clin Res Hepatol Gastroenterol* 2018;42:6-16.
3. Roccarina D, Rosselli M, Genesca J, et al. Elastography methods for the non-invasive assessment of portal hypertension. *Expert Rev Gastroenterol Hepatol* 2018;12:155-164.
4. Augustin S, Pons M, Maurice JB, et al. Expanding the Baveno VI criteria for the screening of varices in patients with compensated advanced chronic liver disease. *Hepatology* 2017;66:1980-1988.
5. Boursier J, Zarski JP, de Ledinghen V, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. *Hepatology* 2013;57:1182-91.
6. Cales P, Buisson F, Ravaioli F, et al. How to clarify the Baveno VI criteria for ruling out varices needing treatment by non-invasive tests. *Liver Int* 2019;39:49-53.
7. Abraldes JG, Bureau C, Stefanescu H, et al. Noninvasive tools and risk of clinically significant portal hypertension and varices in compensated cirrhosis: The "Anticipate" study. *Hepatology* 2016;64:2173-2184.
8. Singal AK, Kamath PS. Model for End-stage Liver Disease. *J Clin Exp Hepatol* 2013;3:50-60.
9. Cales P, Sacher-Huvelin S, Valla D, et al. Large esophageal varice screening by a sequential algorithm using a cirrhosis blood test and optionally capsule endoscopy. *Liver Int* 2018;38:84-93.

- reporting diagnostic accuracy studies on liver fibrosis tests: the Liver-FibroSTARD standards. *J Hepatol* 2015;62:807-15.
11. Petta S, Sebastiani G, Bugianesi E, et al. Noninvasive Prediction of Esophageal Varices by Stiffness and Platelet in Nonalcoholic Fatty Liver Disease Cirrhosis. *J Hepatol* 2018;69:878-885.
 12. Cales P, Boursier J, Oberti F, et al. Cirrhosis diagnosis and liver fibrosis staging: transient elastometry versus cirrhosis blood test. *J Clin Gastroenterol* 2015;49:512-9.
 13. Latorre R, Vaquero J, Rincon D, et al. Determinants of platelet count are different in patients with compensated and decompensated cirrhosis. *Liver Int* 2016;36:232-9.
 14. Dujardin B, Van den Ende J, Van Gompel A, et al. Likelihood ratios: a real improvement for clinical decision making? *Eur J Epidemiol* 1994;10:29-36.
 15. Marot A, Trepo E, Doerig C, et al. Liver stiffness and platelet count for identifying patients with compensated liver disease at low risk of variceal bleeding. *Liver Int* 2017;37:707-716.
 16. Stefanescu H, Marasco G, Cales P, et al. A novel spleen-dedicated stiffness measurement by FibroScan(R) improves the screening of high-risk oesophageal varices. *Liver Int* 2020;40:175-185.
 17. Thabut D, Bureau C, Layese R, et al. Validation of Baveno VI Criteria for Screening and Surveillance of Esophageal Varices in Patients With Compensated Cirrhosis and a Sustained Response to Antiviral Therapy. *Gastroenterology* 2019;156:997-1009.e5.

