2018-2019

THÈSE

pour le

DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE

Qualification en hépato-gastro-entérologie

Personalized platelet/liver stiffness ratio improves and secures the screening of esophageal varices needing treatment

Le ratio plaquettes/élastométrie hépatique personnalisé améliore la sûreté et l'efficacité du dépistage de varices œsophagiennes à risque de saignement

François BUISSON

Né le 30/04/1991 à Epinal

Sous la direction du Pr Paul CALES

Membres du jury

Pr François-Xavier CAROLI-BOSC | Président

Pr Paul CALES | Directeur

Pr Jérôme BOURSIER | Membre

Dr Frédéric OBERTI | Membre

Soutenue publiquement le :12/09/2019

FACULTÉ DE SANTÉ
UNIVERSITÉ D'ANGERS

ENGAGEMENT DE NON PLAGIAT

Je, soussigné Buisson François
déclare être pleinement conscient que le plagiat de documents ou d'une partie d'un document publiée sur toutes formes de support, y compris l'internet, constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée.
En conséquence, je m'engage à citer toutes les sources que j'ai utilisées pour écrire ce rapport ou mémoire.

signé par l'étudiant(e) le 27/07/2019

LISTE DES ENSEIGNANTS DE L'UFR SANTÉ D'ANGERS

Directeur de l'UFR : Pr Nicolas Lerolle
Directeur adjoint de l'UFR et directeur du département de pharmacie : Pr Frédéric Lagarce
Directeur du département de médecine :

PROFESSEURS DES UNIVERSITÉS

ABRAHAM Pierre	Physiologie	Médecine
ANNWEILER Cédric	Gériatrie et biologie du vieillissement	Médecine
ASFAR Pierre	Réanimation	Médecine
AUBE Christophe	Radiologie et imagerie médicale	Médecine
AUGUSTO Jean-François	Néphrologie	Médecine
AZZOUZI Abdel Rahmène	Urologie	Médecine
BARON-HAURY Céline	Médecine générale	Médecine
BAUFRETON Christophe	Chirurgie thoracique et cardiovasculaire	Médecine
BENOIT Jean-Pierre	Pharmacotechnie	Pharmacie
BEYDON Laurent	Anesthésiologie-réanimation	Médecine
BIGOT Pierre	Urologie	Médecine
BONNEAU Dominique	Génétique	Médecine
BOUCHARA Jean-Philippe	Parasitologie et mycologie	Médecine
BOUVARD Béatrice	Rhumatologie	Médecine
BOURSIER Jérôme	Gastroentérologie ; hépatologie	Médecine
BRIET Marie	Pharmacologie	Médecine
CAILLIEZ Eric	Médecine générale	Médecine
CALES Paul	Gastroentérologe ; hépatologie	Médecine
CAMPONE Mario	Cancérologie ; radiothérapie	Médecine
CAROLI-BOSC François-xavier	Gastroentérologie ; hépatologie	Médecine
CHAPPARD Daniel	Cytologie, embryologie et cytogénétique	Médecine
CONNAN Laurent	Médecine générale	Médecine
COUTANT Régis	Pédiatrie	Médecine
COUTURIER Olivier	Biophysique et médecine nucléaire	Médecine
CUSTAUD Marc-Antoine	Physiologie	Médecine
DE BRUX Jean-Louis	Chirurgie thoracique et cardiovasculaire	Médecine
DESCAMPS Philippe	Gynécologie-obstétrique	Médecine
DINOMAIS Mickaël	Médecine physique et de réadaptation	Médecine
DIQUET Bertrand	Pharmacologie	Médecine
DUCANCELLE Alexandra	Bactériologie-virologie ; hygiène hospitalière	Médecine
DUVAL Olivier	Chimie thérapeutique	Pharmacie
DUVERGER Philippe	Pédopsychiatrie	Médecine
EVEILLARD Mathieu	Bactériologie-virologie	Pharmacie
FANELLO Serge	Épidémiologie ; économie de la santé et prévention	Médecine
FAURE Sébastien	Pharmacologie physiologie	Pharmacie
FOURNIER Henri-Dominique	Anatomie	Médecine
FURBER Alain	Cardiologie	Médecine
GAGNADOUX Frédéric	Pneumologie	Médecine
GARNIER François	Médecine générale	Médecine
GASCOIN Géraldine	Pédiatrie	Médecine
GOHIER Bénédicte	Psychiatrie d'adultes	Médecine
GRANRY Jean-Claude	Anesthésiologie-réanimation	Médecine
GUARDIOLA Philippe	Hématologie ; transfusion	Médecine
GUILET David	Chimie analytique	Pharmacie
HAMY Antoine	Chirurgie générale	Médecine
HUNAULT-BERGER Mathilde	Hématologie ; transfusion	Médecine

IFRAH Norbert	Hématologie ; transfusion	Médecine
JEANNIN Pascale	Immunologie	Médecine
KEMPF Marie	Bactériologie-virologie ; hygiène hospitalière	Médecine
LACCOURREYE Laurent	Oto-rhino-laryngologie	Médecine
LAGARCE Frédéric	Biopharmacie	Pharmacie
LARCHER Gérald	Biochimie et biologie moléculaires	Pharmacie
LASOCKI Sigismond	Anesthésiologie-réanimation	Médecine
LEGRAND Erick	Rhumatologie	Médecine
LERMITE Emilie	Chirurgie générale	Médecine
LEROLLE Nicolas	Réanimation	Médecine
LUNEL-FABIANI Françoise	Bactériologie-virologie ; hygiène hospitalière	Médecine
MARCHAIS Véronique	Bactériologie-virologie	Pharmacie
MARTIN Ludovic	Dermato-vénéréologie	Médecine
MENEI Philippe	Neurochirurgie	Médecine
MERCAT Alain	Réanimation	Médecine
MERCIER Philippe	Anatomie	Médecine
PAPON Nicolas	Parasitologie mycologie	Pharmacie
PASSIRANI Catherine	Chimie générale	Pharmacie
PELLIER Isabelle	Pédiatrie	Médecine
PICQUET Jean	Chirurgie vasculaire ; médecine vasculaire	Médecine
PODEVIN Guillaume	Chirurgie infantile	Médecine
PROCACCIO Vincent	Génétique	Médecine
PRUNIER Fabrice	Cardiologie	Médecine
REYNIER Pascal	Biochimie et biologie moléculaire	Médecine
RICHARD Isabelle	Médecine physique et de réadaptation	Médecine
RICHOMME Pascal	Pharmacognosie	Pharmacie
RODIEN Patrice	Endocrinologie, diabète et maladies métaboliques	Médecine
ROHMER Vincent	Endocrinologie, diabète et maladies métaboliques	Médecine
ROQUELAURE Yves	Médecine et santé au travail	Médecine
ROUGE-MAILLART Clotilde	Médecine légale et droit de la santé	Médecine
ROUSSEAU Audrey	Anatomie et cytologie pathologiques	Médecine
ROUSSEAU Pascal	Chirurgie plastique, reconstructrice et esthétique	Médecine
ROUSSELET Marie-Christine	Anatomie et cytologie pathologiques	Médecine
ROY Pierre-Marie	Thérapeutique	Médecine
SAINT-ANDRE Jean-Paul	Anatomie et cytologie pathologiques	Médecine
SAULNIER Patrick	Biophysique pharmaceutique et biostatistique	Pharmacie
SERAPHIN Denis	Chimie organique	Pharmacie
SUBRA Jean-François	Néphrologie	Médecine
UGO Valérie	Hématologie ; transfusion	Médecine
URBAN Thierry	Pneumologie	Médecine
VAN BOGAERT Patrick	Pédiatrie	Médecine
VENIER Marie-Claire	Pharmacotechnie	Pharmacie
VERNY Christophe	Neurologie	Médecine
WILLOTEAUX Serge	Radiologie et imagerie médicale	Médecine

MAÎTRES DE CONFÉRENCES

ANGOULVANT Cécile	Médecine Générale	Médecine
ANNAIX Véronique	Biochimie et biologie moléculaires	Pharmacie
BAGLIN Isabelle	Pharmaco-chimie	Pharmacie
BASTIAT Guillaume	Biophysique et biostatistique	Pharmacie
BEAUVILLAIN Céline	Immunologie	Médecine
BELIZNA Cristina	Médecine interne	Médecine
BELLANGER William	Médecine générale	Médecine
BELONCLE François	Réanimation	Médecine
BENOIT Jacqueline	Pharmacologie et pharmacocinétique	Pharmacie
BIERE Loïc	Cardiologie	Médecine
BLANCHET Odile	Hématologie ; transfusion	Médecine
BOISARD Séverine	Chimie analytique	Pharmacie
CAPITAIN Olivier	Cancérologie ; radiothérapie	Médecine
CASSEREAU Julien	Neurologie	Médecine
CHEVAILLER Alain	Immunologie	Médecine
CHEVALIER Sylvie	Biologie cellulaire	Médecine
CLERE Nicolas	Pharmacologie	Pharmacie
COLIN Estelle	Génétique	Médecine
DE CASABIANCA Catherine	Médecine générale	Médecine
DERBRE Séverine	Pharmacognosie	Pharmacie
DESHAYES Caroline	Bactériologie virologie	Pharmacie
FERRE Marc	Biologie moléculaire	Médecine
FLEURY Maxime	Immunologie	Pharmacie
FORTRAT Jacques-Olivier	Physiologie	Médecine
HAMEL Jean-François	Biostatistiques, informatique médicale	Médicale
HELESBEUX Jean-Jacques	Chimie organique	Pharmacie
HINDRE François	Biophysique	Médecine
JOUSSET-THULLIER Nathalie	Médecine légale et droit de la santé	Médecine
LACOEUILLE Franck	Biophysique et médecine nucléaire	Médecine
LANDREAU Anne	Botanique et Mycologie	Pharmacie
LEGEAY Samuel	Pharmacologie	Pharmacie
LE RAY-RICHOMME Anne-Marie	Valorisation des substances naturelles	Pharmacie
LEPELTIER Elise	Chimie générale Nanovectorisation	Pharmacie
LETOURNEL Franck	Biologie cellulaire	Médecine
LIBOUBAN Hélène	Histologie	Médecine
MABILLEAU Guillaume	Histologie, embryologie et cytogénétique	Médecine
MALLET Sabine	Chimie Analytique et bromatologie	Pharmacie
MAROT Agnès	Parasitologie et mycologie médicale	Pharmacie
MAY-PANLOUP Pascale	Biologie et médecine du développement et de la reproduction	Médecine
MESLIER Nicole	Physiologie	Médecine
MOUILLIE Jean-Marc	Philosophie	Médecine
NAIL BILLAUD Sandrine	Immunologie	Pharmacie
PAPON Xavier	Anatomie	Médecine
PASCO-PAPON Anne	Radiologie et imagerie médicale	Médecine
PECH Brigitte	Pharmacotechnie	Pharmacie
PENCHAUD Anne-Laurence	Sociologie	Médecine
PETIT Audrey	Médecine et santé au travail	Médecine
PIHET Marc	Parasitologie et mycologie	Médecine
PRUNIER Delphine	Biochimie et biologie moléculaire	Médecine
RIOU Jérémie	Biostatistique	Pharmacie
ROGER Emilie	Pharmacotechnie	Pharmacie

SCHINKOWITZ Andréas	Pharmacognosie	Pharmacie
SIMARD Gilles	Biochimie et biologie moléculaire	Médecine
TANGUY-SCHMIDT Aline	Hématologie ; transfusion	Médecine
TRZEPIZUR Wojciech	Pneumologie	Médecine

AUTRES ENSEIGNANTS

AUTRET Erwan	Anglais	Médecine
BARBEROUSSE Michel	Informatique	Médecine
BRUNOIS-DEBU Isabelle	Anglais	Pharmacie
CHIKH Yamina	Économie-Gestion	Médecine
FISBACH Martine	Anglais	Médecine
O'SULLIVAN Kayleigh	Anglais	Médecine

PAST

CAVAILLON Pascal	Pharmacie Industrielle	Pharmacie
LAFFILHE Jean-Louis	Officine	Pharmacie
MOAL Frédéric	Physiologie	Pharmacie

ATER

FOUDI Nabil (M)	Physiologie et communication cellulaire	Pharmacie
HARDONNIERE Kévin	Pharmacologie - Toxicologie	Pharmacie
WAKIM Jamal (Mme)	Biochimie et biomoléculaire	Médecine

AHU

BRIS Céline	Biochimie et biologie moléculaires	Pharmacie
LEROUX Gaël	Toxico	Pharmacie
BRIOT Thomas	Pharmacie Galénique	Pharmacie
CHAPPE Marion	Pharmacotechnie	Pharmacie

CONTRACTUEL

VIAULT Guillaume	Chimie	Pharmacie

REMERCIEMENTS

Au Pr CALES de m'avoir accompagné et aidé tout au long de cette thèse mais également pour la formation qu'il m'a délivré tout au long de mon internat.

Au Pr CAROLI-BOSC pour sa gentillesse, ses conseils lors des moments de doute et pour son accompagnement.

Aux membres du jury de leur participation à cette thèse.

Aux membres de toutes les équipes ayant participé à ce projet (Bologne, Cluj...) et à Carlotta Carboni de sa gentillesse, son écoute, et son aide lors des recherches pour cette thèse.

À Arthur Berger de son aide lors de la réalisation de cette thèse et ce qu'il m'a appris en gastro-entérologie.

À l'ensemble de l'équipe médicale et paramédicale du CHU d'Angers pour ces 4 années, dont je compte bien tirer le meilleur pour leur faire honneur dans mon application quotidienne de la gastro-entérologie.

À l'équipe de proctologie de Bagatelle, pour leur formation intensive médico-chirurgicale que je compte bien mettre en pratique.

À mes parents et ma sœur, soutien inconditionnel quelles que soient les épreuves ou les joies.

À ma famille qui m'accompagne au quotidien malgré la distance (et qui affiche encore un air d'incompréhension très amusant, suivi d'un « mais, pourquoi ? » à l'annonce de ma spécialité!).

À mes co-internes qui ont partagé mon quotidien, m'ont supporté pendant ces 4 ans, et qui savent l'importance qu'ils ont à mes yeux.

À Philippe JAMET et Mathilde FRON, qui m'ont donné l'envie d'être gastro-entérologue.

À Claire BOYER, Claire CARIOU et Claire LEGALL-GRIMAUX, les trois piliers de mes années de médecine, qui m'ont tant apporté, et pour qui ma reconnaissance est sans borne.

Et enfin à tous ceux qui ont embelli ma route durant ces 4 ans et qui sont malheureusement bien trop nombreux pour que je puisse établir une liste exhaustive ici.

ABBREVIATIONS LIST

B6C	Baveno VI criteria
cACLD	Compensated advanced chronic liver disease
CLD	Chronic liver disease
CO	Cut-off
EAST	Expanded amplification spared technique
EB6C	Expanded Baveno VI criteria
EGD	Oesogastroduodenoscopy
EV	Esophageal varices
LSM	Liver stiffness measurement
MELD	Model for end-stage liver disease
NAFLD	Non alcoholic fatty liver disease
NIT	Non-invasive test
PLER	Platelet / liver elastography ratio
PLR	Platelet / LSM ratio
PLRIES	Platelet / LSM ajusted on INR, etiology and INR
VCTE	Vibration-controlled transient elastography
VNT	Varices needing treatment

PLAN

AUTORS

Arthur Berger,*[1] Federico Ravaioli,*[2] Oana Farcau,[3] Davide Festi,[2] Horia Stefanescu,[3] François Buisson,[1] Pierre Nahon,[4] Christophe Bureau,[5] Nathalie Ganne-Carrié,[4] Annalisa Berzigotti,[6] Victor de Ledinghen,[7] Salvatore Petta,[8] Paul Calès[1] and multicenter groups: VO-VCO, ANRS CO12 CIRVIR, M116 and VEB6.[9] Angers, Bologna, Cluj, Bondy, Toulouse, Bern, Bordeaux.

* Author names in bold designate shared 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

[6] Hepatology, Swiss Liver Center, Visceral Surgery and Medicine Clinic (UVCM), Inselspital, 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

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

Writing Assistance:

Kevin L. Erwin for English proofreading (institutional support)

Grant Support:

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

Author contributions:

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

Federico Ravaioli: 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-Carrié: 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

VEB6 group: data collection

Authors' e-mails:

arthur.berger.bx@gmail.com

f.ravaioli@unibo.it

oana.farcau@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

ABSTRACT

Background & Aims: Based on platelets and liver stiffness measurement (LSM by vibration-controlled transient elastography), the Baveno VI criteria (B6C), the expanded B6C (EB6C) or the ANTICIPATE score can be used to rule out varices needing treatment (VNT). We aimed to evaluate and improve these tests.

Methods: 2368 patients were randomized in derivation (n=1579) and validation (n=789) populations with chronic liver diseases (CLD) of various etiologies and severities in a multicenter retro-prospective study. Published tests were compared to two new tests: PLER (platelets/LSM ratio) and PLESIR (Platelets/LSM ratio adjusted on etiology/sex/INR).

Results: Patient characteristics in the derivation population were: VNT: 15.1%, etiologies: viral: 50.2%, NAFLD: 28.9%, alcoholic: 20.8%, MELD score: 9.5 +/- 3.0, LSM≥10kPa: 93.0%. Spared endoscopy rates (with missed VNT rates in parentheses for safety) were, in increasing order: B6C: 23.9% (2.9%), ANTICIPATE: 24.3% (4.6%), PLER: 26.6% (4.6%), PLESIR: 34.8% (3.3%) and EB6C: 41.9% (10.9%). Differences in spared endoscopy rates were significant between tests ($p \leq 0.001$) except for B6C vs ANTICIPATE. Differences in missed VNT rates were significant only between EB6C vs others ($p \leq 0.009$). Test performance and safety were not significantly different between populations. PLESIR was the only safe test (missed VNT $\leq 5\%$) whatever the sex or etiology. A VariScreen algorithm, based successively on platelets or LSM then platelets/LSM then PLESIR in 35% of patients, secured screening (no missed VNT in poor liver function); its spared endoscopy and missed VNT rates were 35.7% and 3.3% respectively.

Conclusions: B6C are safe for missed VNT rate regardless of CLD etiology and severity, and regardless of sex; EB6C are unsafe and no longer recommended.

To improve current VNT screening, we propose the sequential VariScreen algorithm applicable to any main-etiology CLD.

Lay summary

Currently the Baveno VI criteria can be used to avoid unnecessary endoscopy for the screening of varices needing treatment. Here, we describe new tests and a VariScreen algorithm that spare endoscopy in 36% of patients with frequent etiologies of chronic liver disease. The diagnostic strategy is personalized according to sex, etiology and liver function, and it provides safety by not missing varices that need treatment in patients with poor liver function.

INTRODUCTION

Non-invasive tests (NITs) are among the major recent advances made in hepatology. They have been widely validated as accurate predictors of a number of liver pathologies, including not only liver fibrosis but also portal hypertension and its complications such as esophageal varices (EVs). Indeed, several NITs have been developed specifically for EVs, with platelets quickly standing out as a particularly pertinent biomarker [1]. More recently, liver stiffness measurement (LSM), as assessed by vibration-controlled transient elastography (VCTE), has also been shown to be effective for EV assessment [1]. Resultantly, platelets and LSM have been associated in several NITs [1, 2], but the clinical utility of these latter has been disputed. In 2015, the original Baveno VI criteria (B6C), based on platelets and LSM, made it possible to rule out varices needing treatment (VNT) [3]. Although largely validated [4-6], the clinical impact of the B6C was judged modest, providing a spared esophagogastroduodenoscopy (EGD) rate of only about 20% [5]. Therefore, several authors proposed improvements [7-10]. Augustin *et al*. proposed the expanded B6C (EB6C), which, with more permissive cut-offs (COs), enabled a doubling of the spared EGD rate [8]. These NITs are applied in chronic liver disease (CLD) according to severity, with a lower limit defined by LSM $\geq$ 10 kPa and an upper limit by the lack of complications defining compensated advanced CLD (cACLD) [3]. The B6C's aim is to avoid unnecessary EGDs, the rate of which may be exaggerated by NITs for liver fibrosis in cACLD [1].

The current situation raises a multi-tiered question: which current proposal offers the best combination of platelets and LSM, and can it be improved? Thus, our two main objectives with the present study 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 with the goal of improving the performance and safety of the combined use of these biomarkers. Our secondary objectives were to evaluate the impact of the clinical background (e.g. CLD severity and etiology) on the performance and safety of these tests. Therefore, for this study, we included patients with CLD regardless of main disease etiology and liver function severity.

PATIENTS AND METHOD

Participants

In this retro-prospective, international study, data from patients with CLD were collected from a number of centers participating in several studies with VNT usually as the main outcome and VCTE as the measurement outcome. The protocol conformed to the Declaration of Helsinki and was approved by the ethical review board of each participating center. All patients gave informed consent. Patients included were 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 EGD to determine EV size. Platelets count, LSM by VCTE (using the M probe), known EV stage and a maximum delay of six months between EGD and LSM or platelets (see below) 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 47 centers (details in Supplemental Material). Finally, 2368 patients were included: 1579 in the derivation and 789 in the validation populations (Figure 1).