	Derivation				Validation	p^{ab}
	No VNT	VNT	p^{ac}	All	All	
Patients (n)	1340	239	-	1579	789	-
Age (years)	59.4±10.9	59.1±11.1	0.712	59.3±10.9	59.0±11.2	0.436
Sex (% male)	60.4	74.9	<0.001	62.6	65.1	0.239
Etiology (%):	-	-	0.006	-	-	0.999
Viral	50.3	49.8	0.889	50.2	50.2	1
NAFLD	30.1	22.6	0.020	28.9	28.9	1
Alcoholic	19.6	27.6	0.007	20.8	20.9	1
VNT (%)	0	100	-	15.1	15.2	1
BMI (kg/m ²)	28.6±6.1	27.8±4.0	0.024	28.5±5.8	28.1±5.6	0.204
AST (IU/l)	71±52	74±47	0.433	71±51	74±62	0.427
ALT (IU/l)	67±60	57±45	0.003	66±58	68±59	0.277
Normal transaminases (%)	15.9	10.5	0.040	15.0	14.9	0.949
Albumin (g/l)	40.4±5.4	37.0±5.7	<0.001	39.9±5.6	39.6±5.6	0.262
Bilirubin (μmol/l)	17±14	32±48	<0.001	20±23	19±17	0.488
INR	1.1±0.2	1.3±0.2	<0.001	1.2±0.2	1.2±0.2	0.953
Platelets (G/l)	155±73	116±68	<0.001	149±74	149±76	0.940
Creatinine (μmol/l)	73±33	67±18	0.025	72±31	71±35	0.531
MELD score	9.1±2.6	11.5±3.7	<0.001	9.5±3.0	9.4±3.0	0.485
LSM (kPa)	25±17	39±20	<0.001	27±18	28±19	0.459
LSM ≥10 kPa (%)	92.1	98.3	<0.001	93.0	92.4	0.612
LSM reliability (%):	-	-	0.004	-	-	0.717
Very reliable	24.4	34.1	-	26.1	27.9	-
Reliable	66.4	51.2	-	63.8	63.0	-
Poorly reliable	9.2	14.7	-	10.1	9.0	-
LSM cutoff for VNT (kPa)	-	9.0	-	-	10.6 ^d	-
Metavir F stage by LSM (%):	-	-	<0.001	-	-	0.743
<F3 (≤10.8 kPa)	11.4	2.5	<0.001	10.1	10.8	0.596
F3 (10.9-17.6 kPa)	32.5	11.7	<0.001	29.3	28.1	0.549
F3/4 (17.7-25.7 kPa)	21.4	15.9	0.052	20.8	19.4	0.496
F4 (≥25.8 kPa)	34.7	69.9	<0.001	40.0	41.7	0.434

LSM: liver stiffness measurement, VNT: varices needing treatment, ~~CO: cut-off~~. The comparison of the validation and whole populations as a function of VNT is detailed in Table S4 in Table S5.

VNT.^d One outlier value at 6.3 excluded.

	B6C	EB6C	ANTICIPATE	VARISCREEN	p^a
Derivation population (1579 patients):					
Missed VNT (%)	2.9 (0.9-5.2)	10.9 (7.0-15.3)	4.6 (2.0-7.4)	2.9 (0.9-5.1)	$<0.001^b$
Spared EGD (%)	23.9 (21.8-26.2)	41.9 (39.1-44.3)	24.3 (21.9-26.4)	34.5 (31.9-36.8)	$<0.001^c$
Validation population (789 patients):					
Missed VNT (%)	1.7 (0.0-4.3)	11.7 (6.4-17.9)	0.8 (0.0-3.1)	5.0 (1.0-9.3)	$<0.001^d$
Spared EGD (%)	23.7 (20.7-26.6)	44.4 (40.8-47.8)	25.0 (21.9-27.9)	36.2 (32.7-39.4)	$<0.001^e$
Derivation vs validation population (2368 patients) (p^f):					
Missed VNT	0.723	0.860	0.068	0.373	-
Spared EGD	0.919	0.252	0.723	0.411	-

EGD: esophagogastroduodenoscopy, VNT: varices needing treatment

Color legend: missed VNT (% in VNT): ≤ 5 : safe, >10 : unsafe, intermediate; spared EGD (%): ≥ 20 : satisfactory, <10 : unsatisfactory, intermediate

^a Paired Cochran test

^b Differences in missed VNT rate were significant only between EB6C vs others ($p < 0.001$ by paired McNemar test)

^c Differences in spared EGD rates were significant between all tests ($p \leq 0.001$ except for B6C vs ANTICIPATE: $p = 0.751$)

^d Differences in missed VNT rate were significant only between EB6C vs others ($p \leq 0.039$)

^e Differences in spared EGD rates were significant between all tests ($p \leq 0.001$ except for B6C vs ANTICIPATE: $p = 0.320$)

^f Unpaired Fisher test

descriptors are sensitivity (since 1- sensitivity = missed VNT rate) and LR- (= missed VNT rate/specificity).