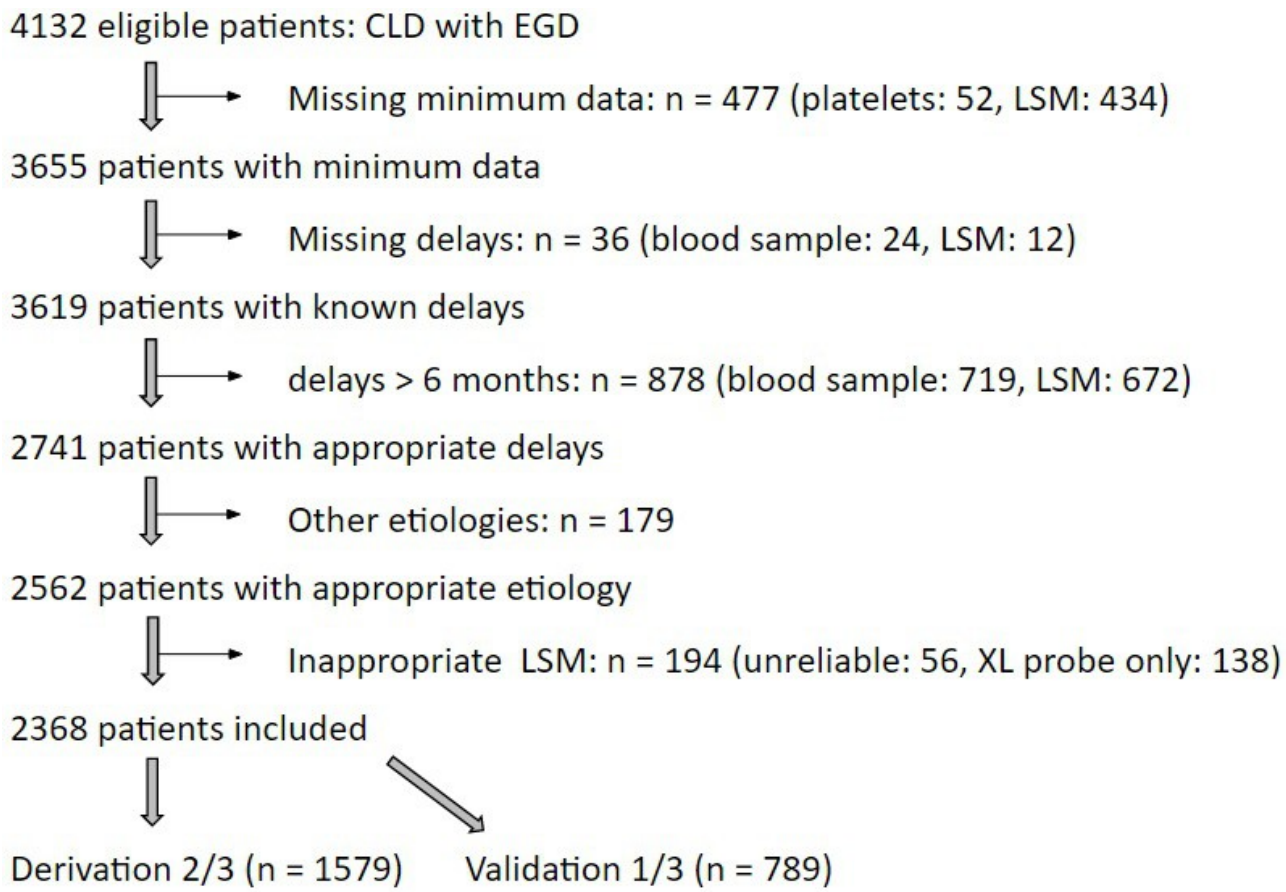


Fig. 1. Patient inclusion flowchart

Methods

Data collection

Dates - A preliminary evaluation testing the impact of delays between examinations on the outcome measurements of B6C and EB6C showed that results stabilized for delays ≤6 months (Table S1). Therefore, patients with delays >6 months were excluded.

Clinical data - The main clinical data were age, sex, height, body weight and CLD etiology. The main laboratory data were liver function tests, blood cell count and serum creatinine levels (measured in each center). The MELD score was calculated according to the original formula including bilirubin, INR and creatinine [11].

Endoscopy - A standard EGD was performed by an experienced operator blinded to other examinations. Grade of EV and presence of red signs on EV were recorded.

LSM - All LSMs were performed by experienced operators with VCTE using M probe-equipped Fibroscan devices (Echosens, Paris, France) after at least three hours of patient fasting according to the manufacturer's recommendations. Patients with LSM using the XL probe only were not included (Figure 1). The technical characteristics are detailed elsewhere [12].

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 [13]).

Outcome measurements - The main outcome measurements were the spared EGD rate as the performance descriptor and the missed VNT rate of NITs as the safety descriptor.

The spared EGD rate was calculated as the ratio between the number of patients with a low VNT risk according to the NIT and the total number of patients. It corresponded to the negative test rate and was dependent on VNT prevalence. We defined spared EGD rates as clinically unsatisfactory when they were $\leq$ 10%, satisfactory when they were $\geq$ 20% (corresponding to the mean rate for B6C in the literature), and intermediate when they were 10-20% (since these included previously reported B6C rates).

The missed VNT rate was calculated with the number of patients with missed VNT as the numerator and the number of patients with VNT (privileged definition, sole reported here), as the denominator [14].

It corresponded to 1 - sensitivity for VNT and was thus theoretically independent of VNT prevalence. We defined missed VNT rates as acceptable or safe when they were $\leq$ 5% (according to B6C) and unacceptable when they were $\geq$ 10%. Rates >5% and <10% were noted as intermediate due to the small number of patients with VNT, any one of whom could represent as much as $\geq$ 5% of a subgroup, despite the size of the parent derivation population. Therefore, we underline that the qualitative grading of missed VNT rates should be interpreted with caution, especially in subgroup analyses. A safe test was called secured when this rate was 0% in CLD with poor liver function (poor CLD hereafter; defined by MELD score $\geq$ 10 [15] corresponding to INR $\geq$ 1.24, see Supplemental Material).

Others - Robustness was the ability of the NIT to maintain satisfactory performance as a function of CLD etiology or severity or sex. *Reliability* evaluated whether certain patient characteristics had a significant impact on test performance. LSM reliability was defined according to Boursier's criteria [16]. *Normal transaminases* were defined by AST and ALT <35 UI/l, which was considered as a surrogate marker of inactive CLD.

Previous tests evaluated

Three published tests were compared: the B6C, the EB6C and the ANTICIPATE score. Details are provided in the Supplemental Material. The ANTICIPATE score, including platelets and LSM, provides a continuous score [17]. In the original paper, a CO for a VNT risk $\leq$ 5% was not provided. Therefore, we calculated it at 0.049 in the present derivation population.

Construction of new NITs

Principles - Obviously, the ideal global NIT must have a high spared EGD rate, but it must also provide a missed VNT rate that both falls under the $\leq$ 5% threshold globally and decreases with CLD severity. Indeed, the risk and the severity of EV bleeding markedly increase as a function of CLD severity [15].

PLER test - Currently, in the B6C, platelets and LSM are used with their own COs, producing a CO pair. However, this latter varies in the literature [3, 7, 8], due to either population characteristics like etiology or simply arbitrary choice since several CO pairs are possible in a same population. The platelets / LSM ratio (PLR) had the strong advantage of circumventing the CO pair choice dilemma, since it confers a single CO for a fixed level of VNT risk (Figure S1). Figure 2 shows that the single PLR CO for a VNT risk <5% manifests as an oblique line corresponding roughly to the extremities of the various CO pairs. The test using PLR with a CO fixed at 10.82 for a missed VNT rate $\leq$ 5% was named *platelet liver elastography ratio* (PLER hereafter). Other PLR characteristics are described in the Supplemental Material.

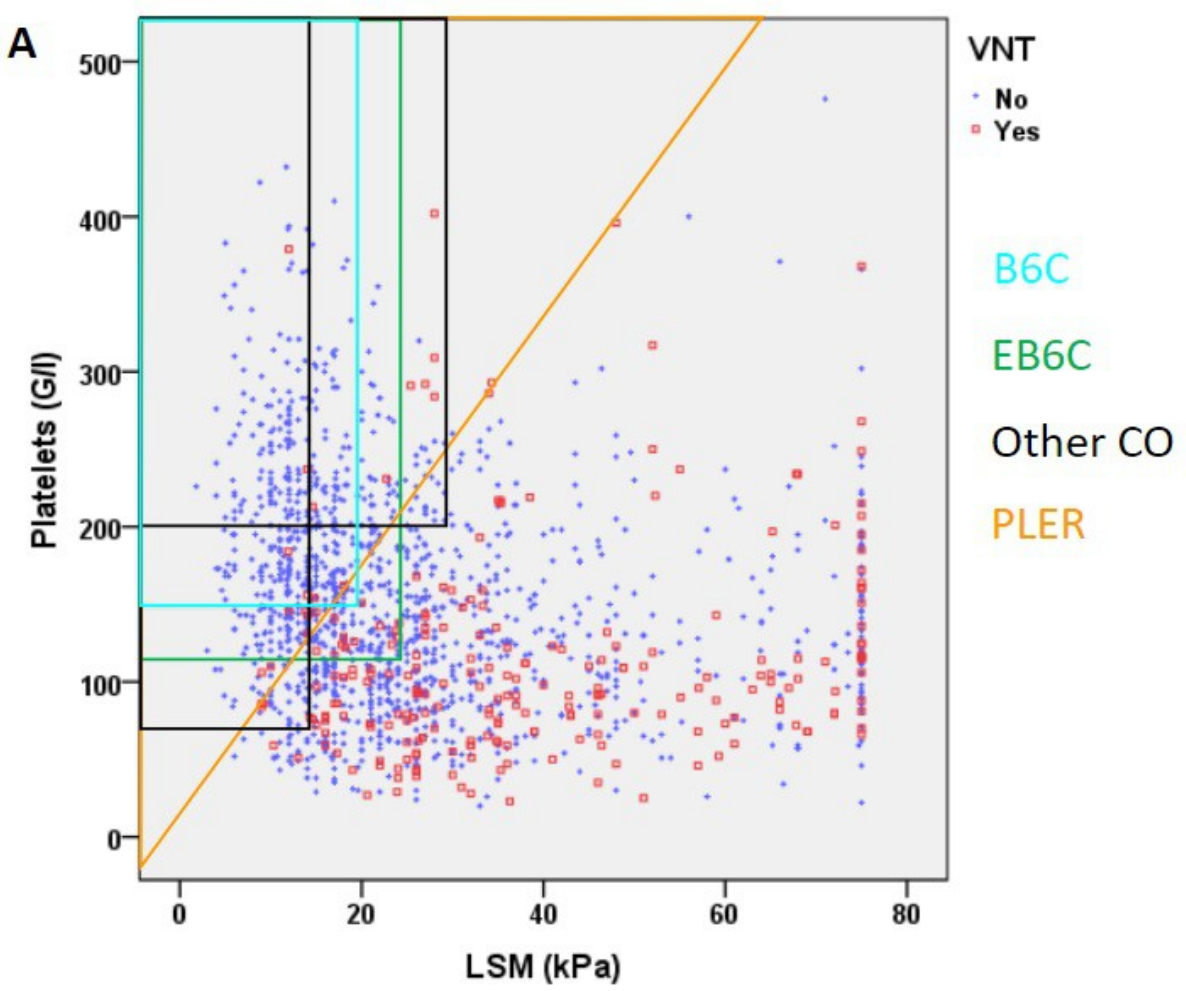
PLESIR test - We observed that sex, etiology and INR or MELD score were independent significant predictors of VNT in addition to PLR.

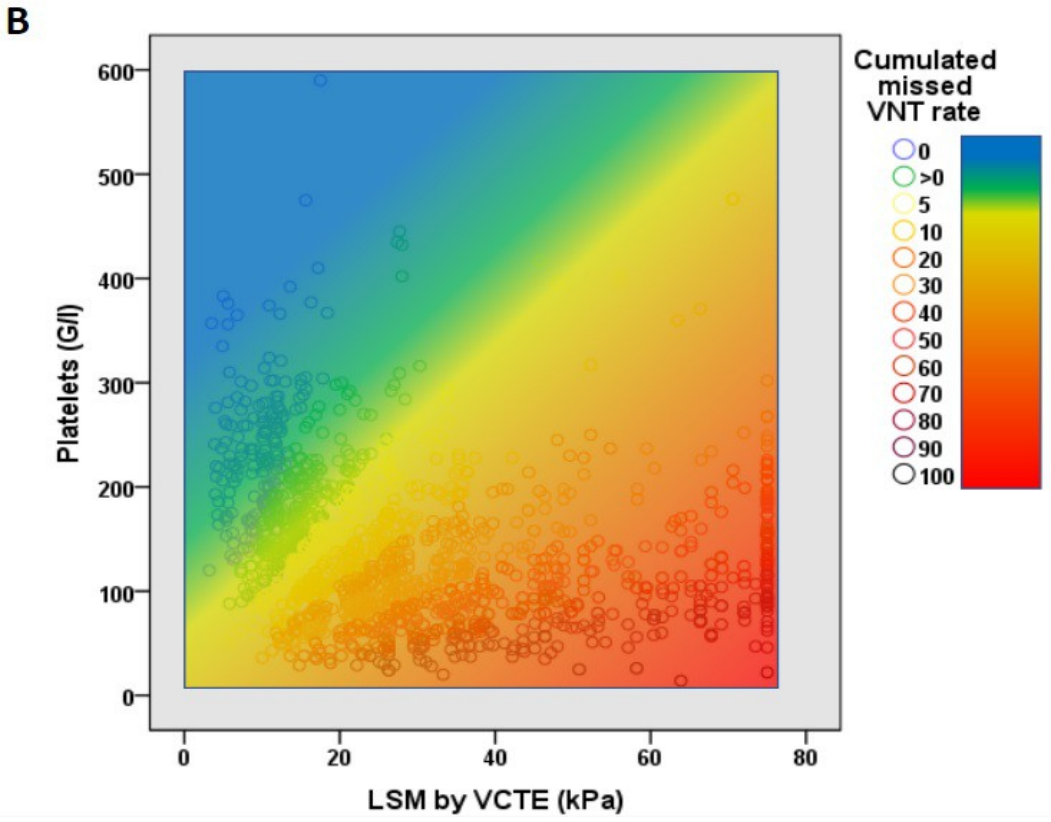
The impact of sex and etiology on VNT prevalence is detailed in Table S2 and Figure S2. In addition, PLR significantly interacted with etiology, sex and INR. So, these confounding factors influenced the PLR CO for missed VNT rate (Table S3).

Therefore, we constructed a test based on the PLR score adjusted on these three variables. However, when the three variables were subjected to analysis of covariance (ANCOVA), the PLR score was only dependent on the INR score (Figure S3). The PLESIR formula is detailed in the Supplemental Material. The binary safety result of PLER or PLESIR is a missed VNT rate $\leq$ 5% or >5%.

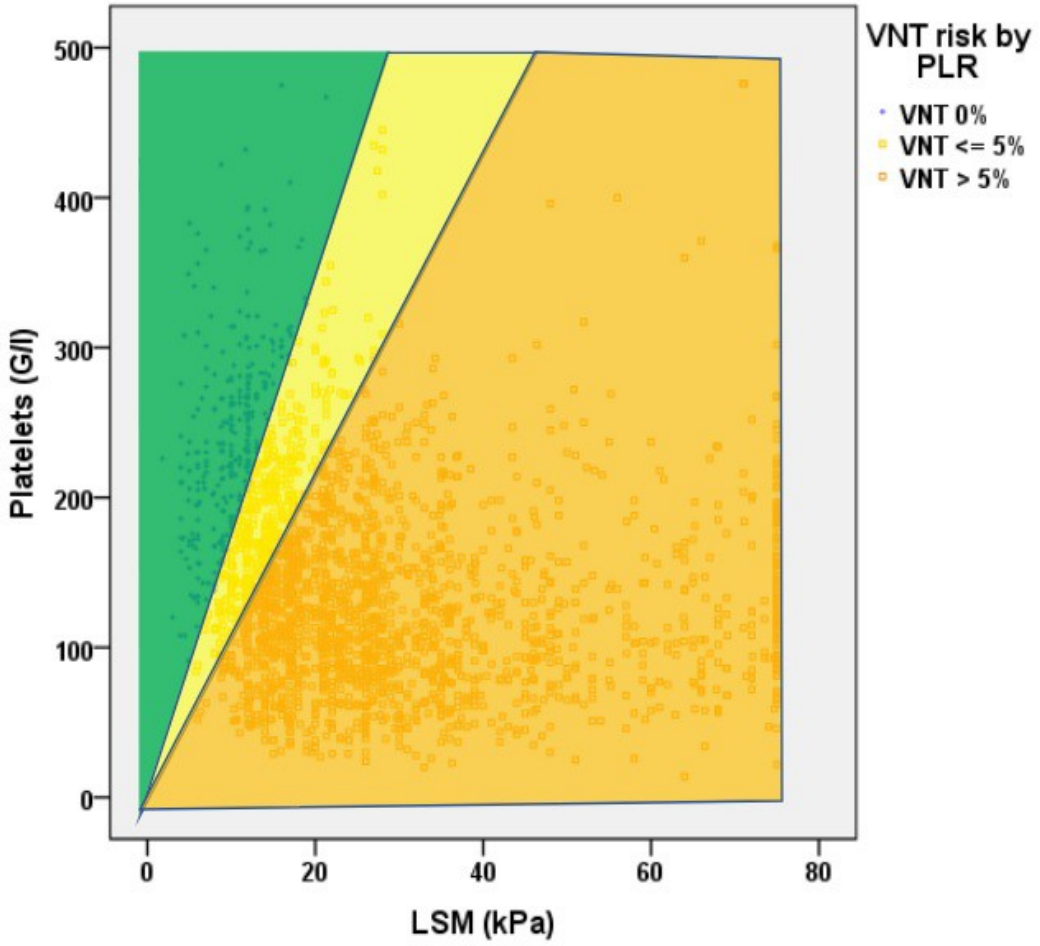
Statistics

Quantitative variables were expressed as mean +/- standard deviation and compared using the Student t test or analysis of variance (ANOVA) or ANCOVA. 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 [18] and Liver FibroSTARD [19] statements, and analyzed on an intention-to-diagnose basis. Thus, all data were included irrespective of reliability criteria of VCTE [16] (except in one NAFLD subpopulation [20]) but missing data were not replaced. Scores including independent predictors were determined by binary logistic regression. The sample size calculation is described in the Supplemental Material. The main statistical analyses were performed using SPSS version 18.0 (IBM, Armonk, NY, USA).


A
VNT
No
Yes
B6C
EB6C
Other CO
PLER
Platelets (G/l)
LSM (kPa)
0
100
200
300
400
500
0
20
40
60
80


B
Cumulated missed VNT rate
0
>0
5
10
20
30
40
50
60
70
80
90
100
Platelets (G/l)
LSM by VCTE (kPa)
0
100
200
300
400
500
600
0
20
40
60
80

C


VNT risk by PLR
VNT 0%
VNT <= 5%
VNT > 5%
Platelets (G/l)
500
400
300
200
100
0
0
20
40
60
80
LSM (kPa)

D

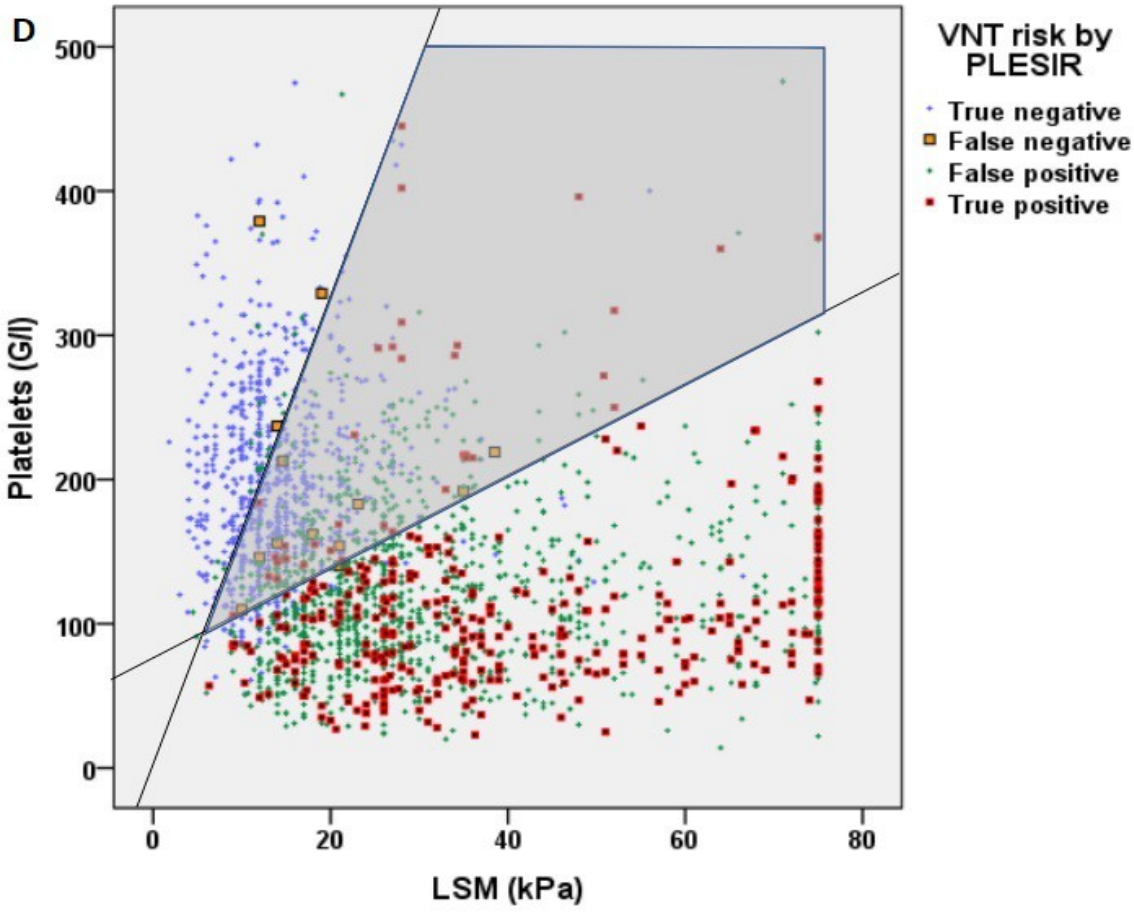

VNT risk by PLESIR
True negative
False negative
False positive
True positive
Platelets (G/l)
500
400
300
200
100
0
0
20
40
60
80
LSM (kPa)

Fig. 2. Scatter plot of platelet (Y axis truncated at 500) and LSM (X axis) values as a function of VNT presence in the derivation population.