	B6C	EB6C	ANTICIPATE	VARISCREEN ^a
Derivation population (1579 patients):				
Accuracy (%) ^b	38.2	53.7	38.0	49.8
Sensitivity (%)	97.1	89.1	95.4	96.7
Specificity (%)	27.7	47.4	27.8	41.4
PPV (%)	19.3	23.2	19.1	22.7
NPV (%)	98.1	96.1	97.1	98.6
LR+	1.34	1.69	1.32	1.65
LR-	0.11	0.23	0.17	0.08
Validation population (789 patients):				
Accuracy (%) ^b	38.4	56.0	39.9	63.1
Sensitivity (%)	98.3	88.3	99.2	95.0
Specificity (%)	27.7	50.2	29.3	57.4
PPV (%)	19.6	24.1	20.1	28.6
NPV (%)	98.9	96.0	99.5	98.5
LR+	1.36	1.77	1.40	2.23
LR-	0.06	0.23	0.03	0.09

PPV: positive predictive value, NPV: negative predictive value, LR: likelihood ratio. The best indices among tests are shown in bold. Color legend: 1- sensitivity or missed VNT (% in VNT): ≤5: safe, >10: unsafe, intermediate; LR-: ≤0.1: excellent (large decrease of post-test VNT probability), >0.1 & ≤0.3: good (significant decrease), >0.3 & ≤0.5: fair (moderate decrease), >0.5 & ≤1: poor (small decrease)

^a Results including reliability analysis

^b Correctly classified

Figure 1. Patient inclusion flowchart.

Figure 2. VariScreen algorithm: VNT screening strategy in CLD, first as a function of biomarkers (platelets, LSM) then ensuing PLER score and PLEASE test as necessary.

Figure 3. VNT diagnosis by NITs as a function of MELD deciles in the whole population.

(A) Spared endoscopy rate (%). (B) Missed VNT number. Whole population. (C) Cumulative LR- of NITs as a function of decreasing MELD deciles. The 9th LR- on the right corresponds to cumulated MELD deciles 10 and 9 and so on in decreasing order. Thus, the LR- of the first MELD decile (#1) corresponds to the global test LR-.

Supplemental Material

The Supplemental Material includes 11 tables and 4 figures.

We thank the co-investigators of the VEB6 (validation of expanded Baveno VI criteria) group as a function of subpopulations:

Monocenter subpopulations:

- Angers (SNIFF 95 study, France): François Buisson, Arthur Berger, Carlotta Carboni (and Bologna), Frédéric Oberti, Isabelle Fouchard, Jérôme Boursier, Adrien Lannes, Gilles Hunault, Paul Calès.
- Bologna (Italy): Elton Dajti, Vanessa Luigina Alemanni, Giovanni Marasco, Federico Ravaioli, Antonio Colecchia, Davide Festi.
- Cluj (Romania): Horia Stefanescu, Oana Farcau, Anca Maniu, Andreea Ardelean, Petra Fischer, Crina Grigoras, Bogdan Procopet.
- Bern (Swiss): Annalisa Berzigotti.
- Toulouse (France): Christophe Bureau, Marie Angèle Robic, Maeva Guillaume.
- Bondy (France): Nathalie Ganne-Carriè, Nathalie Barget.

Multicenter subpopulations:

- VO-VCO (PHRC, France): Sylvie Sacher Huvelin, Paul Calès, Dominique Valla, Christophe Bureau, Anne Olivier, Frédéric Oberti, Jérôme Boursier, Jean Paul Galmiche, Jean Pierre Vinel, Clotilde Duburque, Alain Attar, Isabelle Archambeaud, Robert Benamouzig, Marianne Gaudric, Dominique Luet, Patrice Couzigou, Lucie Planche, Emmanuel Coron, Jean-Baptiste Hiriart, Faiza Chermak, Maude Charbonnier.
- M116 study (Echosens, Italy, Romania, France, United Kingdom): Horia Stefanescu, Giovanni Marasco, Paul Calès, Mirella Fraquelli, Matteo Rosselli, Nathalie Ganne-Carriè,

Andreone, Giuseppe Mazzella, Davide Festi.