(A) PLER test construction, displaying individual cut-offs. The blue and green rectangles correspond to the B6C and EB6C, respectively. The different rectangles correspond to possible cut-off pairs providing zones with a missed VNT rate ≤5%. The orange triangle corresponds to the zone with low VNT risk (missed VNT ≤5%) as calculated with the PLR score CO at 10.8.

(B) Colored zones indicate the PLR score CO for cumulated missed VNT rates: 0%, >0% and ≤5%, >5% and ≤10%, and so on by successive 10% slices.

(C) Graphical nomogram of PLER test for spared endoscopy. The green, yellow and orange triangles indicate respectively 0%, ≤5% and >5% (missed) VNT risks.

(D) VNT (squares) risk as a function of PLESIR test. The grey area indicates the zone where the adjusted PLR score modifies PLR COs compared to a fixed PLR CO like in the PLER test. All negatives correspond to spared endoscopy, false negatives to missed VNT, positives to required endoscopy and false positives to unnecessary endoscopy.

RESULTS

Patient characteristics

Derivation population - This population included 1579 CLD patients (Table 1), nearly two-thirds of whom were men. Viral CLD was the most frequent etiology (50%); other etiologies included NAFLD (29%) or alcoholic CLD (21%). LSMs $\geq$ 10 kPa were observed in 93.0% of patients. VNT prevalence was 15.1%. Severe fibrosis (F3 or F4 Metavir) estimated by LSM [21] was observed in 90% of patients with early cirrhosis in 21% and definitive cirrhosis in 40% (including 70% of VNT). Patients with VNT were more frequently male or of alcoholic etiology and had more severe CLD.

Validation population - This population included 789 CLD patients. No characteristics were significantly different from those of the derivation population (Table 1).

VNT prevalence as a function of sex and etiology - VNT prevalence was always significantly higher in men than in women whatever the etiology (Table S2). However, when the relationship between VNT prevalence and MELD score, sex and etiology was analyzed by ANCOVA (Figure S2), the only significant VNT predictor was MELD score.

Patient characteristics as a function of etiology - These results are detailed in Table S4 in the whole population (due to the similarity of populations). All characteristics were significantly different between etiologies. Briefly, patients with alcoholic CLD were more frequently male, had more severe CLD, higher LSMs and the highest VNT prevalence. Patients with NAFLD had higher BMIs, greater age, the lowest prevalence of VNT, normal transaminases and the lowest AST levels. Patients with viral CLD had the lowest platelet counts despite the lowest MELD scores.

Table 1. Patient characteristics as a function of VNT and populations.

	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	-	50.2	50.2	-
NAFLD	30.1	22.6	-	28.9	28.9	-
Alcoholic	19.6	27.6	-	20.8	20.9	-
VNT (%)	0	100	-	15.1	15.2	*1*
BMI (kg/m²)	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 CO for VNT (kPa)	-	9.0	-	-	10.6 d	-
Metavir F stage by LSM (%):	-	-	*<0.001*	-	-	*0.743*
<F3±1 (≤10.8 kPa)	11.4	2.5	*<0.001*	10.1	10.8	*0.596*
F3±1 (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*

a Student's t test or Chi² test / Fisher test. b Derivation vs validation. c Derivation: no VNT vs VNT. d One outlier value at 6.3 excluded.

LSM: liver stiffness measurement, VNT: varices needing treatment, CO: cut-off

NIT comparison

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 EGD rate (main outcome measurement) 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), PLER: 26.6% (24.3-28.8), PLESIR: 34.8% (32.2-37.2) and EB6C: 41.9% (39.1-44.3).

Validation population - Again, all the NITs had an acceptable missed VNT rate except for the EB6C where it was 11.7% (6.4-17.9) (Table 2) and furthermore significantly higher than the rates of all other NITs. Again, the spared EGD rate was significantly different between tests ($p<0.001$, 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), PLER: 28.8% (25.5-31.8), PLESIR: 35.9% (32.4-39.3) and EB6C: 44.4% (40.8-47.8). Test performance and safety were not significantly different between populations.

Table 2. Missed VNT and spared EGD (% with 95% CI) by tests.

	B6C	EB6C	ANTICIPATE	PLER	PLESIR	*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)	4.6 (2.1-7.4)	3.3 (1.2-5.8)	*<0.001 b*
Spared EGD (%)	23.9 (21.8-26.2)	41.9 (39.1-44.3)	24.3 (21.9-26.4)	26.6 (24.3-28.8)	34.8 (32.2-37.2)	*<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)	1.7 (0.0-4.4)	4.2 (0.9-8.3)	*<0.001 d*
Spared EGD (%)	23.7 (20.7-26.6)	44.4 (40.8-47.8)	25.0 (21.9-27.9)	28.8 (25.5-31.8)	35.9 (32.4-39.3)	*<0.001 e*
Derivation vs validation population (2368 patients) *(p f)*:						
Missed VNT	*0.723*	*0.860*	*0.068*	*0.233*	*0.767*	-
Spared EGD	*0.919*	*0.252*	*0.723*	*0.282*	*0.648*	-

EGD: esophagogastroduodenoscopy, VNT: varices needing treatment

a Paired Cochran test

b Differences in missed VNT rate were significant only between EB6C vs others ($p \leq 0.009$ 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 ($p \leq 0.022$)

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

Color legend:

Missed VNT (% in VNT): ≤ 5: satisfactory, ≥ 10: unsatisfactory, intermediate

Spared EGD (%): ≥ 20: satisfactory, ≤ 10: unsatisfactory, intermediate

Sensitivity analysis

We evaluated the independent predictors of spared EGDs for PLER and PLESIR in a logistic regression including sex, age, etiology, BMI, LSM reliability, normal transaminases, and MELD score as independent variables in the derivation population. For both tests, the independent predictors were MELD score ($p<0.001$), normal transaminases ($p<0.001$) and etiology ($p<0.001$ for PLER and $p=0.014$ for PLESIR). Therefore, we evaluated the impact of these confounding factors on NIT performance in the whole population (Table 3).

Briefly, etiology affected missed VNT rate for B6C and EB6C and spared EGD rate for all tests. Sex affected missed VNT rate for PLER only and spared EGD rate for B6C, EB6C and PLER. The impact of LSM reliability was discordant on spared EGD rate. There was a significant impact of normal transaminases and MELD score on spared endoscopy rate but not on missed VNT rate in all NITs. Thus, spared endoscopy rates were significantly increased in normal transaminases and in lower MELD scores. The impact of the MELD score is detailed as a function of MELD deciles in Figure 3. The spared endoscopy rate decreased progressively and linearly as a function of MELD score in all NITs. The highest rate in the two lowest MELD deciles was observed with PLESIR, reaching 72.7% in the lowest decile. All NITs had a rate <10% in the highest decile. Regarding the missed VNT rate, the profiles were different between NITs with a rather stable rate for EB6C across MELD deciles whereas the rate progressively decreased as a function of MELD score with PLESIR from 37.5% to 0%. The rate of PLESIR might seem high in the lowest decile but, as the VNT prevalence increases progressively as a function of MELD score (Figure 3D), the number of missed VNT remained relatively stable across MELD deciles except for EB6C, this latter showing again an unsafe profile since this number increased as a function of MELD score (Figure 3C).

The impact of etiology as a function of sex (Table S5) or liver severity (Table S6) is reported in the Supplemental Material. Briefly, considering sex by etiology, the only test with a satisfactory missed VNT and spared EGD rates in every clinical setting was PLESIR. Considering liver severity by etiology, importantly, the missed VNT rate decreased as a function of liver severity for all NITs in most etiologies. PLESIR was the only secured test (no missed VNT in INR $\geq$ 1.24, Table S7).

Table 3. Sensitivity analysis in the whole population.

	B6C	EB6C	ANTICIPATE	PLER	PLESIR
Missed VNT rate (%):					
Etiology					
Viral	1.1	10.1	2.2	3.4	3.9
NAFLD	7.4	19.8	7.4	4.9	3.7
ALD	1.0	6.0	2.0	3.0	3.0
p a	***0.006***	***0.012***	*0.069*	*0.761*	*0.922*
Sex					
Male	2.3	10.6	2.3	2.3	3.8
Female	3.2	12.8	6.4	7.4	3.2
p b	*0.702*	*0.570*	*0.088*	***0.046***	*1*
Normal transaminases					
All (n=322)	2.1	10.2	3.3	3.3	3.3
Yes (n=40)	2.5	12.5	2.5	5.0	7.5
No (n=292)	2.1	9.9	3.4	3.1	2.7
p b	*0.596*	*0.581*	*1*	*0.629*	*0.135*
Reliable LSM					
All (n=198)	1.5	7.6	2.5	3.5	3.0
Very reliable (n=67)	0	3.0	0	1.5	0
Reliable (n=106)	2.8	10.4	3.8	4.7	3.8
Poorly reliable (=25)	0	8.0	4.0	4.0	8.0
p a	*0.267*	*0.201*	*0.269*	*0.530*	*0.111*
MELD score					
All (n=278)	1.1	9.0	2.2	2.9	3.2
Tertile 1 (n=30)	3.3	13.3	3.3	3.3	13.3
Tertile 2 (n=76)	1.3	13.2	3.9	5.3	3.9
Tertile 3 (n=172)	0.6	6.4	1.2	1.7	1.2
p a	*0.393*	*0.156*	*0.340*	*0.307*	***0.002***
Spared EGD rate (%):					

Etiology					
Viral	21.6	42.1	25.0	25.9	38.0
NAFLD	33.4	55.8	28.9	35.3	36.9
ALD	16.0	25.9	17.2	19.6	25.9
p a	***<0.001***	***<0.001***	***<0.001***	***<0.001***	***<0.001***
Sex					
Male	22.1	40.4	23.8	*25.7*	35.5
Female	26.9	46.7	25.8	*30.2*	34.6
p b	***0.008***	***0.003***	*0.275*	***0.019***	*0.655*
Normal transaminases					
All (n=2149)	23.7	42.2	24.7	27.4	35.6
Yes (n=322)	29.5	49.4	34.2	37.0	46.6
No (n=1827)	22.7	40.9	23.0	25.7	33.6
p b	***0.010***	***0.005***	***<0.001***	***<0.001***	***<0.001***
Reliable LSM					
All (n=1147)	19.0	35.0	21.7	22.7	31.3
Very reliable (n=306)	14.4	25.5	16.7	19.0	23.5
Reliable (n=729)	21.0	39.1	24.0	24.1	34.6
Poorly reliable (=112)	18.8	34.8	20.5	23.2	31.3
p a	***0.047***	***<0.001***	***0.031***	*0.189*	***0.002***
MELD score					
All (n=1730)	20.8	38.3	23.7	25.0	34.6
Tertile 1 (n=579)	35.9	61.7	40.9	42.0	61.0
Tertile 2 (n=572)	18.0	35.0	21.5	22.9	30.2
Tertile 3 (n=579)	8.3	18.3	8.6	10.0	12.4
p a	***<0.001***	***<0.001***	***<0.001***	***<0.001***	***<0.001***

a Unpaired Chi² test

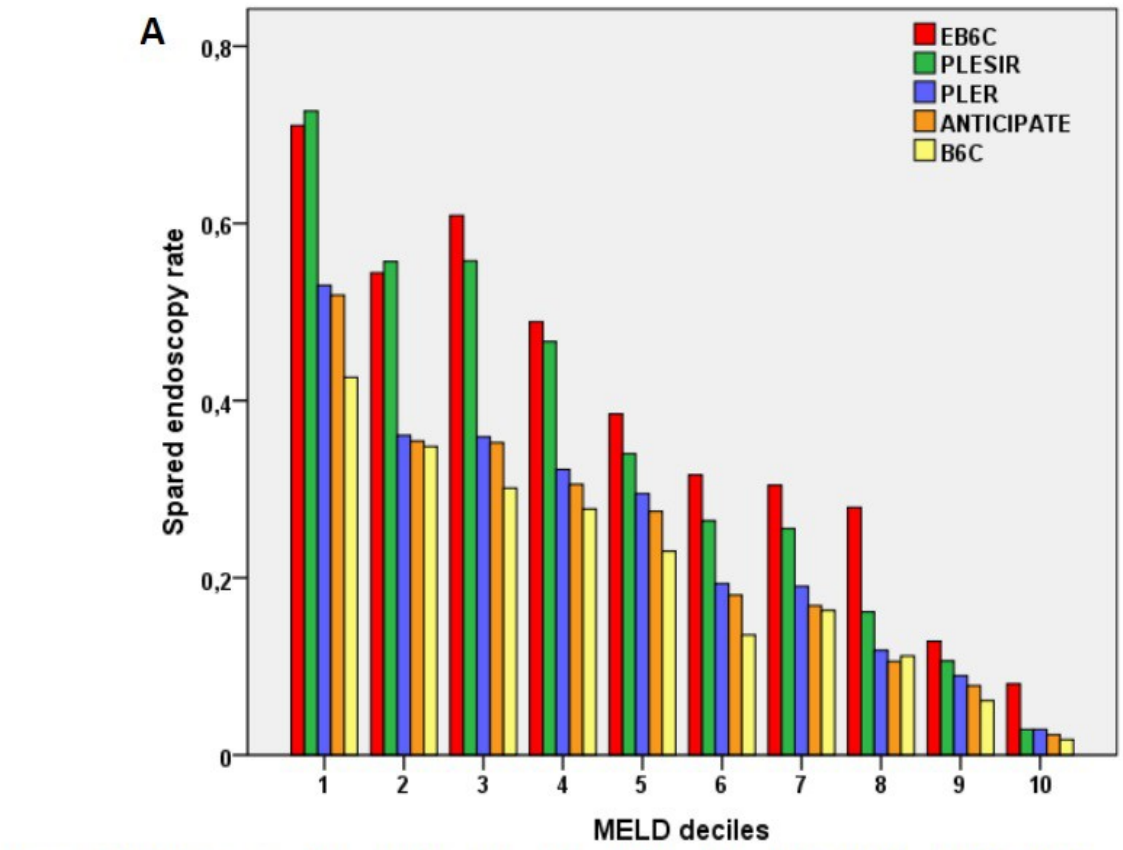
b Unpaired Fisher test

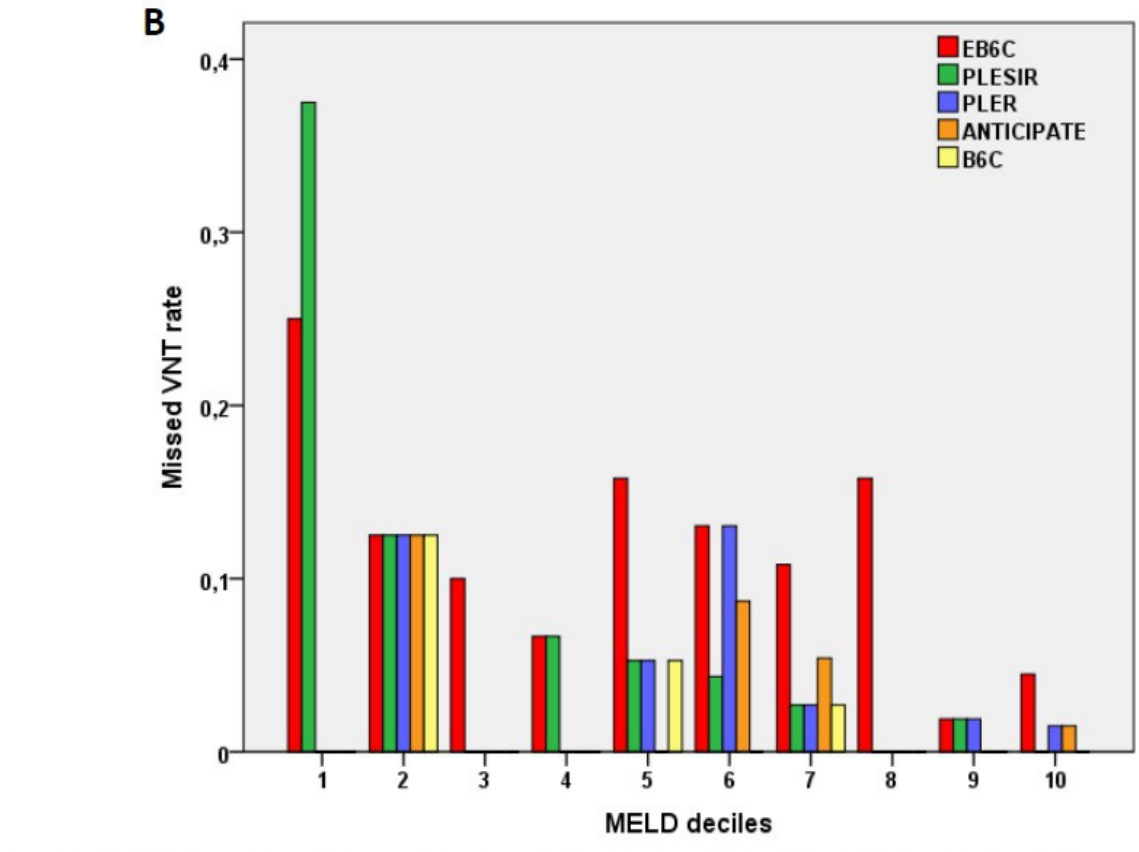
Significant differences in bold characters

Color legend:

Missed VNT (% in VNT): ≤ 5: satisfactory, ≥ 10: unsatisfactory, intermediate

Spared EGD (%): ≥ 20: satisfactory, ≤ 10: unsatisfactory, intermediate


A
EB6C
PLESIR
PLER
ANTICIPATE
B6C
Spared endoscopy rate
0,8
0,6
0,4
0,2
0
1
2
3
4
5
6
7
8
9
10
MELD deciles
Lower MELD score: 6.4 6.8 7.4 7.7 8.2 8.7 9.2 10.0 11.0 13.3


B
EB6C
PLESIR
PLER
ANTICIPATE
B6C
Missed VNT rate
0,4
0,3
0,2
0,1
0
1
2
3
4
5
6
7
8
9
10
MELD deciles
Lower MELD score: 6.4 6.8 7.4 7.7 8.2 8.7 9.2 10.0 11.0 13.3

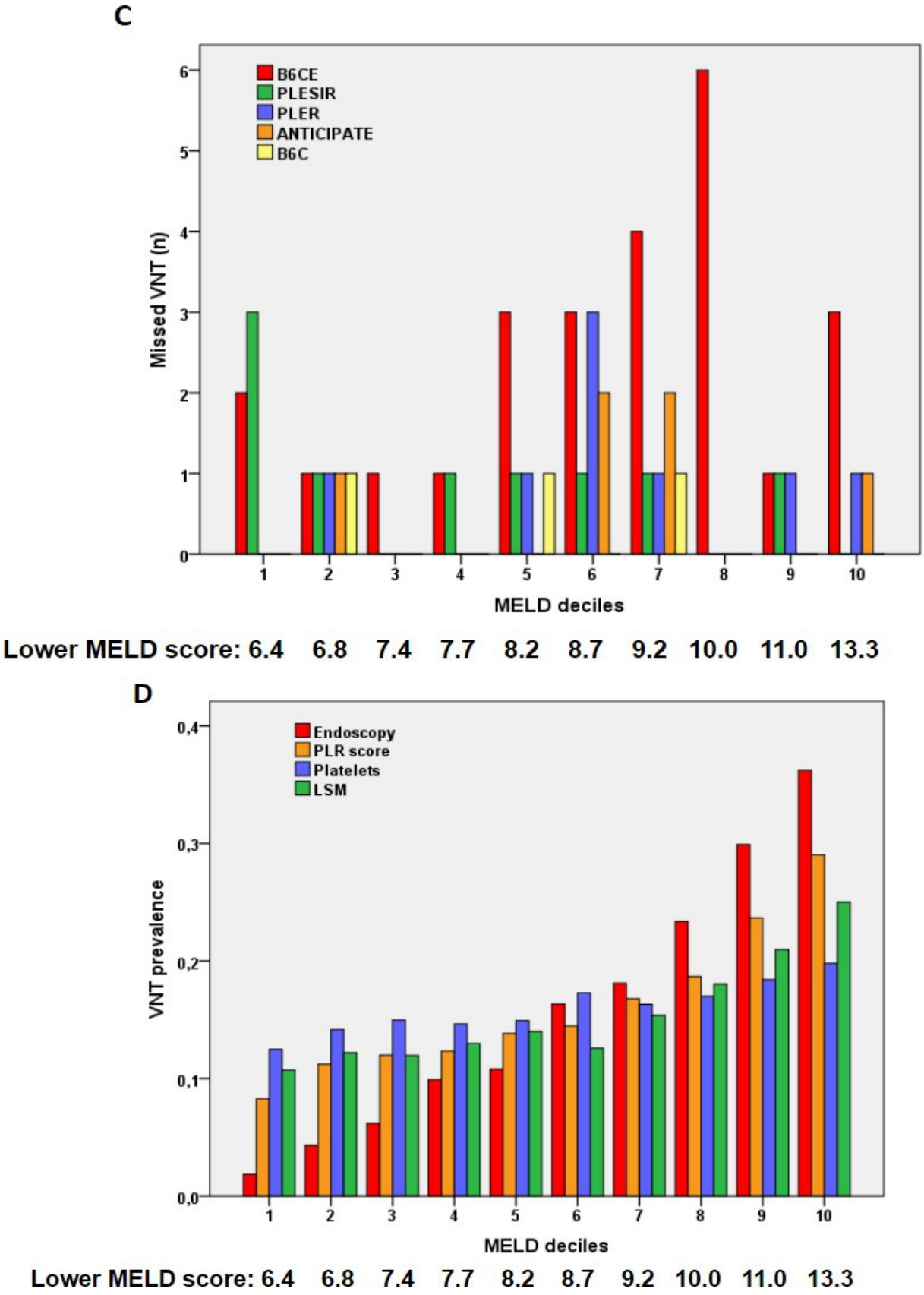


Fig. 3. VNT diagnosis by NITs as a function of MELD deciles.

(A) Spared EGD rate (%). (B) Missed VNT rate (%). (C) Missed VNT number. (D) VNT prevalence according to endoscopy, PLR score (logit), platelets and LSM.