- B6C NAFLD group: Salvatore Petta¹, Giada Sebastiani², Elisabetta Bugianesi³, Mauro Vigano⁴, Vincent Wai-Sun Wong⁵, Annalisa Berzigotti⁶, Anna Ludovica Fracanzani⁷, Quentin M. Anstee^{8,9}, Fabio Marra^{10,11}, Marco Barbara¹, Vincenza Calvaruso¹, Calogero Camma¹,

¹Sezione di Gastroenterologia e Epatologia, Di.Bi.M.I.S, Università di Palermo, Italy;

²Division of Gastroenterology and Hepatology, McGill University Health Centre, Montreal, QC, Canada;

³Division of Gastroenterology, Department of Medical Sciences, University of Torino, Torino, Italy;

⁴Hepatology Unit, Ospedale San Giuseppe, University of Milan, Milan, Italy;

⁵Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong;

⁶Swiss Liver Center, Hepatology, University Clinic for Visceral Surgery and Medicine, Inselspital, University of Bern, Switzerland;

⁷Department of Pathophysiology and Transplantation, Ca' Granda IRCCS Foundation, Policlinico Hospital, University of Milan, Italy;

⁸Institute of Cellular Medicine, Faculty of Medical Sciences, Newcastle University, Newcastle Upon Tyne, United Kingdom; ⁹Liver Unit, Newcastle Upon Tyne Hospitals NHS Foundation Trust, Freeman Hospital, Newcastle upon Tyne, United Kingdom;

¹⁰Dipartimento di Medicina Sperimentale e Clinica, University of Florence, Italy;

¹¹Research Center DENOTHE, University of Florence, Italy;

¹²Centre d'Investigation de la Fibrose Hépatique, INSERM U1053, Hôpital Haut-Lévêque, Bordeaux University Hospital, Pessac, France

Stanislas Pol⁴, Hélène Fontaine⁴, Dominique Larrey⁵, Victor De Lédighen⁶, Denis Ouzan⁷, Fabien Zoulim⁸, Dominique Roulot⁹, Albert Tran¹⁰, Jean-Pierre Bronowicki¹¹, Jean-Pierre Zarski¹², Vincent Leroy¹², Ghassan Riachi¹³, Paul Calès¹⁴, Jean-Marie Péron¹⁵, Laurent Alric¹⁶, Marc Bourlière¹⁷, Philippe Mathurin¹⁸, Sebastien Dharancy¹⁸, Jean-Frédéric Blanc¹⁹, Armand Abergel²⁰, Lawrence Serfaty²¹, Ariane Mallat²², Jean-Didier Grangé²³, Pierre Attali²⁴, Yannick Bacq²⁵, Claire Wartelle²⁶, Thông Dao²⁷, Yves Benhamou²⁸, Christophe Pilette²⁹, Christine Silvain³⁰, Christos Christidis³¹, Dominique Capron³², Gérard Thieffin³³, Sophie Hillaire³⁴, Vincent Di Martino³⁵.

¹AP-HP, Hôpital Jean Verdier, Service d'Hépatologie, Bondy, Université Paris 13, Bobigny et INSERM U1162, Université Paris 5, Paris; ²AP-HP, Hôpital Beaujon, Service d'Hépatologie, Clichy; ³CHU Pontchaillou, Service d'Hépatologie, Rennes; ⁴AP-HP, Hôpital Cochin, Département d'Hépatologie et INSERM UMS20 et U1223, Institut Pasteur, Université Paris Descartes, Paris; ⁵Hôpital Saint Eloi, Service d'Hépatologie, Montpellier; ⁶Hôpital Haut-Lévêque, Service d'Hépatologie, Bordeaux; ⁷Institut Arnaud Tzanck, Service d'Hépatologie, St Laurent du Var; ⁸Hôpital Hôtel Dieu, Service d'Hépatologie, Lyon; ⁹AP-HP, Hôpital Avicenne, Service d'Hépatologie, Bobigny; ¹⁰CHU de Nice, Service d'Hépatologie, et INSERM U1065, Université de Nice-Sophia-Antipolis, Nice; ¹¹Hôpital Brabois, Service d'Hépatologie, Vandoeuvre-les-Nancy; ¹²Hôpital Michallon, Service d'Hépatologie, Grenoble; ¹³Hôpital Charles-Nicolle, Service d'Hépatologie, Rouen; ¹⁴CHU d'Angers, Service d'Hépatologie, Angers; ¹⁵Hôpital Purpan, Service d'Hépatologie, Toulouse; ¹⁶CHU Toulouse, Service de Médecine Interne-Pôle Digestif UMR 152, Toulouse; ¹⁷Hôpital Saint Joseph, Service d'Hépatologie, Marseille; ¹⁸Hôpital Claude Huriez, Service d'Hépatologie, Lille; ¹⁹Hôpital St André, Service d'Hépatologie, Bordeaux; ²⁰Hôpital Hôtel Dieu, Service d'Hépatologie, Clermont-

Henri Mondor, Service d'Hépatologie, Créteil; ²³AP-HP, Hôpital Tenon, Service d'Hépatologie, Paris; ²⁴AP-HP, Hôpital Paul Brousse, Service d'Hépatologie, Villejuif; ²⁵Hôpital Trousseau, Unité d'Hépatologie, CHRU de Tours; ²⁶Hôpital d'Aix-En-Provence, Service d'Hépatologie, Aix-En-Provence; ²⁷Hôpital de la Côte de Nacre, Service d'Hépatologie, Caen; ²⁸AP-HP, Groupe Hospitalier de La Pitié-Salpêtrière, Service d'Hépatologie, Paris; ²⁹CHU Le Mans, Service d'Hépatologie, Le Mans; ³⁰CHU de Poitiers, Service d'Hépatologie, Poitiers; ³¹Institut Mutualiste Montsouris, Service d'Hépatologie, Paris; ³²Hôpital Amiens Nord, Service d'Hépatologie, Amiens; ³³Hôpital Robert Debré, Service d'Hépatologie, Reims; ³⁴Hôpital Foch, Service d'Hépatologie, Suresnes; ³⁵Hôpital Jean Minjoz, Service d'Hépatologie, Besançon. France.

We also thank Kevin L. Erwin for writing assistance (English proofreading), Gilles Hunault for implementing the calculator, and the sponsors of the subpopulations (who had no other role in the present study): VO-VCO study: the PHRC (Health Ministry, France); M116 study: Echosens, Paris, France; CO12 CirVir cohort: ANRS, France.

4132 eligible patients: CLD with EGD

↓ → Missing minimum data: n = 477 (platelets: 52, LSM: 434)

3655 patients with minimum data

↓ → Missing delays: n = 36 (blood sample: 24, LSM: 12)

3619 patients with known delays

↓ → delays > 6 months: n = 878 (blood sample: 719, LSM: 672)