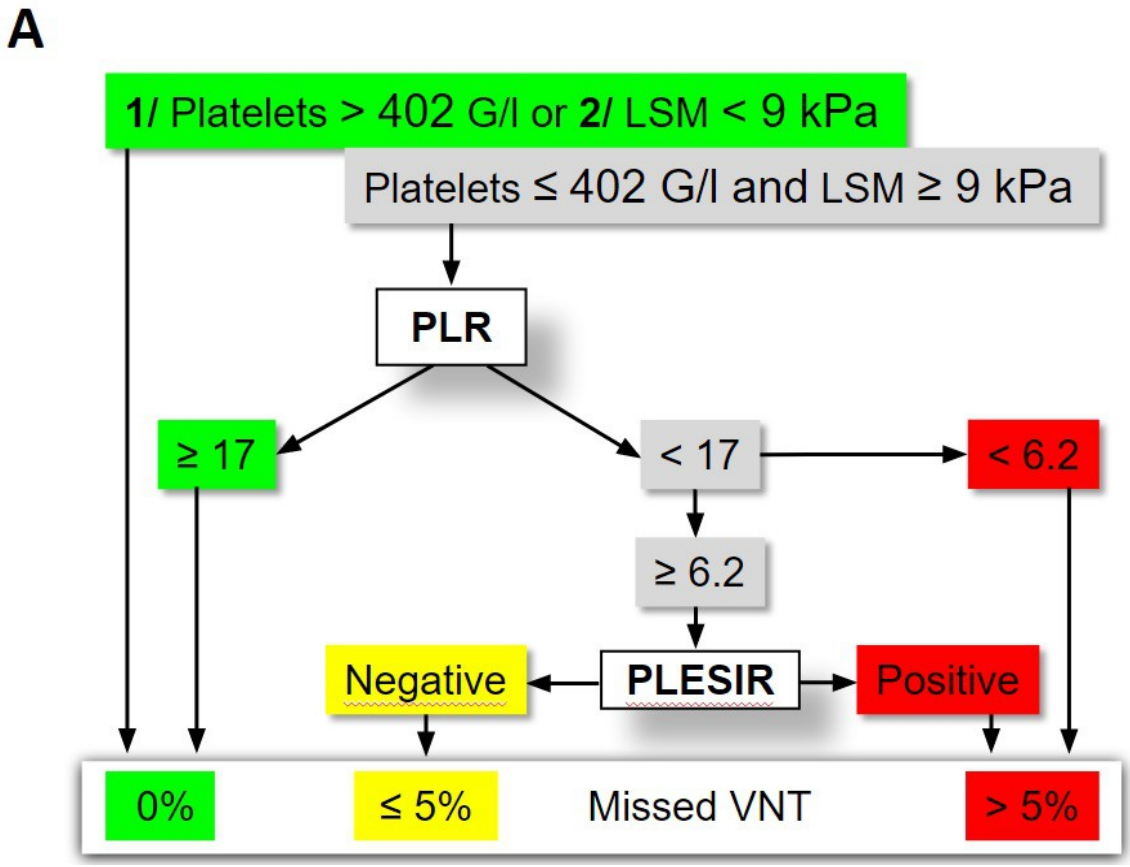
Clinical application

Our results point to two new NITs for clinical application, one requiring only platelets and LSM (PLER) and one requiring one additional blood marker (INR in PLESIR). PLER was clinically simpler than PLESIR. Besides the PLER test, the PLR score provides information on the VNT risk, as shown in Figure S1: it was 0% at high PLR values and remained so down to a PLR CO of 17.0; thereafter, it began increasing to reach 5% at a PLR CO of 10.82. Otherwise, we can define a simple clinical rule for PLR score: when platelets (expressed in usual G/l units) are $\geq$10 times (in viral or alcoholic CLD) or $\geq$15 times (in NAFLD) the LSM value (expressed in kPa with VCTE), VNT risk is below ≈5% and EGD can be avoided.

Moreover, we can propose a simple clinical algorithm for VNT screening (Figure 4A). First, it is necessary to calculate PLR only in patients at risk of VNT, i.e. those having platelets $\leq$402 (i.e.≈400) G/l and even LSM $\geq$ 9 kPa. This first step eliminated the need for PLR calculation in 5.6% of the derivation population (Figure 4B). Second, when PLR is $\geq$17, the VNT risk is 0% whatever the clinical setting (sex or etiologies evaluated) and endoscopy is unnecessary; thus 7.3% of the derivation population would have been spared PLESIR calculation. By contrast, when PLR is <6.2, the VNT risk is >5% in all clinical settings and endoscopy is mandatory; this would have removed 51.6% of the derivation population from PLESIR calculation. Between these two COs, the grey zone of the PLR score requires PLESIR calculation; 35.4% of the derivation population would have been concerned. This sequential algorithm, called *VariScreen*, performs the same as the PLESIR approach alone (Table 4). However, VariScreen is clinically more relevant.

First, it avoids calculations, e.g., PLR would have been avoided in 5.6% (5.5-6.7%) and PLESIR in 64.5% (62.1-67.0%) of patients in the derivation population.

Second, VariScreen was safer than PLESIR since the missed VNT rate was 0% in MELD score ≥ 10 (or INR ≥ 1.24) in both populations. Finally, a reliability analysis provided a more reliable version of VariScreen where the spared EGD rate was increased slightly from 35.1% to 36.1% (p<0.001) in the whole population while not affecting safety (Tables S10/11). VNT risks by the new NITs are provided by a calculator (http://forge.info.univ-angers.fr/~gh/wstat/pler-plesir-variscreen.php).



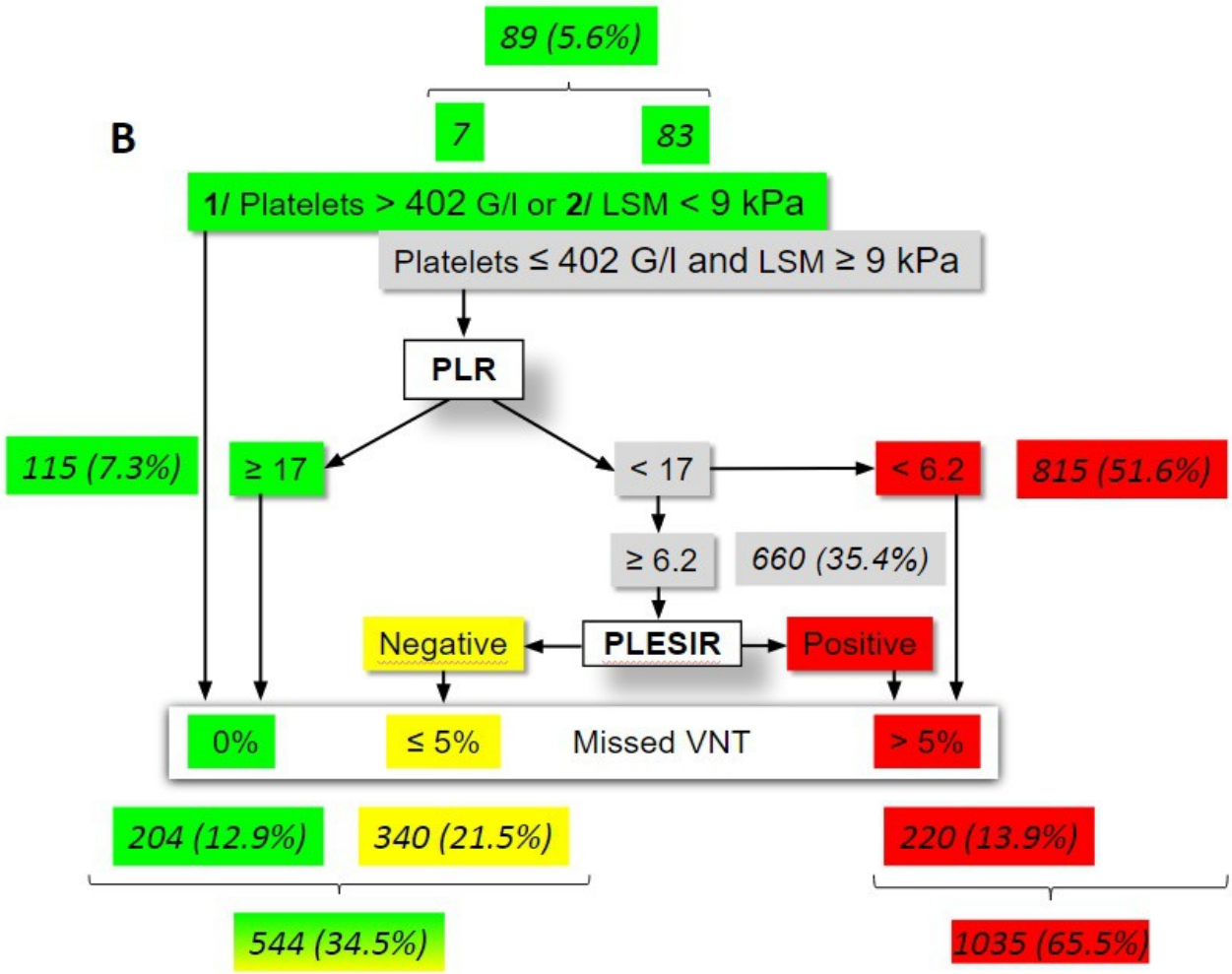


Figure 4. VariSreen algorithm: VNT screening strategy in CLD, first as a function of biomarkers (platelets, LSM) and then, optionally, of ensuing PLR score then PLESIR test.

Panel A: algorithm only.

Panel B: algorithm with patient figures (in italics) from derivation population.

Table 4. Characteristics of VariScreen algorithm with comparison to PLESIR test.

	Whole population	Derivation	Validation	*p a*
Patients (n)	2368	1579	789	
VNT risk as a function of PLR (%):				
All PLR values	15.2	15.1	15.2	*1*
PLR ≥17	0.7	0.5 b	1.0	*1*
6.2 ≤ PLR <17 (grey zone)	5.8	6.2	5.0	*0.544*
PLR <6.2	25.3	24.8	26.5	*0.526*
Spared PLR score (%):	5.9	5.6	6.5	*0.460*
Spared PLESIR (%):	64.3	64.5	63.9	*0.785*
Spared EGD rate (%):				
PLESIR	35.2	34.8	35.9	*0.648*
VariScreen	35.1	34.5	36.2	*0.411*
p c	*0.791*	*0.430*	*0.629*	-
Missed VNT rate (%):				
In all patients:				
PLESIR	3.6	3.3	4.2	*0.767*
VariScreen	3.6	2.9	5.0	*0.373*
p c	*0.500*	*1*	*1*	-
In MELD score ≥10 d:				
PLESIR	0.6	1.0	0	*1*
VariScreen	0	0	0	*NA*
p c	*1*	*1*	*1*	-
In MELD score ≥12 (%) d:				
PLESIR	1.1	1.6	0	*1*
VariScreen	0	0	0	*NA*
p c	*1*	*1*	*1*	-

NA: not available
a Unpaired Fisher test
b 1 patient considered as outlier
c Paired McNemar test
d Secured safety criterion
[e] One patient with viral CLD, MELD=12.5 (bilirubin =80, creatinine=62, INR=1.0), albumin=35 and AST=122

DISCUSSION AND CONCLUSION

Main results

Among the published NITs, the original B6C offer extensive validation, a spared endoscopy rate of 20-23% [5] and a low missed VNT rate. Here, we too found the B6C to be robust and safely applicable regardless of etiology and liver severity. However, the missed VNT rate was borderline in NAFLD and the spared VNT rate was only 16% in alcoholic CLD. In contrast, despite their high spared EGD rate, we disqualified the EB6C because of an unacceptably 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 that retrospectively calculated that aspect based on sensitivity, the statistical standard in test construction, and equally reported excessively high rates for it, ranging from 6.5% to 33.3% [14, 22].

In this work, we also validated the ANTICIPATE score for the first time, showing that it offers good performance and robustness, apart the same limits mentioned above for the B6C (borderline missed VNT rate in NAFLD and low spared VNT rate in alcoholic CLD). Moreover, we provided a CO delimiting a VNT risk $\leq$ 5% for it. ANTICIPATE differs from other published tests in that it is a continuous score providing the precise VNT risk in a given patient. However, ANTICIPATE does require a calculator, and although it performed well, it did fall short of the new tests.

Among the new NITs presented here, the simple PLER test performed better than the B6C. Compared to this latter and the EB6C, PLER also has the advantage of employing a CO that is not only unique but also objective in that the VNT risk $\leq$ 5% is automatically calculated; the CO depends only on population characteristics.

The only putative disadvantage for PLER is that clinicians may find the figures it provides ineloquent, but this can be solved by a very simple clinical rule (see below).

Finally, PLESIR was the best-performing test and furthermore robust and safe; so, PLESIR cumulated advantages. Our validation and derivation 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 wherein the robustness of NITs could be calculated according to the three main influencing factors: sex, CLD etiology and severity. The relationship between VNT and sex or etiology has been poorly evaluated. On that issue, previous studies have provided negative results but with low sample sizes [9]. Furthermore, we propose an original test, called PLER, that simply uses the ratio of platelets to LSM.

With our large population, we showed that sex, CLD etiology and severity were confounders of NIT performance. We found that the influences of etiology and sex were only moderate but that of severity was marked, probably because all the involved factors vary with it. Thus, VNT prevalence and biomarkers (platelets and LSM) are strongly related to liver severity [23]. This results in the reduction of the performance of NITs in severe CLD. The Baveno VI statements suggest applying B6C in cACLD. In contrast to previous studies, we were more able to examine the applicability of NITs according to CLD severity. For the B6C, we found an unsatisfactory spared EGD rate (8%) in severe CLD (third MELD tertile), confirming the Baveno VI statement that these criteria underperform in severe CLD.

Our results do suggest however that new NITs may be able to partially mitigate that problem. Indeed, PLESIR is the PLR adjusted on sex, etiology and even INR. This adjustment provides the important advantages of increased performance and safety in every clinical setting.

For example, PLESIR showed an acceptable or safe missed VNT rate in both sexes, in contrast to the B6C in NAFLD and ANTICIPATE or PLER in women. In addition, PLESIR provided safety (0.6% missed VNT when the MELD score was $\geq$ 10), robustness and better performance in severe CLD (12% spared EGD in third MELD tertile).

The present study is the first to validate the B6C in alcoholic CLD. Our results suggest that several NITs can be used in this setting albeit with a significantly lower spared EGD rate (<20% except for PLESIR at 26%). Concerning NAFLD, the spared EGD rate was very satisfactory but the missed VNT rate was acceptable only for PLER and PLESIR.

Limits

Our study comprised numerous centers, which necessarily induced variability. However, the performances of published tests observed here were very similar to those of previous studies [1, 4, 5], which validates our data record. Our study was retrospective in design but data were prospectively recorded and VNT was the main outcome as concerned data collection (except in the prospective CIRVIR cohort and NAFLD-B6C study). According to the Baveno VI statements, VNT are defined as large EV and grade 1 EV with red signs [3]. However, we did not include EV with red signs for several reasons. First, many data had been collected before the Baveno VI statements were published, and thus information on red signs was missing in a substantial proportion of them. More importantly, in an earlier fully prospective subpopulation, we observed significant inter-center variability in the prevalence of grade 1 EV with red signs [24], which statistically affected outcome measurements. Those results cast doubts on the real prevalence of this EV category, which may be affected by large interobserver variability [25]. So, for the present work, we did not consider grade 1 EV with red signs, even in a subgroup analysis. This issue merits a fully prospective study.

We also underline that our results on test performance align quite well with those of previous studies using different VNT definitions. We decided to analyze the impact of liver severity using the MELD score (or INR) and not clinical complications (the upper limit of the B6C). 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 CO determination. Additionally, Child-Pugh, including complications, and MELD scores have shown similar prognostic values in most cases [26]. 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 [6]. Moreover, normal transaminases were considered as a surrogate marker of inactive CLD in the present study; as expected, spared endoscopy rates were significantly higher in normal transaminases, but this affected the overall rate only slightly due to their low prevalence (15%). Finally, we had enough patients to evaluate the new NITs in NAFLD (29%) and alcoholic CLD (21%). An evaluation of the published NITs in a large NAFLD series suggested that their COs should be adapted to that condition [20]. The present study shows that COs have to be adapted to each etiology and sex. The majority of NAFLD patients in the present work were provided by a recently published study [20] that, concerning the performance of NITs, reported more optimistic results than ours. That difference can be traced to the employed definition of missed VNT, as discussed earlier and detailed elsewhere [14], and furthermore illustrates why the definition of missed VNT should be harmonized to allow true comparisons between studies. We recently presented arguments for our preferred definition (missed VNT/all VNT) [14]. The main limit of the privileged definition for the present study is that it is the most sensitive to the number of patients with missed VNT, which limits subgroup analysis interpretation.

We also take this opportunity to underline that the performance of the VCTE XL probe needs further study.

The most important limit to consider is clinical application. In our study, PLESIR had high missed VNT rates in patients with MELD scores <10, but that finding merits cautious interpretation for three reasons. First, it was derived from a subgroup analysis with a low VNT number [27], whereas the target was a missed VNT rate $\leq 5\%$ in the whole population. Consequently, the crude number of missed VNT was the highest in the second MELD tertile (Figure 3C). Second, the missed VNT rate was satisfactory in the third MELD tertile where there is a far higher VNT prevalence. Third, we defined an ideal NIT offering a decreasing missed VNT rate as a function of CLD severity, since bleeding risk and severity increase with that latter. The test adjusted on INR provided the best profile as expected. Thus, PLESIR was safe (missed VNT $\leq 1\%$) -and the VariScreen algorithm even more so (no missed VNT)- in severe CLD (MELD score ≥ 10) regardless of etiology. This implies that the missed VNT rate can be >5% in less severe CLD, since the target ($\leq$5%) is the mean of all subgroups. However, we consider that a secured strategy, i.e., one without missed VNT in poor CLD, can be applied despite a low spared EGD rate for three reasons. Firstly, it is safe as concerns risky VNT (because no VNT are missed). Secondly, it also increases safety in clinical use since there is no MELD limit to inadvertently neglect (a possibility with unsafe tests). And finally, all NITs estimate VNT risk >5% (and more precisely for PLR score), which can serve as a more convincing argument for the EGD indication in a reluctant patient knowing that EV endoscopic screening is applied in only 1/3 of patients [28].

Clinical application

The new NITs offering satisfactory performance can be ranked as follows (easiest to hardest): PLER, ANTICIPATE (considered as new due to CO determination in the present study), and PLESIR. PLER can be used in several ways, none being particularly complex. The first, which requires no materials, should prime: when the platelet count (G/l) is $\geq$ 10-fold the LSM value (kPa with VCTE) in viral and alcoholic CLD, EGD is not necessary. For example, a patient with platelets at 250 G/l and LSM at 25 kPa would not need EGD. The CO in NAFLD is 15. PLER can also be used with a graphical nomogram (Figure 2C). The PLESIR test offers better performance and safety but requires a calculator. In reality, we recommend the VariScreen algorithm, which did not require a calculator in 65% of patients and 35% of endoscopies in the present population.

Moreover, as the first steps and 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.

The ANTICIPATE test offers little or no advantage over PLER, and, if a calculator is to be used, it would be more logical to turn to a NIT offering markedly higher performance. 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.

Conclusion

In the present study, we show that the Baveno VI criteria are robust whatever the etiology or severity of CLD (with a safety limit in NAFLD), but that the expanded Baveno VI criteria are unsafe due to a high missed VNT rate. We propose two possibilities to improve the valuable Baveno VI criteria by using platelets and LSM either alone or adjusted on INR, etiology and sex. We present an easy-to-use score (PLR), deployable everywhere with very simple rules (platelets ≥ 10-fold LSM = no endoscopy in viral and alcoholic CLD) and a better-performing personalized test (PLESIR) that requires a calculator. We recommend the sequential VariScreen algorithm (successive optional steps: platelet count => LSM => PLR score => PLESIR => endoscopy) for its applicability in all main-etiology CLDs and especially its secureness in poor CLD.

Acknowledgments

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-Carrië, 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-Carrië, Victor de Ledinghen, Federico Ravaioli, Antonio Colecchia, Corina Rusu, Pietro Andreone, Giuseppe Mazzella, Davide Festi.
- B6C NAFLD group : Salvatore Petta[1], Giada Sebastiani[2], Elisabetta Bugianesi[3], Mauro Vigano[4], Vincent Wai-Sun Wong[5], Annalisa Berzigotti[6], Anna Ludovica Fracanzani[7], Quentin M. Anstee[8,9], Fabio Marra[10,11], Marco Barbara[1], Vincenza Calvaruso[1], Calogero Camma[1],

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

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

  [3]Division of Gastroenterology, Department of Medical Sciences, University of Torino,

  Torino, Italy;

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

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

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

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

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

[10]Dipartimento di Medicina Sperimentale e Clinica, University of Florence, Italy; [11]Research Center DENOTHE, University of Florence, Italy;

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

- ANRS CO12 CirVir group: Pierre Nahon[1], Patrick Marcellin[2], Dominique Guyader[3], Stanislas Pol[4], Hélène Fontaine[4], Dominique Larrey[5], Victor De Lédinghen[6], Denis Ouzan[7], Fabien Zoulim[8], Dominique Roulot[9], Albert Tran[10], Jean-Pierre Bronowicki[11], Jean-Pierre Zarski[12], Vincent Leroy[12], Ghassan Riachi[13], Paul Calès[14], Jean-Marie Péron[15], Laurent Alric[16], Marc Bourlière[17], Philippe Mathurin[18], Sebastien Dharancy[18], Jean-Frédéric Blanc[19], Armand Abergel[20], Lawrence Serfaty[21], Ariane Mallat[22], Jean-Didier Grangé[23], Pierre Attali[24], Yannick Bacq[25], Claire Wartelle[26], Thông Dao[27], Yves Benhamou[28], Christophe Pilette[29], Christine Silvain[30], Christos Christidis[31], Dominique Capron[32], Gérard Thiefin[33], Sophie Hillaire[34], Vincent Di Martino[35].

  [1]AP-HP, Hôpital Jean Verdier, Service d'Hépatologie, Bondy, Université Paris 13, Bobigny et INSERM U1162, Université Paris 5, Paris; [2]AP-HP, Hôpital Beaujon, Service d'Hépatologie, Clichy; [3]CHU Pontchaillou, Service d'Hépatologie, Rennes; [4]AP-HP, Hôpital Cochin, Département d'Hépatologie et INSERM UMS20 et U1223, Institut Pasteur, Université Paris Descartes, Paris; [5]Hôpital Saint Eloi, Service d'Hépatologie, Montpellier; [6]Hôpital Haut-Lévêque, Service d'Hépatologie, Bordeaux; [7]Institut Arnaud Tzanck, Service d'Hépatologie, St Laurent du Var; [8]Hôpital Hôtel Dieu, Service d'Hépatologie, Lyon; [9]AP-HP, Hôpital Avicenne, Service d'Hépatologie, Bobigny; [10]CHU de Nice, Service d'Hépatologie,

et INSERM U1065, Université de Nice-Sophia-Antipolis, Nice; [11]Hôpital Brabois, Service d'Hépatologie, Vandoeuvre-les-Nancy; [12]Hôpital Michallon, Service d'Hépatologie, Grenoble; [13]Hôpital Charles-Nicolle, Service d'Hépatologie, Rouen; [14]CHU d'Angers, Service d'Hépatologie, Angers; [15]Hôpital Purpan, Service d'Hépatologie, Toulouse; [16]CHU Toulouse, Service de Médecine Interne-Pôle Digestif UMR 152, Toulouse; [17]Hôpital Saint Joseph, Service d'Hépatologie, Marseille; [18]Hôpital Claude Huriez, Service d'Hépatologie, Lille; [19]Hôpital St André, Service d'Hépatologie, Bordeaux; [20]Hôpital Hôtel Dieu, Service d'Hépatologie, Clermont-Ferrand; [21]AP-HP, Hôpital Saint-Antoine, Service d'Hépatologie, Paris; [22]AP-HP, Hôpital Henri Mondor, Service d'Hépatologie, Créteil; [23]AP-HP, Hôpital Tenon, Service d'Hépatologie, Paris; [24]AP-HP, Hôpital Paul Brousse, Service d'Hépatologie, Villejuif; [25]Hôpital Trousseau, Unité d'Hépatologie, CHRU de Tours; [26]Hôpital d'Aix-En-Provence, Service d'Hépatologie, Aix-En-Provence; [27]Hôpital de la Côte de Nacre, Service d'Hépatologie, Caen; [28]AP-HP, Groupe Hospitalier de La Pitié-Salpêtrière, Service d'Hépatologie, Paris; [29]CHU Le Mans, Service d'Hépatologie, Le Mans; [30]CHU de Poitiers, Service d'Hépatologie, Poitiers; [31]Institut Mutualiste Montsouris, Service d'Hépatologie, Paris; [32]Hôpital Amiens Nord, Service d'Hépatologie, Amiens; [33]Hôpital Robert Debré, Service d'Hépatologie, Reims; [34]Hôpital Foch, Service d'Hépatologie, Suresnes; [35]Hôpital 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 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.

REFERENCE

[1] 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.
[2] Ravaioli F, Montagnani M, Lisotti A, Festi D, Mazzella G, Azzaroli F. Noninvasive Assessment of Portal Hypertension in Advanced Chronic Liver Disease: An Update. Gastroenterol Res Pract 2018;2018:4202091.
[3] 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;63:743-752.
[4] Marot A, Trepo 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;37:707-716.
[5] Roccarina D, Rosselli M, Genesca J, Tsochatzis EA. Elastography methods for the non-invasive assessment of portal hypertension. Expert Rev Gastroenterol Hepatol 2018;12:155-164.
[6] Thabut D, Bureau C, Layese R, Bourcier V, Hammouche M, Cagnot C, 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.e1005.
[7] Ding NS, Nguyen T, Iser DM, Hong T, Flanagan E, Wong A, et al. Liver stiffness plus platelet count can be used to exclude high-risk oesophageal varices. Liver Int 2016;36:240-245.
[8] 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;66:1980-1988.
[9] 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;37:1177-1183.
[10] Colecchia A, Ravaioli F, Marasco G, Colli A, Dajti E, Di Biase AR, et al. A combined model based on spleen stiffness measurement and Baveno VI criteria to rule out high-risk varices in advanced chronic liver disease. J Hepatol 2018;69:308-317.
[11] Kamath PS, Kim WR, Advanced Liver Disease Study G. The model for end-stage liver disease (MELD). Hepatology 2007;45:797-805.
[12] Boursier J, Konate A, Gorea G, Reaud S, Quemener E, Oberti F, et al. Reproducibility of liver stiffness measurement by ultrasonographic elastometry. Clin Gastroenterol Hepatol 2008;6:1263-1269.
[13] Cales P, Oberti F, Bernard-Chabert B, Payen JL. Evaluation of Baveno recommendations for grading esophageal varices. J Hepatol 2003;39:657-659.
[14] Cales P, Buisson F, Ravaioli F, Berger A, Carboni C, Marasco G, 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.
[15] Singal AK, Kamath PS. Model for End-stage Liver Disease. J Clin Exp Hepatol 2013;3:50-60.
[16] Boursier J, Zarski JP, de Ledinghen V, Rousselet MC, Sturm N, Lebail B, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. Hepatology 2013;57:1182-1191.
[17] 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;64:2173-2184.

[18] Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;351:h5527.
[19] Boursier J, de Ledinghen V, Poynard T, Guechot J, Carrat F, Leroy V, et al. An extension of STARD statements for reporting diagnostic accuracy studies on liver fibrosis tests: the Liver-FibroSTARD standards. J Hepatol 2015;62:807-815.
[20] Petta S, Sebastiani G, Bugianesi E, Vigano M, Wong VW, Berzigotti A, et al. Noninvasive Prediction of Esophageal Varices by Stiffness and Platelet in Nonalcoholic Fatty Liver Disease Cirrhosis. J Hepatol 2018;69:878-885.
[21] Cales P, Boursier J, Oberti F, Bardou D, Zarski JP, de Ledinghen V. Cirrhosis diagnosis and liver fibrosis staging: transient elastometry versus cirrhosis blood test. J Clin Gastroenterol 2015;49:512-519.
[22] Dajti E, Ravaioli F, Colecchia A, Marasco G, Calès P, Festi D. "Are the Expanded Baveno VI Criteria really safe to screen compensated cirrhotic patients for high-risk varices?". Dig Liver Dis 2019;51:456-457.
[23] Latorre R, Vaquero J, Rincon D, Puerto M, Ponce MD, Sarnago F, et al. Determinants of platelet count are different in patients with compensated and decompensated cirrhosis. Liver Int 2016;36:232-239.
[24] Bastard C, Miette V, Cales P, Stefanescu H, Festi D, Sandrin L. A Novel FibroScan Examination Dedicated to Spleen Stiffness Measurement. Ultrasound Med Biol 2018;44:1616-1626.
[25] Cales P, Buscail L, Bretagne JF, Champigneulle B, Bourbon P, Duclos B, et al. [Interobserver and intercenter agreement of gastro-esophageal endoscopic signs in cirrhosis. Results of a prospective multicenter study]. Gastroenterol Clin Biol 1989;13:967-973.
[26] Peng Y, Qi X, Guo X. Child-Pugh Versus MELD Score for the Assessment of Prognosis in Liver Cirrhosis: A Systematic Review and Meta-Analysis of Observational Studies. Medicine (Baltimore) 2016;95:e2877.
[27] Dujardin B, Van den Ende J, Van Gompel A, Unger JP, Van der Stuyft P. Likelihood ratios: a real improvement for clinical decision making? Eur J Epidemiol 1994;10:29-36.
[28] Flemming JA, Saxena V, Shen H, Terrault NA, Rongey C. Facility- and Patient-Level Factors Associated with Esophageal Variceal Screening in the USA. Dig Dis Sci 2016;61:62-69.

Supplemental Materials
The Supplemental Material includes 11 tables and 7 figures.

LIST OF FIGURES

TABLE LIST

TABLE OF CONTENTS

ANNEXES

Supplemental material

Personalized platelets/liver stiffness ratio improves and secures the screening of esophageal varices needing treatment

Arthur Berger,* 1 **Federico Ravaioli,*** 2 Oana Farcau, 3 Davide Festi,2 Horia Stefanescu, 3 François Buisson, 1 Pierre Nahon, 4 Christophe Bureau, 5 Nathalie Ganne-Carrià, 4 Annalisa Berzigotti, 6 Victor de Ledinghen, 7 Salvatore Petta, 8 Paul Calès 1 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.

Contents

Introduction

Tables S1 to S7 and Figures S1 to S3 are supplementary to data introduced in the main text. Tables S8 to S11 and Figures S4 to S7 present additional data as described below.

Subpopulations

Ten subpopulations were included in the whole population. Seven subpopulations had been previously published within the setting of the evaluation of different diagnostic tools for VNT. Video capsule endoscopy was initially evaluated in the prospective multicenter VCO-VO study (NCT00941421) performed in Nantes, Angers, Toulouse, Clichy, Lomme, Bobigny, Paris Cochin and Bordeaux (France) [1, 2] . Spleen stiffness measurement by a conventional VCTE device was initially evaluated in Colecchia's retrospective study with derivation and validation populations recruited respectively in Bologna (Italy) and Bern (Switzerland) [3] . Spleen stiffness measurement by a new specific device was initially evaluated in the prospective multicenter M116 study (NCT02180113) performed in Cluj (Romania), Bologna, Verona and Milan (Italy), Angers, Bordeaux and Bondy (France), and London (UK) [4] . One part of the CIRVIR cohort [5] was included, i.e., that involved in an ancillary study on fibrosis tests conducted by the Grenoble (France) center. The CIRVIR cohort included histologically-proven Child-Pugh A cirrhosis due to HCV or HBV. The remaining part of the CIRVIR cohort [5] was also included later on. Most patients with NAFLD were provided by a recently published multicenter study wherein unreliable LSM were excluded [6] ; we also excluded patients having only LSM done with the XL probe. Any patient who had already been included in a previous study involved in the development of a test compared here could not be included in a subpopulation of the present derivation population. This aspect indeed had to be monitored, because two centers, Cluj and Toulouse, had participated in studies for the ANTICIPATE score [7] and/or the EB6C [8] . However, the

subpopulation from Toulouse was that included in the previous ANTICIPATE study [7] ; this implies a small optimism bias in favor of ANTICIPATE as a new tests in the derivation population, all the more so in that its CO for VNT was determined in this population.
The three other previously-unpublished subpopulations were prospectively recorded but analyzed retrospectively for the present study. These were: i) the population of the SNIFF 95 study performed in Angers, ii) a subpopulation recruited in Cluj and iii) a subpopulation from Bondy.
Duplicated patients were not included in the study. A data sheet including 43 variables was sent to all principalthe 47 investigators of centers between May 2018 and May 2019. All subpopulation studies were approved by the local ethical review boards. The distribution of patients as a function of centers is summarized in Table S8. The high number of patients excluded in the CIRVIR population was mainly due to missing data and unacceptable delays between examinations.
Finally, the 47 centers were located as follows: 35 in France, 6 in Italy, 2 in the United Kingdom, and 1 each in Romania, Switzerland, Canada, and Hong Kong (i.e., 7 countries in all).

Sample size calculation

The sample size was initially calculated for paired proportions using the spared EGD rate of the best previous reference NIT (B6C) estimated at 20% and the best new test estimated at 30%. EB6C were not considered as a reference comparator since we recently found that such an approach was unsafe due to an unacceptably high missed VNT rate [9] . To conclude that the spared EGD rate with the best new NIT was higher than that of the B6C (20%), a total sample size of 1112 patients was required (two-sided McNemar test, alpha=0.5%, 1-beta=90%, proportion of total expected to switch from + to -: 25%, proportion of total expected to switch from - to +: 15%) [10] .

The alpha risk was set at 0.5%, considering recentdiscussions on statistical significance levels [11] . Ultimately, we decided to include even more patients in the derivation population ($\geq$ 1500) due to the definition of missed VNT and the planned subgroup analyses.

Tests evaluated

Platelets and LSM constitute the two major biomarkers for VNT and offer excellent reproducibility [12, 13] . The literature presents several ways to combine them. A first method is to use both biomarkers without transformation, i.e., with crude values. This is the basis of the original B6C and subsequent versions [8, 14] . This first method has the advantage of eloquent COs directly usable in clinical practice but the disadvantage of binary diagnosis with low and high VNT risk. A second method is to use a mathematical transformation, as done by Albrades et al in the ANTICIPATE study, resulting in a logistic score varying from 0 to 1 [7] .

Here, the advantage is a precise VNT risk for each patient and the disadvantage a need for a calculator, partially compensated for by a nomogram. Curiously, no other expression of both biomarkers was used as a single combination. Thus, we compared three published tests for the present study. Patients were categorized according to the B6C [15] to be at low risk for VNT (<5%) when LSM <20 kPa and platelets >150 G/l or to be at high risk for VNT in other cases (i.e., when LSM $\geq$ 20 kPa or platelets $\leq$150 G/L). The COs for the EB6C were respectively 25 kPa and 110 G/l. The ANTICIPATE score is detailed below.

PLR score

Figure 2 furnishes a scatter plot of platelets and LSM values in the derivation population and illustrates how different CO pairs delineate rectangles where patients are at low VNT risk. However, it is reasonable to ask if that is the best way to handle these VNT predictors. When we introduced platelets and LSM in a first logistic regression model, both were strong predictors ($p<0.001$ in the derivation population). In a second model, we added the product of platelets x LSM [16] as an interaction term and found that it too was a strong predictor ($p<0.001$).

However, considering that platelets and LSM vary in opposite senses as a function of liver severity, we constructed a third model using a platelets-to-LSM ratio (PLR). In it, the only significant independent predictor was PLR ($p<0.001$).

The PLR score is a continuous variable ranging from 0.2 to 126 in the whole population; this score is inversely proportional to the VNT risk (Figure S4). Thus, the PLR logit provides the VNT probability (Figures S1 and S4). The logit AUROCs for VNT in the derivation population were, PLR: 0.761 (95% CI: 0.731-0.792), platelets + LSM: 0.751 (0.719-0.784, $p=0.027$ vs PLR by Delong test), LSM: 0.726 (0.693-0.759, $p=0.003$), platelets: 0.684 (0.647-0.722, $p<0.001$); but the PLR provided a much larger patient group for 95% VNT sensitivity than platelets and/or LSM did (Figure S5). It should be noted that AUROC values were the same for raw variables or their logit, but only logit permitted a comparison of LSM + platelets vs PLR.

In the construction of PLER, the PLR CO was specific for each of the six sex / etiology pairs and adjusted for INR in four of these six pairs. Indeed, in one pair (females with viral CLD), PLR adjustment on INR brought no advantage and PLR performance was inferior to a CO pair for platelets and LSM (like in the B6C). In the other pair (males with NAFLD), PLR adjustment on INR brought no advantage and VNT risk at5% was better distinguished by a simple PLR CO.

Surprisingly, and to the best of our knowledge, PLR has not been previously evaluated. Berzigotti et al reported a logistic score called VRS where platelets are multiplied by LSM but within a calculation also including spleen diameter [16] . However, a multiplicative approach amplifies the value range less than a ratio approach since these two variables vary in opposite directions. Moreover and in contrast to the product of a multiplication, this ratio counteracted the interaction between the two variables in the present study.

Employed formulas

ANTICIPATE model:

The logit score was as previously published [7] : -4.458421+1.3193115*ln(LSM) - 0.016306902*platelet count (platelet values over 150 should be introduced as 150). Platelets in G/l, LSM in kPa. The CO for VNT of ANTICIPATE was determined at 0.049 in the present derivation population.

PLER test: PLR score = platelets / LSM

Platelets in G/l, LSM in kPa. PLR COs used in the PLER test are detailed as a function of main etiologies and sex in Table S3.

Estimated VNT risk by PLR score: Logit of score: -0.393 - (0.231*PLR).

PLESIR test:

```
compute PLESIR = 0.
do if (Etiology = virus) and (Sex = M) and (PLR < (-12.793 + (17.244*INR))).
compute PLESIR = 1.
else if (Etiology = virus) and (Sex = F) and (Platelets < 144).
compute PLESIR = 1.
else if (Etiology = virus) and (Sex = F) and (LSM > 24.9).
compute PLESIR = 1.
else if (Etiology = NAFLD) and (Sex = M) and (PLR < 10.58).
compute PLESIR = 1.
else if (Etiology = NAFLD) and (Sex = F) and (PLR < (-191 + (198.571*INR))).
compute PLESIR = 1.
else if (Etiology = Alcohol) and (Sex = M) and (PLR < (-32.268 + (36.321*INR))).
compute PLESIR = 1.
else if (Etiology = Alcohol) and (Sex = F) and (PLR < (-20.264 + (24.671*INR))).
compute PLESIR = 1.
end if.
Execute.
```

Where platelets (G/l), LSM (kPa by VCTE and M probe), 0 (negative test): VNT risk $\leq$5% (or spared endoscopy), 1 (positive test): VNT risk >5%.

This SPSS syntax (IBM, Armonk, NY, USA) is easily translatable in other programming languages. In any case, what are important to note are the values of the variable COs.

The PLER and PLESIR calculations for individual data or databases can be performed free of

charge at the following URL:

http://forge.info.univ-angers.fr/~gh/wstat/pler-plesir-variscreen.php

Discrepancies: when PLER and PLESIR tests are discrepant, the test that rules out VNT is the more reliable one. However, it is preferable to consider VariScreen (see below).

Influence of delays

The dates of all examinations were recorded. Patients with missing dates were excluded. The date of inclusion was that of the EGD. No limits to delays were set in the eligibility criteria to enable an assessment of the maximum allowable delays between EGD and platelets count or LSM. The influence of delays between endoscopy and LSM or platelet count on main outcome measurements is reported in Table S1. This delay influenced the missed VNT rate more than it did the spared EGD rate. The missed VNT rate was proportional to delay as expected, but it stabilized for delays of six months and less. This suggests that a negative test should be repeated every six months. The paradoxically higher rate at the lowest delay (1 month) can be explained by the low number of VNT, which may increase the impact of one patient with missed VNT. A detailed analysis showed that the lowest rate varied between one and six months according to NIT, with the most frequent lowest rate at two months. However, the present method is only suggestive since patients are partially different as a function of delays. The best method would be to measure LSM and platelets at different intervals in all patients, but such an approach would be very difficult to apply in clinical practice. Finally, the present results suggest that for clinical research, the optimal maximum delay is two months. For clinical practice, the optimal delay is six months, furthermore advantageous in that it corresponds to the interval for hepatocellular carcinoma screening in cirrhosis.

Comparison of NITs as a function of etiology and sex or liver severity

Etiology and sex

Test performance as a function of both sex and etiology is detailed in Table S5. Briefly, spared EGD rates were significantly different as a function of etiology in both sexes in all tests (except for ANTICIPATE in females). The spared EGD rate was insufficient (<20%) in alcoholic CLD for B6C, ANTICIPATE and PLER (except in females in the last two tests). The spared EGD rate was higher in NAFLD than in viral CLD in all tests (except for PLESIR in men). The missed VNT rates were significantly different as a function of etiology only in men for B6C, EB6C and ANTICIPATE. This was due to higher missed VNT rates in NAFLD. The missed VNT rates were unsatisfactory in most settings for EB6C. The missed VNT rates were intermediate in NAFLD for B6C and ANTICIPATE in both sexes and in females for PLER and ANTICIPATE for all etiologies (except PLER in alcoholic CLD). Finally, the only test with satisfactory missed VNT and spared EGD rates in every clinical setting was PLESIR. Thus, PLESIR compensated for the limits of PLER in most settings, especially in men with alcoholic CLD (for spared EGD rate) and in females with viral CLD or NAFLD (for missed VNT rate). The only weakness of PLESIR was a relatively low spared EGD rate (23.1%) in men with alcoholic CLD but this was an improvement compared to other acceptable NITs; in all other settings, this PLESIR rate was >30%.

Etiology and liver severity

Liver severity was assessed by MELD or INR. Indeed, as creatinine was missing in one large subpopulation with NAFLD [6] , NIT results as a function of MELD were not reliable in that CLD etiology. Therefore, we also used INR as a surrogate marker of liver severity. Indeed, INR was the composite marker that best correlated with MELD (rs : 0.806 vs bilirubin: 0.647,creatinine: -0.033) (Figure S6). MELD and INR tertiles were concordant in 58.0% of patients and discrepancy >1 tertile was observed in 1.6% of patients.

Test performance as a function of both etiology and liver severity is detailed in Table S6.

Briefly, spared EGD rates significantly decreased as a function of liver failure for all NITs in all etiologies. Importantly, the missed VNT rate decreased as a function of liver severity for all NITs in most etiologies. In the third tertile (CLD with MELD $\geq$ 9.7 or INR $\geq$ 1.2), in addition to EB6C, B6C might be unsafe in NAFLD (missed VNT: 6.1%) together with an intermediate spared EGD rate (11.5%), that is an unfavorable setting. PLESIR was the only secured test (no missed VNT in INR $\geq$ 1.24, Table S7) with an intermediate spared EGD rate in viral CLD and NAFLD (Figure 3). INR COs in Table S7 were estimated by linear regression of the MELD score due to their excellent correlation.

VNT: special cases

VNT in LSM <10 kPa

Four patients with VNT had LSM <10kPa (three at 9.0 kPa); all were female with viral CLD and three had BMI $\geq$28.4 kg/m² and ALT $\leq$33 UI/l; the fourth had a BMI of 23.4 kg/m² and ALT at 112 UI/l.