2741 patients with appropriate delays

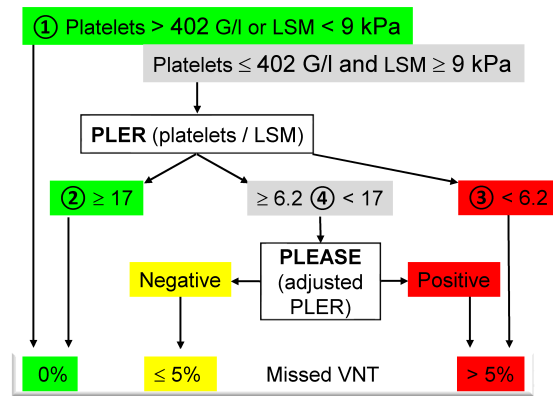
↓ → Other etiologies: n = 179

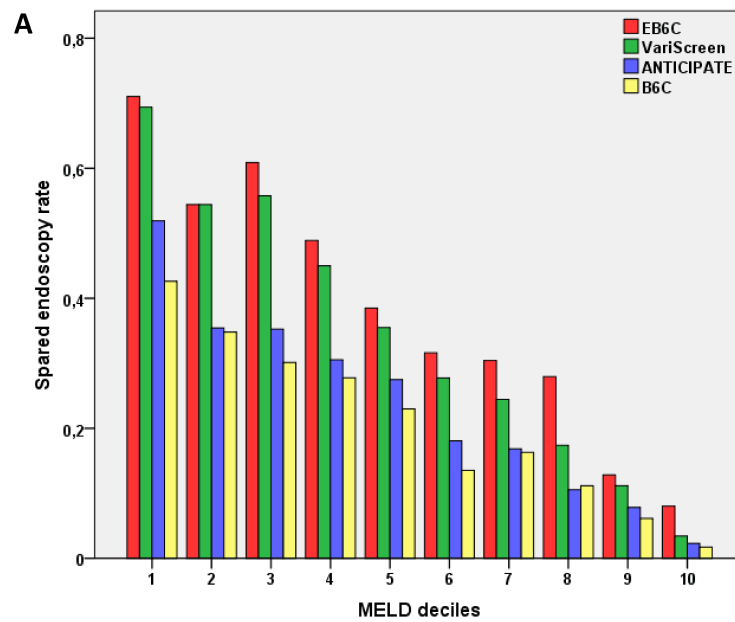
2562 patients with appropriate etiology

↓ → Inappropriate LSM: n = 194 (unreliable: 56, XL probe only: 138)

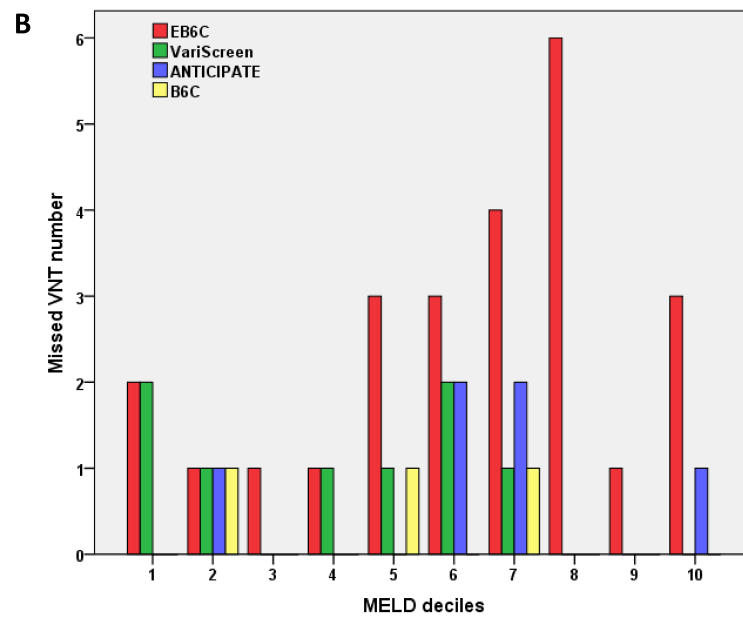
2368 patients included

↓
Derivation 2/3 (n = 1579) ↘ Validation 1/3 (n = 789)

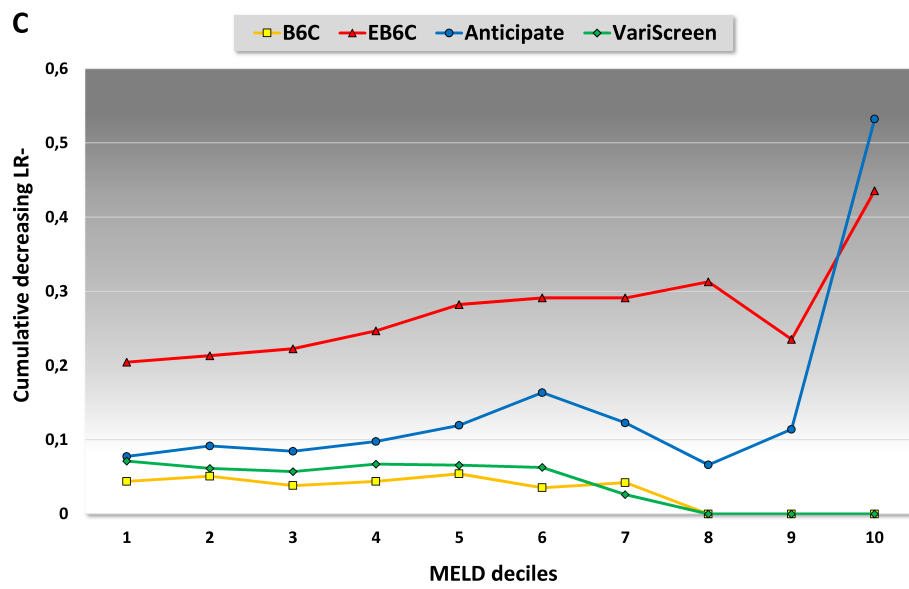




Lower MELD score: 6.4 6.8 7.4 7.7 8.2 8.7 9.2 10.0 11.0 13.3



Lower MELD score: 6.4 6.8 7.4 7.7 8.2 8.7 9.2 10.0 11.0 13.3



Background: on platelets and liver stiffness measurements, the Baveno VI criteria (B6C), the expanded B6C (EB6C), and the ANTICIPATE score can be used to rule out varices needing treatment (VNT) in patients with compensated chronic liver disease.

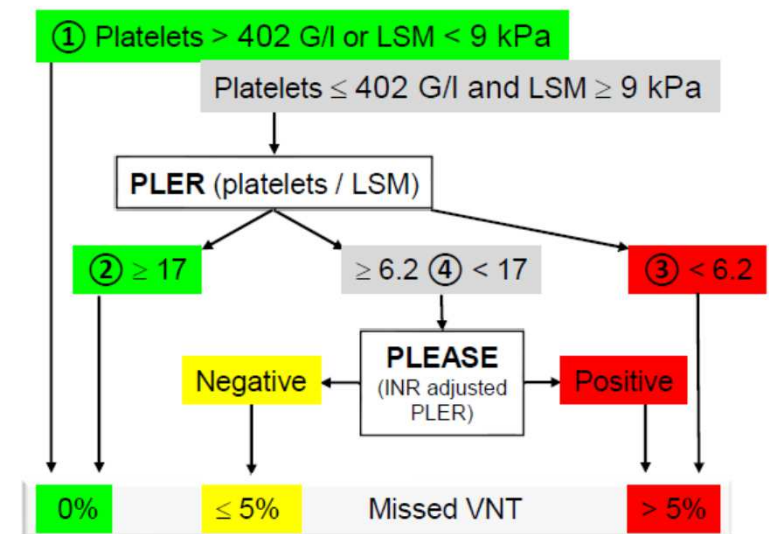
Findings: A retrospective study of data from 2368 patients with chronic liver disease found that the B6C are safe whereas the EB6C are unsafe, based on missed VNT. The VariScreen algorithm performs well in patients with chronic liver disease of any etiology or severity.

Implications for patient care: The VariScreen is the only test that safely rules out VNTs and can be used in clinical practice regardless of severity.

The non-invasive VariScreen algorithm rules out varices needing treatment irrespective of the severity of chronic liver diseases in around 35% of patients



VariScreen uses blood platelets, liver stiffness measurement (LSM) and optionally INR:



Clinical Gastroenterology
and Hepatology