Characteristics of missed VNT

The characteristics of missed VNT were compared to other VNT as a function of NITs in the whole population. As expected, platelet count and LSM were significantly different. Age, BMI, AST, ALT and bilirubin were not significantly different (data not shown). In missed VNT, albumin was significantly increased in all tests but MELD score was significantly increased only in the PLESIR test and VariScreen (Table S9). The lack of a significant difference in the B6C is attributable to a lack of statistical power since the difference was of the same amplitude as in PLESIR. However, missed VNT by VariScreen had the lowest MELD score. Thus, PLESIR -and even more so VariScreen- secure screening since the occurrence and severity of variceal bleeding are linked to liver failure degree [17] despite higher spared VNT and missed VNT rates than with the B6C reference (Figure S7).

VariScreen algorithm

PLESIR requirement

PLESIR requirement could be reduced from 37.0% to 21.0% ($p<0.001$) by using a PLR CO adapted to each etiology and sex (Table S3). This came at the expense of a reduced spared EGD rate (34.7% to 31.3% in the whole population, $p<0.001$) but also conferred a slightly improved missed VNT rate (3.1% vs 2.2%, $p=0.375$).

Reliability analysis

The misclassification of VNT by VariScreen was tested by logistic regression using the composite markers of VariScreen (plus age and PLER test) as independent variables. This was evaluated in the whole population due to the small event number (n=13) and population similarity. The significant independent predictors were the three interaction pairs between the three tests (PLER, PLESIR, VariScreen). Therefore, we analyzed safety as a function of the six discrepancy cases between these three tests (Table S10).

Among these six cases, one led to using PLESIR instead of VariScreen; and two others led to indicating EGD instead of VariScreen. Finally, this more reliable version of VariScreen provided slightly better performance, increasing the spared EGD rate from 35.1% (33.2-36.8) to 36.1% (34.2-37.8) ($p<0.001$) in the whole population. Safety was maintained, albeit with amissed VNT in one more patient, i.e., a rate increasing slightly from 3.6% (1.7-5.7) to 3.9%(1.9-5.9) ($p=1$) (Table S11). However, VariScreen was apparently less secured since this additional patient with missed VNT had a MELD score at 12.5. We do signal that the MELD score might have overestimated liver function in that case (Table S11). It should be noted that this reliability analysis on test discrepancies is useful and transparent for clinicians in calculator interpretation.

Table S1. Influence of delays between endoscopy and LSM or blood biology on main outcome measurements (%) as a function of three tests. Population including eligible patients.

Delay ≤ (months)	n pts	B6C		EB6C		PLER
		Spared EGD	Missed VNT	Spared EGD	Missed VNT	Spared EGD
No limit	3361	26.9	5.0	46.1	14.5	30.8
12	3049	25.6	4.1	44.4	12.7	29.2
9	2909	25.7	4.0	44.4	12.7	29.0
6	2748 [a]	25.7	3.2	44.4	11.9	29.0
3	2504	25.3	3.4	44.2	11.7	28.6
2	1761	22.4	2.0	40.1	9.9	26.1
1	765	25.5	3.0	41.2	12.7	29.0

[a] The difference with the number of included patients is due to exclusion causes

Color legend:

Missed VNT (% in VNT): ≤5: satisfactory, ≥10: unsatisfactory, intermediate; the lowest rate is indicated in bold.

Spared EGD has no color code because it is not a pertinent outcome measurement here.

Table S2. VNT prevalence (%) as a function of etiology and sex.

	Patients (n)	All	Virus	NAFLD	Alcohol	*p a*
Derivation						
Patients (n)		1579	793	457	323	-
All	1579	15.1	15.0	11.8	20.1	***0.006***
Men	989	18.1	17.7	13.6	23.4	***0.015***
Female	590	10.2	10.8	9.5	9.1	*0.847*
p b	-	***<0.001***	***0.008***	***0.242***	***0.005***	-
Validation						
Patients (n)		789	396	228	165	-
All	789	15.2	14.9	11.8	20.6	*0.056*
Men	514	16.7	15.0	14.7	22.4	*0.148*
Female	275	12.4	14.7	7.6	15.0	*0.236*
p b	-	*0.119*	*1*	*0.143*	*0.375*	-
Both						
Patients (n)		2368	1189	685	494	-
All	2368	15.2	15.0	11.8	20.2	***<0.001***
Men	1503	17.6	16.8	14.0	23.1	***0.003***
Female	865	10.9	12.0	8.9	11.1	*0.412*
p b	-	***<0.001***	***0.030***	***0.044***	***0.005***	-

a Chi² test, b Fisher test with significant values in bold characters

Overall figures in red characters

Table S3. PLR cut-offs for missed VNT rate <5% or 0% as a function of etiology and sex in the derivation population.

	Virus		NAFLD		Alcohol		All	
Missed VNT rate (≤ %)	5	0	5	0	5	0	5	0
Males (>)	6.187	12.2	10.6	14.6	11.03	17	10.18	17
Females (>)	11.8	14.4	15.6	15.6a	10	10	11.8	15.6 a
All (>)	10.22	14.4	14.95	15.6	10.82	>17	10.82	17

a one outlier patient excluded with PLR = 31.6

Table S4. Patient characteristics as a function of etiology in the whole population.

	Virus	NAFLD	Alcohol	p a			
				All	V vs N	V vs A	*N vs A*
n patients	1189	685	494	-	-	-	-
Age (years)	58.6±11.4	**62.4±10.4**	56.1±11.0	<0.001	<0.001	<0.001	<0.001
Sex (% male)	61.6	57.5	**76.3**	<0.001	*0.087*	<0.001	<0.001
VNT (%)	15.0	**11.8**	**20.2**	<0.001	*0.061*	0.009	<0.001
BMI (kg/m²)	26.3±4.2	**31.7±6.1**	26.4±4.9	<0.001	<0.001	0.814	<0.001
AST (IU/l)	75±67	**56±66**	68±52	0.003	*0.004*	0.031	*0.065*
ALT (IU/l)	80±67	56±41	50±47	<0.001	<0.001	<0.001	*0.055*
Normal transaminases (%)	17.2	**3.2**	20.9	<0.001	<0.001	0.083	<0.001
Albumin (g/l)	39.8±5.1	41.6±4.4	37.3±7.1	<0.001	<0.001	<0.001	<0.001
Bilirubin (µmol/l)	18±19	15±13	29±31	<0.001	*0.007*	<0.001	<0.001
INR	1.2±0.2	1.1±0.1	1.3±0.3	<0.001	<0.001	<0.001	<0.001
Platelets (G/l)	**131±63**	176±83	152±79	<0.001	<0.001	<0.001	<0.001
Creatinine (µmol/l)	73±28	79±21	67±42	<0.001	*0.129*	<0.001	*0.002*
MELD score	8.9±2.3	9.2±2.4	**10.9±3.9**	<0.001	*0.379*	<0.001	<0.001
LSM (kPa)	23±13	25±16	**40±24**	<0.001	*0.042*	<0.001	<0.001
LSM ≥10 kPa (%)	92.0	94.7	92.1	0.069	*0.030*	1	*0.071*
LSM reliability (%)	-	-	-	<0.001	*0.320*	<0.001	*0.284*
Very reliable	20.5	26.8	37.7	-	-	-	-
Reliable	70.3	60.7	51.8	-	-	-	-
Poorly reliable	9.1	12.5	10.5	-	-	-	-
LSM CO for VNT (kPa)	9.0	10.3 b	01/10/00	9.0	-	-	-

LSM: liver stiffness measurement, VNT: varices needing treatment

a Student's t test / ANOVA or Chi² test / Fisher test, d One outlier value at 6.3 excluded

Specific characteristic in bold characters

Table S5. Test performance as a function of etiology and sex in the whole population.

Test	Missed VNT rate (%)					Spared EGD rate (%)				
	All	Virus	NAFLD	ALD	*p a*	All	Virus	NAFLD	ALD	*p a*
B6C										
All	2.5	1.1	7.4	1.0	***0.006***	23.9	21.6	33.4	16.0	***<0.001***
Men	2.3	0.8	7.3	1.1	***0.019***	22.1	21.0	31.0	14.9	***<0.001***
Female	3.2	1.8	7.7	0.0	*0.291*	26.9	22.5	36.8	19.7	***<0.001***
p b	*0.702*	*0.524*	*1*	*1*	-	***0.008***	*0.563*	*0.120*	*0.248*	-
EB6C										
All	11.1	10.1	19.8	6.0	***0.012***	42.7	42.1	55.8	25.9	***<0.001***
Men	10.6	9.8	21.8	4.6	***0.005***	40.4	41.0	54.8	24.1	***<0.001***
Female	12.8	10.9	15.4	15.4	*0.814*	46.7	44.0	57.0	31.6	***<0.001***
p b	*0.570*	*0.793*	*0.565*	*0.173*	-	***0.003***	*0.334*	*0.586*	*0.117*	-
Anticipate										
All	3.3	2.2	7.4	2.0	*0.069*	24.5	25.0	28.9	17.2	***<0.001***
Men	2.3	0.8	7.3	1.1	***0.019***	23.8	25.0	28.9	15.9	***<0.001***
Female	6.4	5.5	7.7	7.7	*0.909*	25.8	24.9	28.9	21.4	*0.246*
p b	*0.088*	*0.088*	*1*	*0.244*	-	*0.275*	*1*	*1*	*0.206*	-
PLER										
All	3.6	3.4	4.9	3.0	*0.761*	27.3	25.9	35.3	19.6	***<0.001***
Men	2.3	0.8	3.6	3.4	*0.335*	25.7	25.8	33.8	17.0	***<0.001***
Female	7.4	9.1	7.7	0.0	*0.532*	30.2	26.0	37.5	28.2	***0.004***
p b	***0.046***	***0.011***	*0.590*	*1*	-	***0.019***	*0.946*	*0.332*	***0.011***	-
PLESIR										
All	3.6	3.9	3.7	3.0	*0.922*	35.2	38.0	36.9	25.9	***<0.001***
Men	3.8	4.1	3.6	3.4	*0.972*	35.5	42.5	34.5	23.1	***<0.001***
Female	3.2	3.6	3.8	0.0	*0.779*	34.6	30.9	40.2	35.0	***0.032***
p b	*1*	*1*	*1*	*1*	-	*0.655*	***<0.001***	*0.129*	***0.011***	-

a Chi² test

b Fisher test

Significant difference in bold characters

Color legend:

Missed VNT (% in VNT): ≤ 5: satisfactory, ≥ 10: unsatisfactory, intermediate

Spared EGD (%): ≥ 20: satisfactory, ≤ 10: unsatisfactory, intermediate

Table S6A. Test performance as a function of both etiology and INR tertiles (0.79-1.07, 1.08-1.19, 1.20-3.98) in the whole population.

Test	Missed VNT rate (%)					Spared EGD rate (%)				
	All	Virus	NAFLD	ALD	p a	All	Virus	NAFLD	ALD	*p a*
Patients (n)	359	178	81	100		2368	1189	685	494	-
B6C										
Tertile 1	9.3	5.3	16.7	0	0.343	42.6	37.9	47.6	42.2	***0.034***
Tertile 2	2.3	0	3.3	8.3	0.213	22.8	23.7	20.1	24.1	***0.558***
Tertile 3	1.3	0	6.1	0	**0.029**	6.0	6.4	11.5	3.0	***0.005***
p b	***0.009***	*0.174*	*0.216*	***0.025***		***<0.001***	***<0.001***	***<0.001***	***<0.001***	
EB6C										
Tertile 1	30.2	21.1	44.4	16.7	0.222	66.8	64.1	71.7	56.7	***0.010***
Tertile 2	11.6	9.1	13.3	16.7	0.720	45.1	46.0	45.1	42.3	*0.759*
Tertile 3	7.4	8.7	12.1	3.7	0.219	16.4	20.5	23.0	7.1	***<0.001***
p b	***<0.001***	*0.246*	***0.012***	*0.109*		***<0.001***	***<0.001***	***<0.001***	***<0.001***	
Anticipate										
Tertile 1	16.3	15.8	22.2	0	0.441	42.7	44.3	40.5	45.6	*0.496*
Tertile 2	2.3	0	3.3	8.3	0.213	23.3	25.1	19.1	24.1	*0.244*
Tertile 3	1.3	0.9	3.0	1.2	0.626	7.5	9.2	8.8	4.1	***0.039***
p b	***<0.001***	***<0.001***	***0.025***	*0.243*		***<0.001***	***<0.001***	***<0.001***	***<0.001***	
PLER										
Tertile 1	18.6	21.1	16.7	16.7	0.935	48.2	46.1	49.7	50.0	*0.581*
Tertile 2	2.3	0	3.3	8.3	0.213	24.4	24.4	22.5	27.0	*0.643*
Tertile 3	1.3	1.7	0	1.2	0.737	9.3	11.1	11.5	5.6	***0.038***
p b	***<0.001***	***<0.001***	***0.028***	*0.052*		***<0.001***	***<0.001***	***<0.001***	***<0.001***	
PLESIR										
Tertile 1	23.3	31.6	11.1	33.3	0.277	63.5	67.1	57.6	74.4	***0.002***
Tertile 2	2.3	0	3.3	8.3	0.213	33.0	40.8	14.7	36.5	***<0.001***
Tertile 3	0.4	0.9	0	0	0.605	*9.0*	11.8	9.7	4.1	***0.003***
p b	***<0.001***	***<0.001***	*0.132*	***<0.001***		***<0.001***	***<0.001***	***<0.001***	***<0.001***	

a Chi² test

b Fisher test

Significant difference in bold characters

Color legend:

Missed VNT (% in VNT): ≤ 5: satisfactory, ≥10: unsatisfactory, intermediate. This must be interpreted with caution in this subgroup analysis

Spared EGD (%): ≥20: satisfactory, ≤ 10: unsatisfactory, intermediate

Table S6B. Test performance as a function of both etiology and MELD tertiles (6.4-7.8, 7.9-9.6, 9.7-28.3) in the whole population. MELD results are not reliable in NAFLD due to missing data.

Test	Missed VNT rate (%)					Spared EGD rate (%)				
	All	Virus	NAFLD	ALD	p a	All	Virus	NAFLD	ALD	*p a*
Patients (n)	278	172	8	98	-	1730	1178	78	484	-
B6C										
Tertile 1	3.3	4.5	0	0	0.829	35.9	35.5	28.6	39.2	*0.605*
Tertile 2	1.3	1.6	0	0	0.901	18.0	15.5	37.0	21.8	***0.008***
Tertile 3	0.6	0	0	1.3	0.561	8.3	9.0	43.3	3.2	***<0.001***
p b	*0.393*	*0.191*	*NA*	*0.893*		***<0.001***	***<0.001***	*0.582*	***<0.001***	
EB6C										
Tertile 1	13.6	13.3	0	16.7	0.832	61.7	62.5	57.1	58.8	*0.717*
Tertile 2	13.2	14.3	0	8.3	0.792	35.0	33.7	55.6	34.6	*0.070*
Tertile 3	6.4	6.9	20.0	5.0	0.398	18.3	23.3	60.0	7.2	***<0.001***
p b	*0.156*	*0.301*	*0.710*	*0.487*		***<0.001***	***<0.001***	*0.942*	***<0.001***	
Anticipate										
Tertile 1	3.3	4.5	0	0	0.829	40.9	41.2	23.8	43.1	*0.251*
Tertile 2	3.9	4.8	0	0	0.725	21.5	19.2	48.1	23.3	***0.002***
Tertile 3	1.2	0	0	2.5	0.312	8.6	9.3	46.7	3.2	***<0.001***
p b	*0.340*	*0.123*	*NA*	*0.795*		***<0.001***	***<0.001***	*0.171*	***<0.001***	
PLER										
Tertile 1	3.3	4.5	0	0	0.829	42.0	41.7	23.8	47.1	*0.139*
Tertile 2	5.3	6.3	0	0	0.647	22.9	20.9	44.4	24.8	***0.016***
Tertile 3	1.7	1.1	0	2.5	0.765	10.0	10.3	46.7	5.2	***<0.001***
p b	*0.307*	*0.221*	*NA*	*0.795*		***<0.001***	***<0.001***	*0.215*	***<0.001***	
PLESIR										
Tertile 1	13.3	13.6	0	16.7	0.832	61.0	62.5	23.8	61.8	***0.002***
Tertile 2	3.9	4.8	0	0	0.725	30.2	28.2	40.7	34.6	*0.178*
Tertile 3	1.2	1.1	0	1.3	0.968	12.4	15.3	33.3	6.4	***<0.001***
p b	***0.002***	***0.028***	*NA*	***0.031***		***<0.001***	***<0.001***	*0.467*	***<0.001***	

a Chi² test

b Fisher test

Significant difference in bold characters

Color legend:

Missed VNT (% in VNT): ≤5: satisfactory, ≥ 10: unsatisfactory, intermediate. This must be interpreted with caution in this subgroup analysis

Spared EGD (%): ≥20: satisfactory, ≤10: unsatisfactory, intermediate

Table S7. Missed VNT rate (%) of NITs as a function of etiology in patients with poor CLD and VNT of the whole population. This table evaluates secureness.

	Patients (n)	B6C	EB6C	ANTICIPATE	PLER	PLESIR	VariScreen
MELD ≥ 10							
All	158	0	10.8	0.6	1.3	0.6	0
Virus	77	0	9.1	0	1.3	0	0
NAFLD	4	0	25.0	0	0	0	0
Alcohol	77	0	11.7	1.3	1.3	1.3	0
p a	-	*NA*	*0.566*	*0.569*	*0.566*	*0.569*	*NA*
MELD ≥ 12							
All	95	0	6.3	1.1	1.1	1.1	0
Virus	42	0	2.4	0	0	0	0
NAFLD	1	0	0	0	0	0	0
Alcohol	52	0	9.6	1.9	1.9	1.9	0
p a	-	*NA*	*0.346*	*0.658*	*0.658*	*0.658*	*NA*
INR ≥ 1.24 b							
All	180	1.1	11.1	1.7	1.7	0	1.1
Virus	82	0	12.2	1.2	2.4	0	1.2
NAFLD	22	8.3	12.5	4.2 c	0	0	4.2 c
Alcohol	74	0	9.5	1.4	1.4	0	0
p a	-	***0.001***	*0.840*	*0.589*	*0.687*	*NA*	*0.237*
INR ≥ 1.37 b							
All	98	1.0	8.2	1.0	2.0	0	1.0
Virus	37	0	5.4	0	2.7	0	0
NAFLD	6	14.3 c	14.3 c	0	0	0	14.3 c
Alcohol	54	0	9.3	1.9	1.9	0	0
p a	-	***0.001***	*0.666*	*0.663*	*0.888*	*NA*	***0.001***

a Chi² test

b CO estimated by linear regression of MELD score due to excellent correlation (see Figure S6)

c One patient

Significant difference in bold characters

Color legend:

Missed VNT (% in VNT): 0: secured, ≤ 5: satisfactory, ≥ 10: unsatisfactory, intermediate

Table S8. Distribution of patients as a function of centers or multicenter cohort/study.

Center number	Center / multicenter study	Main investigator	Patients	
			n	% a
1	Bologna	D. Festi	498	21.0
2	Verona/ Bern	1. Colecchia&A.Berzigotti	103	4.3
3	Multicenter M116 study	H. Stefanescu	217	9.2
4	Multicenter VO-VCO study	S. Huvelin	203	8.6
5	Multicenter CIRVIR cohort	P. Nahon & JP. Zarski	169	7.1
6	Angers	P. Calès	232	9.8
7	Cluj	H. Stefanescu	72	3.0
8	Toulouse C. Bure	C. Bureau	219	9.2
9	Bondy	N. Ganne-Carriè	48	2.0
10	Multicenter NAFLD-B6C study	S. Petta	607	25.6

a Rates per population

Table S9. Liver failure degree as a function of VNT status in the whole VEB6 population.

	B6C	EB6C	ANTICIPATE	PLER	PLESIR	VariScreen
Albumin (g/l) in 342 patients:						
Missed VNT	41.8±3.3	39.7±5.5	41.3±4.6	40.5±5.2	40.3±4.3	40.2±5.0
Other VNT	36.6±5.7	36.4±5.7	36.6±5.7	36.6±5.7	36.6±5.7	36.6±5.7
p a	***0.007***	***0.001***	***0.005***	***0.020***	***0.024***	***0.028***
MELD score in 278 patients:						
Missed VNT	8.3±1.5	10.2±3.7	9.9±2.8	10.5±4.5	8.2±2.0	8.0±1.3
Other VNT	11.5±3.7	11.6±3.7	11.5±3.7	11.5±3.7	11.6±3.7	11.5±3.7
P a	*0.138*	*0.077*	*0.290*	*0.470*	***0.007***	***<0.001***

a Student's t test

Significant difference in bold characters

Table S10. Reliability analysis of VariScreen algorithm. Safety as a function of discordances between PLER, PLESIR and VariScreen in the whole population.

Tests	**Code**	**n**	**PLER**		**PLESIR**		**VariScreen**		**Choice** a
			Code	**VNT (%)**	**Code b**	**VNT (%)**	**Code**	**VNT (%)**	
Discordant									
VariScreen:									
	0	0	1	**0**	1	**0**	0	0	VariScreen (EGD)
	1	6	0	16.7	0	16.7	1	**16.7 c**	*EGD*
p d	-	-	-	*1*	-	*1*	-	*1*	
PLESIR:									
	0	21	1	**4.8**	0	4.8	1	**4.8**	VariScreen
	1	30	0	6.7	1	**6.7 e**	0	6.7	*PLESIR*
p d	-	-	-	*1*	-	*1*	-	*1*	-
PLER:									
	0	213	0	1.9	1	**1.9**	1	**1.9**	VariScreen
	1	36	1	**13.9**	0	13.9	0	13.9	EGD
p d	-	-	-	*0.004*	-	*0.004*	-	*0.004*	-
Concordant									
	0	1133	0	23.0	0	23.0	0	23.0	EGD
	1	590	1	**1.2**	1	**1.2**	1	**1.2**	VariScreen
p d	-	-	-	*<0.001*	-	*<0.001*	-	*<0.001*	-

a Test used in the improved VariScreen algorithm. Test in italics correspond to change vs original VariScreen. 3 patients change for EGD status: 1 required and 2 spared in improved VariScreen (details below)

b 0: EGD required, 1: EGD spared

c 1 patient: female with NAFLD, ALT=59 and INR=1.55

d Fisher test

e 2 patients: 6.7% (0.0-16.1): males with viral CLD, ALT=58/72 and INR=1.00 both; so, intermediate missed VNT risk in preserved liver function is acceptable.

Bold characters: missed VNT rate

Color legend:

Missed VNT (% in VNT): ≤5: satisfactory, ≥10: unsatisfactory, intermediate

Table S11. Comparison of VariScreen algorithm and its improved reliable version.

	Whole population	Derivation	Validation	*p a*
Patients (n)	2368	1579	789	-
Spared EGD rate (%):				
VariScreen	35.1	34.5	36.2	*0.411*
Reliable VariScreen	36.1	35.7	36.9	*0.586*
p b	*<0.001*	*<0.001*	*0.180*	-
Missed VNT rate (%):				
In all patients:				
VariScreen	3.6	2.9	5.0	*0.373*
Reliable VariScreen	3.9	3.3	5.0	*0.564*
p b	*1*	*1*	*1*	-
In MELD score ≥10 d:				
VariScreen	0	0	0	*NA*
Reliable VariScreen	0.6	1.0 c	0	*1*
p b	*1*	*1*	*1*	-
In MELD score ≥12 (%) d:				
VariScreen	0	0	0	*NA*
Reliable VariScreen	1.1	1.6 c	0	*1*
p b	*1*	*1*	*1*	-

NA: not available

a Unpaired Fisher test

b Paired McNemar test

c One patient with viral CLD, MELD=12.5 (bilirubin =80, creatinine=62, INR=1.0), albumin=35 and AST=122

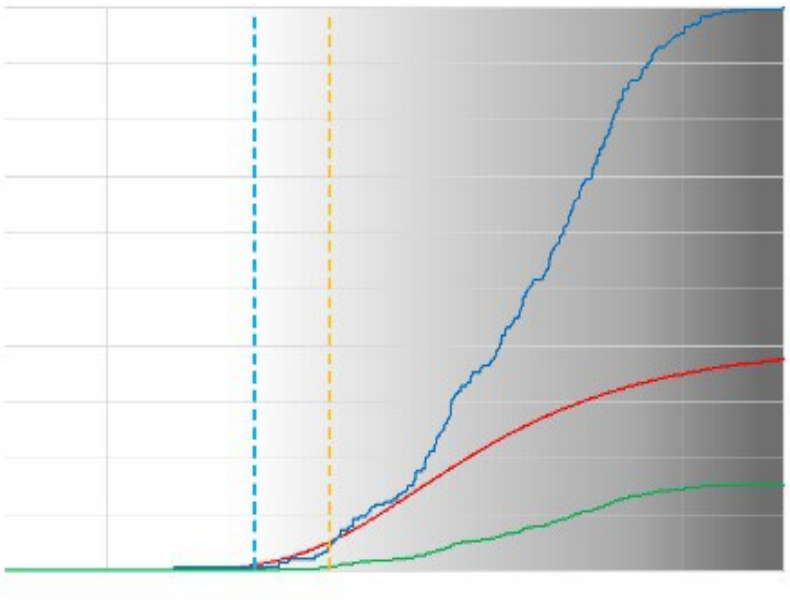


Figure S1. *Different VNT rates (Y axis) as a function of PLR score (X axis) (inverted sense) in the derivation population. The blue, red and green curves indicate respectively the cumulated missed VNT rate, the VNT risk estimated by logit of PLER, and the cumulated VNT prevalence. This plot shows that the cut-offs for missed VNT rate (with a maximum spared endoscopy rate on the left side) are easy to determine and indicated here by the vertical blue dashed line for a 0% cut-off and the vertical orange dashed line for a 5% cut-off.*

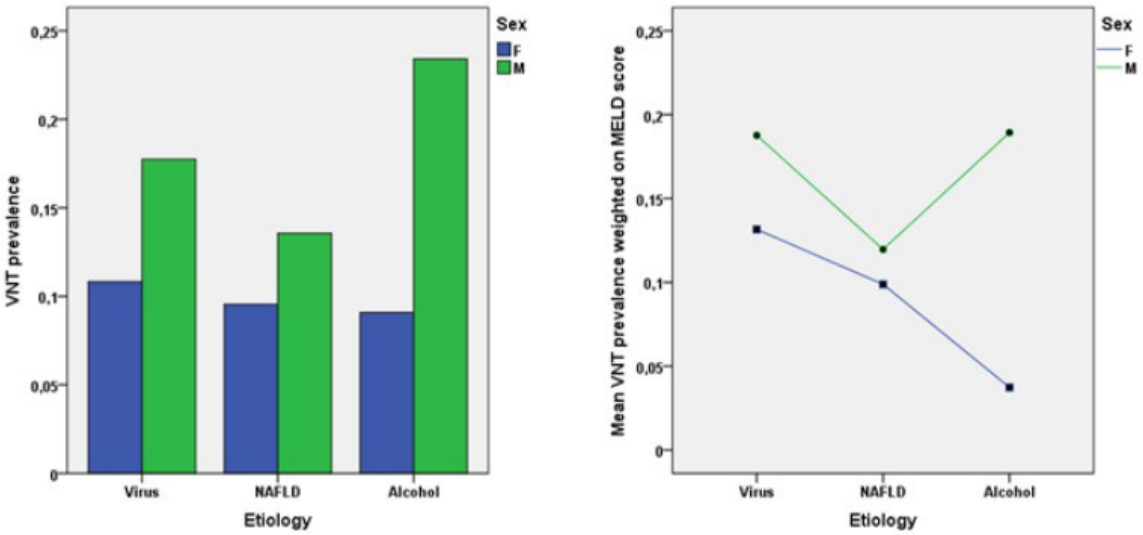


Figure S2. VNT prevalence as a function of sex and etiology in the derivation population. Left panel: crude prevalence. Right panel: weighted prevalence on MELD score. MELD was the only independent predictor (p<0.001) by ANCOVA.

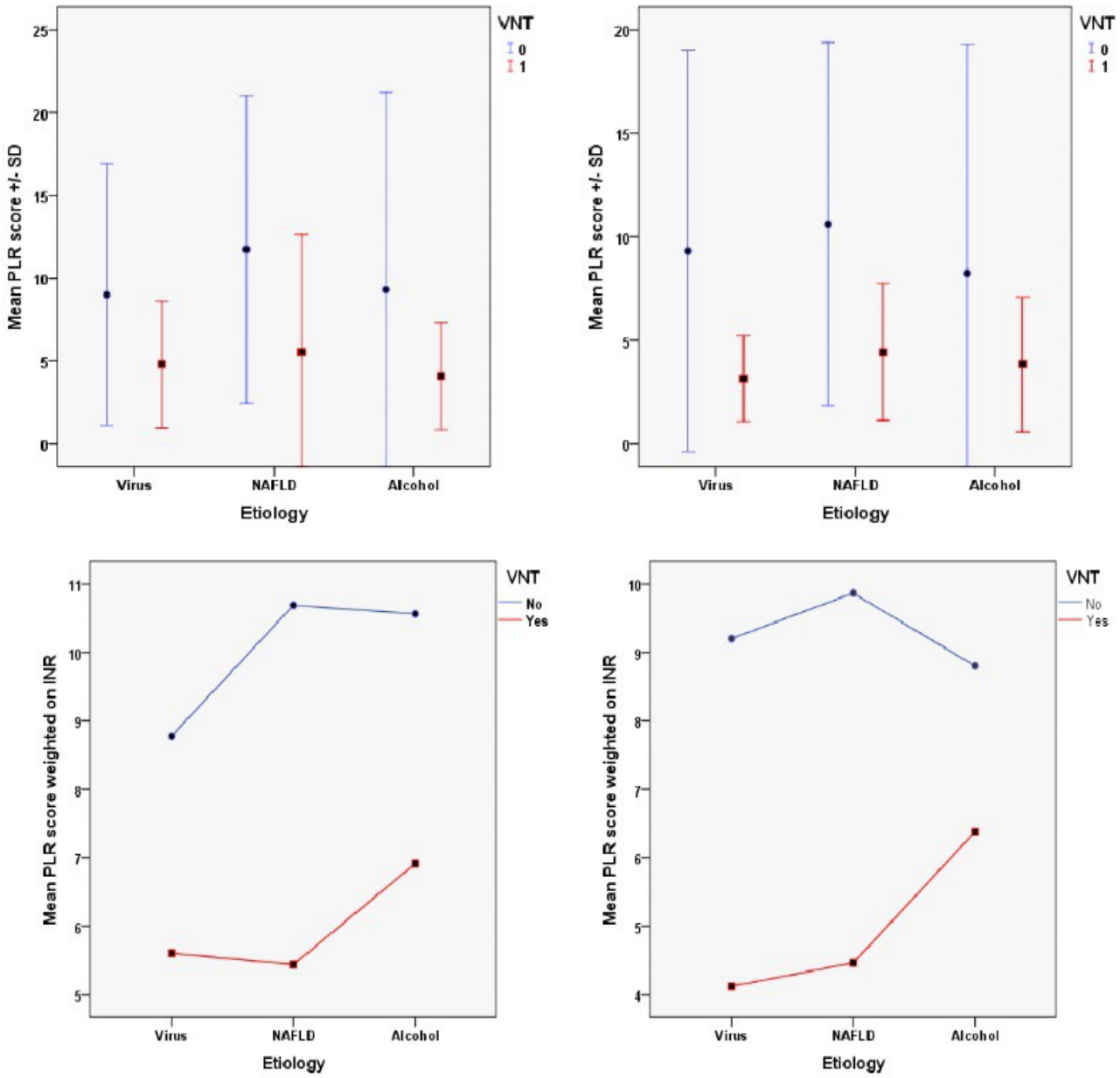


Figure S3. Mean PLR score as a function of sex and etiology in the derivation population. Panel A: crude score. Panel B: weighted score on MELD score. MELD was the only independent predictor (p<0.001) by ANCOVA

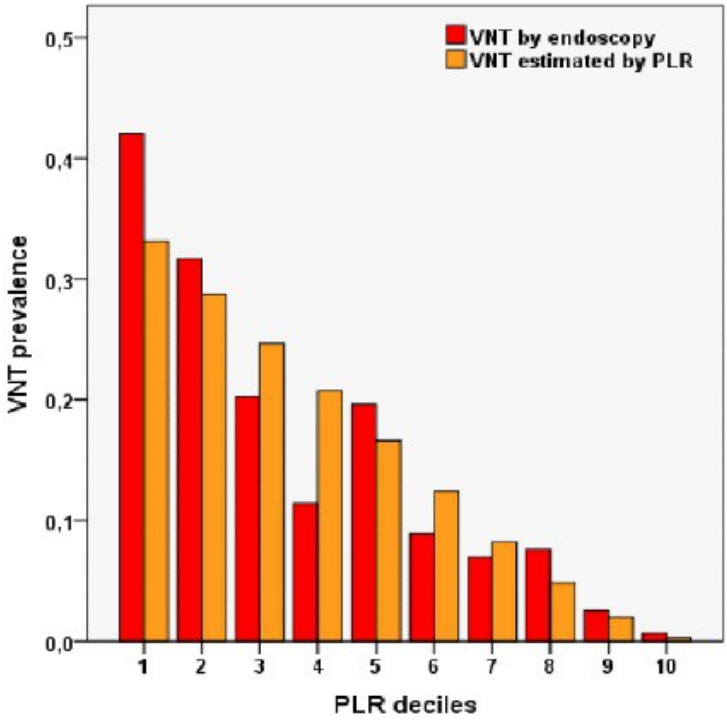


Figure S4. VNT prevalence estimated by endoscopy and logit of PLR score as a function of logit PLR deciles in the derivation population.

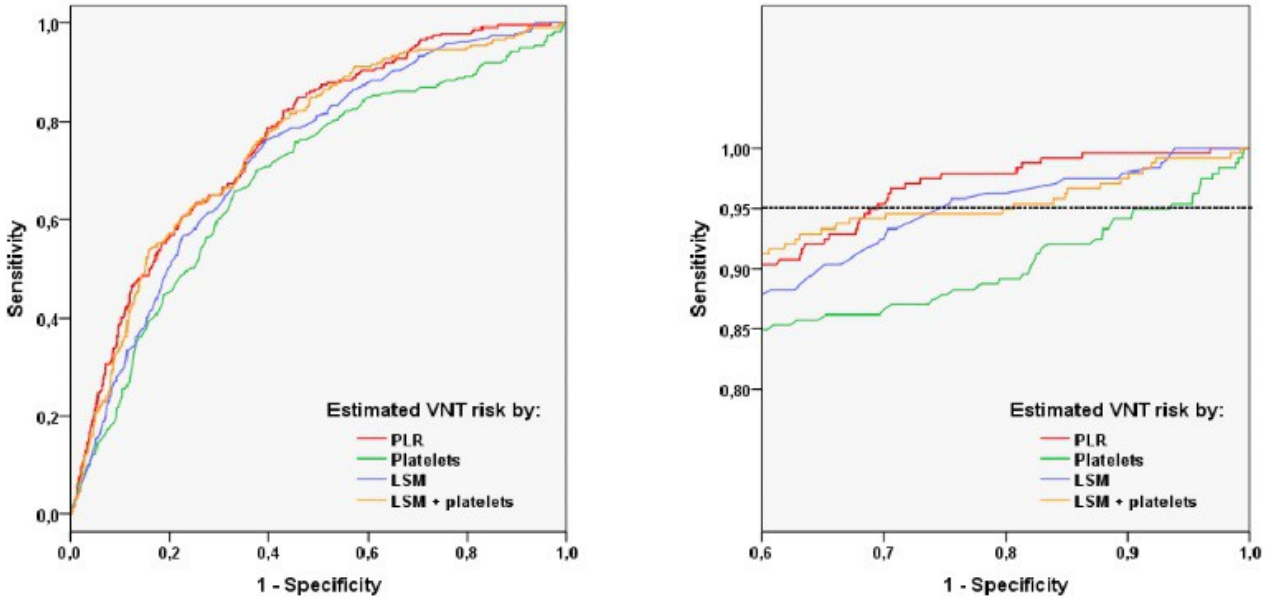


Figure S5. AUROCs for VNT of PLR score and its composite markers (logit) in the derivation population. Right panel: details showing that PLR score had the largest population for 95% sensitivity for VNT, i.e. 5% missed VNT rate (dashed line).

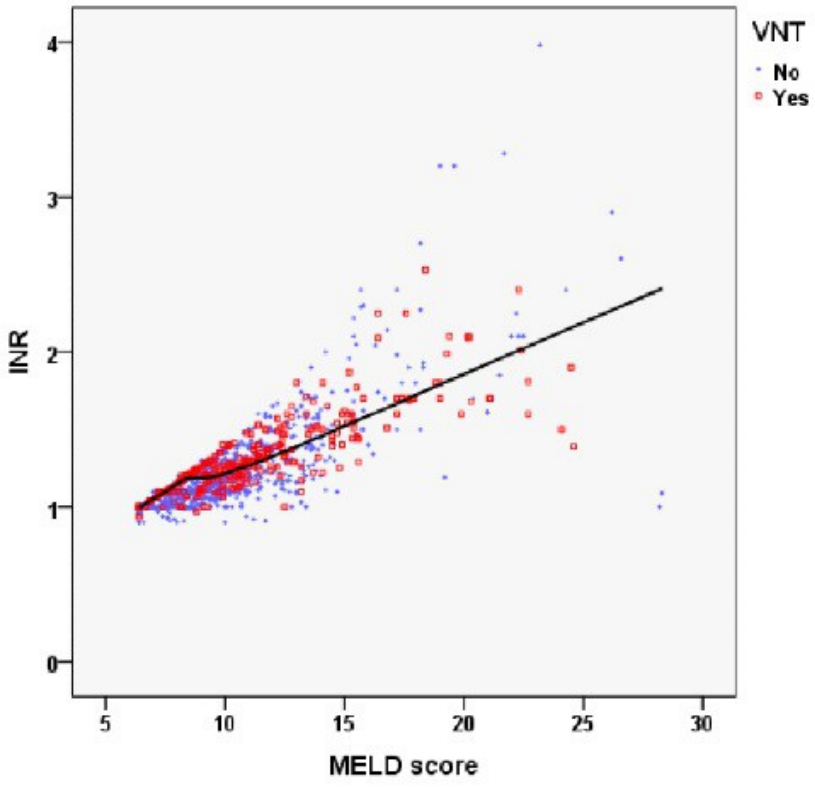


Figure S6. Correlation of MELD score and INR in the whole population. The black line indicates the non-linear regression by LOWESS showing an excellent trend for colinearity

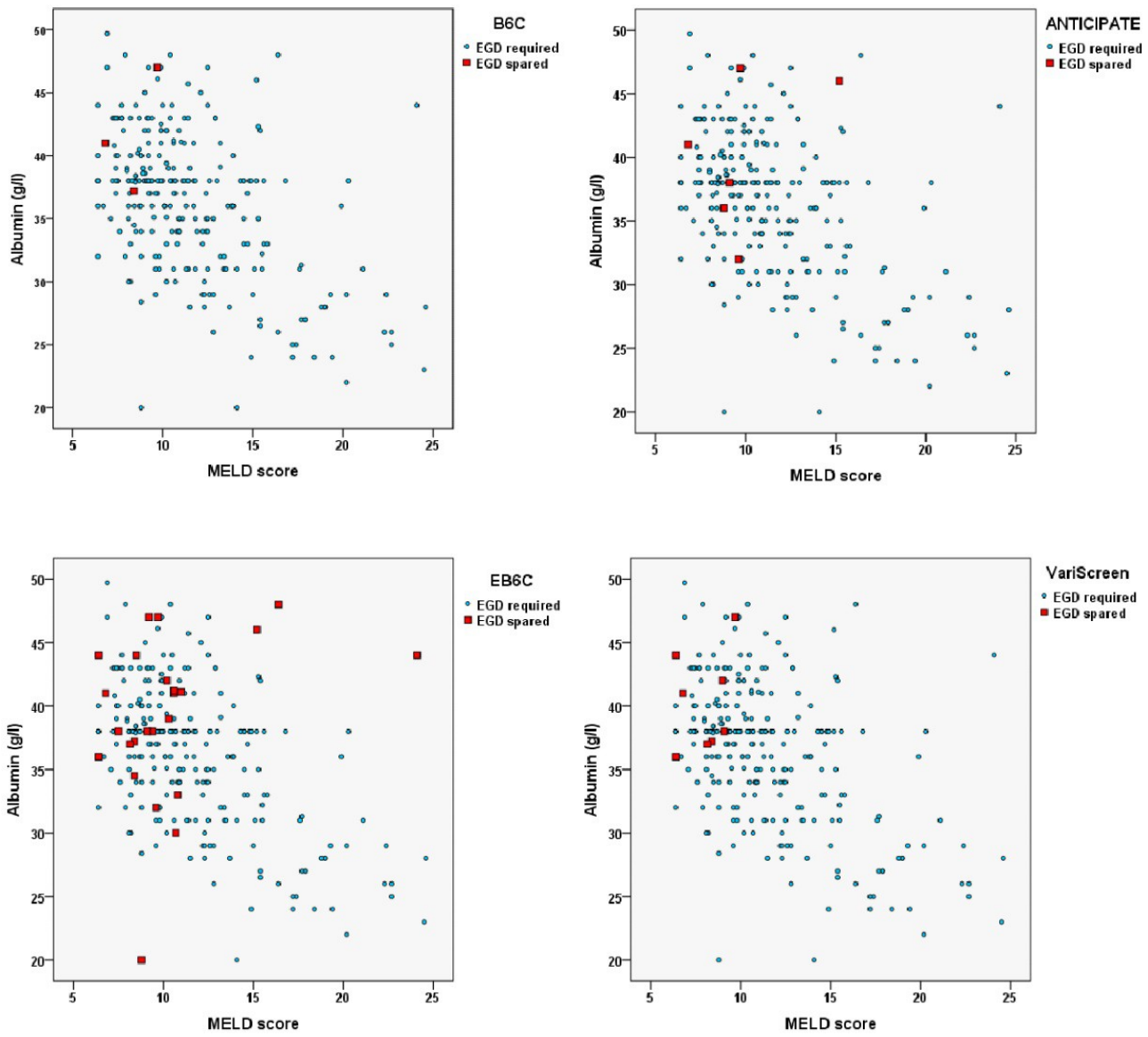


Figure S7. VNT status as a function of four tests in the whole population (262 patients with VNT). These figures show that VariScreen is as safe as B6C despite a higher missed VNT rate (in red); by contrast, ANTICIPATE and even EB6C are unsafe due not only to a high missed VNT rate but also to several missed VNT with higher liver failure degree.

Received: 4 June 2018 | Revised: 20 July 2018 | Accepted: 4 August 2018

DOI: 10.1111/liv.13945

BRIEF DEFINITIVE REPORT

WILEY Liver INTERNATIONAL

How to clarify the Baveno VI criteria for ruling out varices needing treatment by noninvasive tests

Paul Calès[1] | François Buisson[1] | Federico Ravaioli[2] | Arthur Berger[1] | Carlotta Carboni[2] | Giovanni Marasco[2] | Davide Festi[2]

[1]Hepatology department, University Hospital, Angers, France

[2]Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy

Correspondence
Paul Calès, Hepatology, CHU, Angers, France.
Email: paul.cales@univ-angers.fr

Handling Editor: Mario Mondelli

Abstract

Background & Aims: Baveno VI criteria enabled the screening of varices needing treatment (VNT) without endoscopy but created confusion by not stating the method used to calculate the 5% missed VNT limit, resulting in different calculations across validation studies. We analysed those calculations to clarify their diagnostic meaning.

Methods: (a) Literature review and recalculation of the missed VNT rates according to the three definitions encountered. (b) Contingency table comparison of these latter to determine their diagnostic meanings. (c) Real case analysis. 4/Simulation of variations in the three main statistical descriptors (VNT, missed VNT or spared endoscopies).

Results: Missed VNT rates in the three definitions varied five- to 10-fold across 7 papers. The contingency table showed that the definitions based on VNT prevalence and spared endoscopy as reference corresponded, respectively, to sensitivity and negative predictive value (NPV). The whole population-based definition corresponded to diagnostic accuracy (not pertinent in that setting). Real case analysis showed that concerning liver stiffness, the 95% sensitivity and NPV cut-offs for VNT were, respectively, 14.1 and 26.5 kPa. The VNT-based definition offered a more statistically powerful paired comparison between diagnostic tests, whereas the definition based on spared endoscopies was hampered by an unpaired comparison. Case simulation showed that the VNT-based definition was the most sensitive to descriptor variations.

Conclusion: The definitions of missed VNT rate placing VNT or spared endoscopy as the denominator are appropriate, providing, respectively, sensitivity and NPV for VNT. We privilege the first since it corresponds to the true proportion of missed VNT.

KEYWORDS

Baveno VI criteria, noninvasive diagnosis, oesophageal varices, sensitivity

Abbreviations: LSM, liver stiffness measurement; NPV, negative predictive value; VCTE, vibration-controlled transient elastography; VNT, varices needing treatment.

1 | INTRODUCTION

For the first time, the Baveno VI Consensus Workshop[1] has stated that the association of blood platelets >150 g/L and liver stiffness measurement (LSM) by vibration-controlled transient elastography (VCTE) <20 kPa is sufficient to rule out varices needing treatment (VNT) noninvasively, thus sparing patients with compensated advanced chronic liver disease from upper gastrointestinal endoscopy ("endoscopy" hereafter). These Baveno VI criteria fixed a maximum rate of missed VNT at 5%, but also created an ambiguity by not stating the method used to calculate that rate.[1,2] Thus, the papers aiming to validate these criteria have used three different methods to calculate missed VNT. In all of them, the numerator was the number of missed VNT, but the denominator was either: the number of spared endoscopies, as done by Bae et al[3] in a recent issue of this journal; the number of VNT, as we did recently also in this journal[4] and elsewhere[5]; or the whole patient population, as done by Augustin et al.[6] Therefore, our aim with the present work was to perform a literature review and analyse the calculation methods to clarify their diagnostic meaning.

Key points

- Baveno VI criteria enable the screening of varices needing treatment (VNT) without endoscopy.
- The criterion maximum missed VNT rate is ≤5%, but the calculation for this rate was not stated, resulting in three definitions in the literature; consequently, the range between the three rates varies five- to 10-fold across 7 papers.
- The missed VNT rate calculated among VNT corresponds to sensitivity for VNT while that calculated among spared endoscopy corresponds to negative predictive value for them.
- We privilege the VNT-based definition, since it provides the true missed VNT rate and offers a more powerful statistical comparison between diagnostic tests.

2 | MATERIALS AND METHODS

2.1 | Literature review

We reviewed the literature, found three missed VNT calculation methods ("definitions" hereafter), and recorded or recalculated the rates of missed VNT according to them. We labelled these definitions as follows:

1. "VNT," where the missed VNT rate = number of patients (np) with missed VNT/np VNT;
2. "spared endoscopy," where the missed VNT rate = np missed VNT/np spared endoscopy;
3. "whole population," where the missed VNT rate = np missed VNT/whole np.

2.2 | Diagnostic meaning

Using the classical 2 × 2 contingency table providing diagnostic indices (Table S1), we aimed to deduce the diagnostic meaning of the three definitions encountered in the literature by comparing them to the classical diagnostic indices: sensitivity, specificity, negative (NPV) and positive predictive values, and diagnostic accuracy.

2.3 | Real case analysis

In our VO-VCO population,[4] we compared the missed VNT rates of the expanded Baveno VI criteria[6] with those of a blood test (CirrhoMeter, Echosens, Paris, France) as a function of the three definitions. Additionally, to provide an example of the influence of a definition on diagnostic cut-offs, we calculated the curves of diagnostic indices of LSM by VCTE (Fibroscan™, Echosens) for VNT.

2.4 | Case simulation

We performed a simulation (Table S2) wherein one of the three main statistical descriptors (VNT prevalence, missed VNT or spared endoscopies) was varied while the two others were kept constant (if not affected by the former), subsequently providing three cases. Three different VNT prevalences were simulated in each case. We simulated an additional case where two descriptors varied.

3 | RESULTS

3.1 | Literature review

Table 1 shows that the extreme rates of missed VNT between definitions varied from five- to 10-fold in the 7 articles analysed. Thus, rates considered as acceptable (≤5%) in one definition could be largely unacceptable (eg 33%) in another. It should also be noted that the missed VNT rates decreased according to the definitions in the order: VNT < spared endoscopy < whole population. The spared endoscopy definition was the most frequently used method.

3.2 | Diagnostic meaning

All definitions placed missed VNT cases as the numerator of the ratio (Table S1). In the VNT definition (np VNT as denominator), the ratio corresponded to 1–sensitivity for VNT, that is, a minimum 95% sensitivity. In the spared endoscopy definition (np spared endoscopy as denominator), the ratio corresponded to 1–NPV for VNT, that is, a minimum 95% NPV. In the whole population definition (number of all patients as denominator), the ratio corresponded to 1–accuracy; it provided the prevalence of missed VNT which is the prevalence of misclassified patients for VNT.

TABLE 1 Rates of missed VNT according to the reference used as the denominator

			Missed VNT (%)			
Author	Test used	Pts	VNT	Spared endoscopy	Pts	Ratio[a]
Augustin[6]	Expanded B6C	925	*6.5*	**1.6**	**0.6**	10.1
	B6C	925	3.3	**1.5**	**0.3**	10.1
	Expanded B6 + MELD=6	883	NA	**1.7**	**0.8**	NA
Jangouk[7]	B6C	161	0.0	**0.0**	0.0	NA
	Plt>150G/l + MELD=6	161	0.0	**0.0**	0.0	NA
	B6C	101	0.0	**0.0**	0.0	NA
	Plt>150G/l + MELD=6	101	*5.9*	**3.3**	1.0	5.9
Roccarina[8], [b]	B6C	4569	2.1	1.0	**0.2**	10.4
Calès[4]	VariScreen	287	**6.7**	1.7	1.2	5.7
Bae[3]	Expanded B6C	282	*18.2*	**6.8**	3.5	5.1
	B6C	282	5.5	**3.8**	1.1	5.1
Petta[9]	Expanded B6C	338	*17.7*	**4.4**	2.4	7.4
	B6C	338	*11.1*	**4.4**	1.1	10.1
Colecchia[5]	B6C	498[c]	**1.0**	1.0	0.2	5.0
	B6C	115[d]	**0.0**	0.0	0.0	NA
	Expanded B6C	498[c]	**8.0**	3.9	1.6	5.0
	Expanded B6C	115[d]	**33.3**	8.6	4.4	7.6
	B6C/SSM ≤46 kPa	498[c]	**3.0**	1.4	0.6	5.0
	B6C/SSM ≤46 kPa	115[d]	**0.0**	0.0	0.0	NA

B6C, Baveno VI criteria; MELD, Model for End-Stage Liver Disease; NA, not available or applicable; Plt, platelet; Pts, number of patients; SSM, spleen stiffness measurement; VNT, varices needing treatment.
Missed VNT: in bold for method used in the article and in italics when > 5%.
[a]Ratio between extreme values.
[b]This review includes 13 studies, most of which were published in abstract form; two were fully published.[4,7]
[c]Retrospective cohort.
[d]Prospective cohort.

3.3 | Real case

3.3.1 | Test comparison

The comparison between the expanded Baveno VI criteria and the blood test showed two noteworthy results (Table S3). Firstly, the missed VNT rate by the blood test was acceptable in the whole population definition (4.6%) or approached acceptability in spared endoscopy definition (9.4%) but was quite unacceptable in the VNT definition (27.0%). Secondly, the VNT and whole population definitions had a strong statistical advantage over the spared endoscopy definition. Indeed, in the VNT and whole population definitions, the statistical comparison of missed VNT rates between diagnostic tests was paired (ie in the same patient group), which favoured the statistical significance of the rate difference (eg by using a McNemar test). By contrast, in the spared endoscopy definition, the statistical comparison was unpaired (ie between two different patient groups), which underpowered the comparison (eg by using a chi square test). Thus, in the example, the difference in missed VNT rate between the two diagnostic tests was significant with the VNT and whole population definitions ($P = 0.012$) but borderline with the spared endoscopy definition ($P = 0.052$) although the test difference was two-fold (8.2%) that of the whole population definition (4.1%).

3.3.2 | Test curves

As mathematically expected, 100% sensitivity and 100% NPV for VNT corresponded to the same test value or cut-off. Thereafter, sensitivity and NPV progressed in the same decreasing trend, but their difference in values increased as a function of LSM increase. For example, in our previous VO-VCO population,[4] the 100% cut-off was 9.2 kPa for VCTE (Figure 1). The 95% sensitivity cut-off for VNT was 14.1 kPa (corresponding to a 98% NPV) and the 95% NPV cut-off for VNT was 26.5 kPa (corresponding to 92% sensitivity). The 90% sensitivity cut-off was 17.2 kPa, and the 90% NPV cut-off was 47.3 kPa.

3.4 | Case simulation

There were no parallel trends in clinical descriptors between the three calculation methods when the three simulation cases were

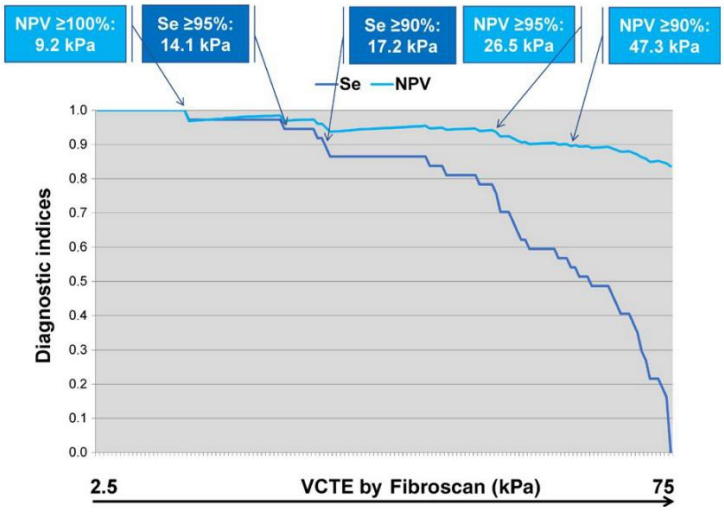


FIGURE 1 Diagnostic indices (Y axis) of LSM by VCTE (X axis) for VNT. Cut-offs of 100%, 95% and 90% sensitivity and negative predictive values are shown in the top boxes. These curves facilitate accurate cut-off determination. Unpublished data from previous VO-VCO study (218 patients with cirrhosis)[4]

compared (Figure S1). The whole population definition had a stable rate in most cases, whereas the two other definitions rarely varied in parallel. In a single additional case where the numbers of missed VNT and spared endoscopies were doubled, the missed VNT rate was paradoxically constant with the most frequent definition (spared endoscopy), whereas the raw number of missed VNT was doubled compared to the baseline case, leaving half of the concerned patients without expected primary prevention. Finally, the VNT definition was the most sensitive to variations in clinical descriptors (Figure S1).

4 | DISCUSSION

In the Baveno VI criteria, the most relevant clinical descriptor is spared endoscopy rate provided that the missed VNT rate is ≤5%.[1] Missed VNT describes false-negative cases and enables the calculation of either the NPV for VNT in patients with negative Baveno VI criteria (sparing endoscopy) or the sensitivity for VNT in patients with VNT. In other words, the maximum 5% missed VNT rate means either a minimum 95% sensitivity or a minimum 95% NPV for VNT. Sensitivity is calculated in the population with VNT, but it applies to the whole population. By contrast, the NPV is calculated in the population with spared endoscopy and it applies only to that subpopulation. Sensitivity is an intrinsic characteristic of the test and NPV an extrinsic characteristic of it depending on the sample, that is, the VNT prevalence. Thus, both have clear and separate meanings.

However, the most used definition, where spared endoscopy is the denominator, does not correctly reflect the benefit of Baveno VI criteria. For example, in simulation case 1, where VNT prevalence varied (Table S2), the clinical benefit also varied since the simulation shows a stable number of missed VNT despite a growing VNT prevalence. However, the calculated missed VNT rate remained constant. Thus, this spared endoscopy definition was somewhat insensitive to VNT prevalence, in a setting—clinical statistics—where predictive value is thought to be dependent on outcome prevalence. Another example is more eloquent: in the additional single simulation case 4 (Table S2) where the numbers of missed VNT and spared endoscopies had been doubled compared to the baseline conditions, the spared endoscopy definition gave a missed VNT rate similar to that of the baseline conditions (12.5%) despite the doubling of the number of missed VNT, leaving half of patients without expected primary prevention.

Despite its statistical advantage, the VNT definition (based on sensitivity) was more restrictive (and less optimistic) than the spared endoscopy definition. With the VNT definition, for example, the cut-off for LSM by VCTE was lower and consequently the spared endoscopy rate as well.

As concerns the whole population definition, our results showed that it was the least informative and most biased. It corresponds to accuracy which is not a pertinent diagnostic index in this clinical issue. Thus, this definition should be avoided.

The choice of definition also has important consequences on the determination of cut-offs in diagnostic tests, since the main result for the missed VNT rate should agree with the statistical method used to determine the cut-off. In other words, both metrics should be based on the same diagnostic index, that is, sensitivity or NPV. The example of LSM by VCTE clearly showed that the cut-offs for excluding a maximum of 5% VNT were 14.1 kPa for sensitivity and 26.5 kPa for NPV in our VO-VCO population[4] (Figure 1). This showed that the definition used for the missed VNT rate of ≤5% strongly influenced the cut-off determination of the diagnostic test used.

Finally, in our opinion, two definitions should be reported. Firstly, the VNT definition, since it provides the true proportion of missed VNT and in turn has epidemiological meaning (rate of patients left without prevention). The VNT definition corresponds to the sensitivity

of the Baveno VI criteria for VNT. In clinical studies, this definition also offers the advantage of a more powerful comparison between diagnostic tests than the spared endoscopy definition. However, this latter should also be reported secondarily, because it has the pragmatic advantage of corresponding to a clinical issue. Indeed, the physician wants to know the error risk, that is, the NPV for VNT in the population without endoscopy. For example, a 5% rate of missed VNT corresponds to a 95% NPV for VNT in patients with spared endoscopy; that corresponds probably the most to the clinical meaning of the Baveno VI criteria.[1,2] We do, however, privilege the first definition, especially for test construction. Statistical proposals for test evaluation in derivation and validation populations are summarized in the Table S4.

Thus, in conclusion, we believe that all upcoming studies evaluating endoscopy-sparing diagnostic tests or algorithms in the setting of VNT screening should take into account the VNT (np missed VNT/np VNT) and the spared endoscopy (np missed VNT/np spared endoscopy) definitions for calculating the missed VNT rate.

CONFLICT OF INTEREST

The authors do not have any disclosures to report.

ORCID

Paul Calès http://orcid.org/0000-0003-4866-5274

Federico Ravaioli http://orcid.org/0000-0002-1142-8585

REFERENCES

1. De Franchis R, Baveno VIF. Expanding consensus in portal hypertension: report of the Baveno VI Consensus Workshop: stratifying risk and individualizing care for portal hypertension. *J Hepatol*. 2015;63:743-752.
2. Augustin S, Pons M, Santos B, Ventura M, Genescà J. Identifying compensated advanced chronic liver disease: when (not) to start screening for varices and clinically significant portal hypertension. In deFranchis R, ed. *Portal Hypertension VI*. Proceedings of the Sixth Baveno Consensus Workshop: Stratifying Risk and Individualizing Care. Berlin, Germany: Springer;2016:39-49.
3. Bae J, Sinn DH, Kang W, et al. Validation of the Baveno VI and the expanded Baveno VI criteria to identify patients who could avoid screening endoscopy. *Liver Int*. 2018;38:1442-1448.
4. 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.
5. Colecchia A, Ravaioli F, Marasco G, et al. A combined model based on spleen stiffness measurement and Baveno VI criteria to rule out high-risk varices in advanced chronic liver disease. *J Hepatol*. 2018;69:308-317.
6. 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.
7. 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;37:1177-1183.
8. Roccarina D, Rosselli M, Genesca J, Tsochatzis EA. Elastography methods for the non-invasive assessment of portal hypertension. *Expert Rev Gastroenterol Hepatol*. 2018;12:155-164.
9. 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(4):878-885.

SUPPORTING INFORMATION

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

How to cite this article: Calès P, Buisson F, Ravaioli F, et al. How to clarify the Baveno VI criteria for ruling out varices needing treatment by noninvasive tests. *Liver Int*. 2019;39:49–53. https://doi.org/10.1111/liv.13945

BUISSON François

Le ratio plaquettes/élastométrie hépatique personnalisé améliore la sûreté et l'efficacité du dépistage des varices oesophagiennes à risque de saignement

RÉSUMÉ

Introduction : Basés sur le taux de plaquettes et l'élastométrie hépatique (LSM by vibration-controlled transient elastography), les critères de Baveno VI (B6C), les critères de Baveno VI étendus (EB6C) ou le score ANTICIPATE peuvent être utilisés pour éliminer la présence de varices à traiter (VNT). L'objectif était d'évaluer et d'améliorer ces tests.

Sujets et Méthodes : 2368 patients avec une hépatopathie chronique d'étiologie et de sévérité variées ont été randomisés en cohorte de dérivation (n=1579) et de validation (n=789) au sein d'une étude multicentrique rétro-prospective. Les tests publiés étaient comparés à deux nouveaux tests : PLER (platelets/LSM ratio) et PLESIR (Platelets/LSM, adjusted on Etiology/Sex/INR, ratio).

Résultats : les caractéristiques des patients de la population de dérivation étaient : VNT: 15.1%, étiologies : virus: 50.2%, NAFLD: 28.9%, alcool: 20.8%, MELD score: 9.5±3.0, LSM≥10kPa: 93.0%. Les taux d'endoscopies évitées (avec taux de VNT manquée entre parenthèse pour sureté d'usage) étaient en ordre croissant : B6C : 23.9% (2.9%), ANTICIPATE : 24.3% (4.6%), PLER: 26.6% (4.6%), PLESIR : 34.8% (3.3%) et EB6C : 41.9% (10.9%). Les différences entre les taux d'endoscopies évitées étaient significatives entre les tests ($p \leq 0.001$) sauf pour B6C vs ANTICIPATE. Le taux de VNT manquées était différent seulement entre EB6C vs autres tests ($p \leq 0.009$). La performance et la sureté d'emploi des tests n'étaient pas significativement différentes entre les populations. PLESIR était le seul test avec un taux de VNT manquées ≤5% quel que soit le sexe ou l'étiologie. Un algorithme VariScreen, basé successivement sur les plaquettes ou LSM puis plaquettes/LSM puis PLESIR chez 35% des patients, sécurisait totalement le dépistage (pas de VNT manquées chez les hépatopathies avancées). Le taux d'endoscopies évitées était de 35,7%, le taux de VNT manquées de 3,3%

Conclusion: L'usage de B6C est sûr par rapport aux VNT manquées, qu'importent le sexe, l'étiologie ou la sévérité de l'atteinte hépatique ; Le taux de VNT manquées de EB6C est trop élevé pour qu'il soit recommandé à l'usage. Pour améliorer le dépistage des VNT, nous proposons l'usage de l'algorithme séquentiel VariScreen, applicable à toute pathologie hépatique chronique.

Mots-clés : hypertension portale ; varices oesophagiennes ; diagnostic non-invasif ; critères de Baveno VI ; élastographie impulsionnelle à vibration contrôlée

Personalized platelets/liver stiffness ratio improves and secures the screening of esophageal varices needing treatment

ABSTRACT

Background & Aims: Based on platelets and liver stiffness measurement (LSM by vibration-controlled transient elastography), the Baveno VI criteria (B6C), the expanded B6 criteria (EB6C) or the ANTICIPATE score can be used to rule out varices needing treatment (VNT). We aimed to evaluate and improve these tests.

Methods: 2368 patients were randomized in derivation (n=1579) and validation (n=789) populations with chronic liver disease of various etiologies and severity in a multicenter retro-prospective study. Published tests were compared to two new tests: PLER (platelets/LSM ratio) and PLESIR (Platelets/LSM, adjusted on Etiology/Sex/INR, ratio).

Results: Patient characteristics were in the derivation population: VNT: 15.1%, etiologies: virus: 50.2%, NAFLD: 28.9%, alcohol: 20.8%, MELD score: 9.5±3.0, LSM≥10kPa: 93.0%. Spared endoscopy rates (with missed VNT rates in parentheses for safety) were in increasing order: B6C: 23.9% (2.9%), ANTICIPATE: 24.3% (4.6%), PLER: 26.6% (4.6%), PLESIR: 34.8% (3.3%) and EB6C: 41.9% (10.9%). Differences in spared endoscopy rates were significant between tests ($p \leq 0.001$) except for B6C vs ANTICIPATE and for missed VNT rate only between EB6C vs others ($p \leq 0.009$). The test performance and safety were not significantly different between populations. PLESIR was the only safe test (missed VNT ≤5%) whatever sex and etiology. A VariScreen algorithm, based successively on platelets or LSM then platelets/LSM then PLESIR in 35% of patients, secured screening (no missed VNT in poor liver function); spared endoscopy rate was 35.7% and missed VNT rate was 3.3%.

Conclusions: B6C are safe for missed VNT rate regardless of CLD etiology and severity, and sex; EB6C are no longer recommended as unsafe. To improve the current VNT screening, we propose the sequential VariScreen algorithm applicable to any CLD of main etiologies.

Keywords : portal hypertension; esophageal varices; non-invasive diagnosis; Baveno VI criteria, vibration-controlled transient elastography,

FACULTÉ DE SANTÉ
UNIVERSITÉ D'ANGERS