2020-2021

THÈSE

pour le

DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE

Qualification en hépato-gastro-entérologie

Individual and population screening of varices needing treatment by a simple, safe and accurate test : LIP

Dépistage individuel et en population des varices nécessitant un traitement par un test simple, sûr et performant : LIP

CANDILLIER Camille

Née le 16/03/1993 à Seclin (59)

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 : 07/09/2021

FACULTÉ DE SANTÉ

UNIVERSITÉ D'ANGERS

ENGAGEMENT DE NON PLAGIAT

Je, soussignée CANDILLIER Camille
déclare être pleinement consciente 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'étudiante le 07/07/2021

LISTE DES ENSEIGNANTS DE LA FACULTÉ DE SANTÉ D'ANGERS

Doyen de la Faculté : Pr Nicolas Lerolle
Vice-Doyen de la Faculté et directeur du département de pharmacie : Pr Frédéric Lagarce
Directeur du département de médecine : Pr Cédric Annweiler

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
BAUFRETON Christophe	Chirurgie thoracique et cardiovasculaire	Médecine
BELLANGER William	Médecine Générale	Médecine
BENOIT Jean-Pierre	Pharmacotechnie	Pharmacie
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
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
COPIN Marie-Christine	Anatomie et cytologie pathologiques	Médecine
COUTANT Régis	Pédiatrie	Médecine
CUSTAUD Marc-Antoine	Physiologie	Médecine
DE BRUX Jean-Louis	Chirurgie thoracique et cardiovasculaire	Médecine
DE CASABIANCA Catherine	Médecine Générale	Médecine
DESCAMPS Philippe	Gynécologie-obstétrique	Médecine
D'ESCATHA Alexis	Médecine et santé au travail	Médecine
DINOMAIS Mickaël	Médecine physique et de réadaptation	Médecine
DIQUET Bertrand	Pharmacologie	Médecine

DUBEE Vincent	Maladies Infectieuses et Tropicales	Médecine
DUCANCELLE Alexandra	Bactériologie-virologie ; hygiène hospitalière	Médecine
DUVAL Olivier	Chimie thérapeutique	Pharmacie
DUVERGER Philippe	Pédopsychiatrie	Médecine
EVEILLARD Mathieu	Bactériologie-virologie	Pharmacie
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
GUARDIOLA Philippe	Hématologie ; transfusion	Médecine
GUILET David	Chimie analytique	Pharmacie
HAMY Antoine	Chirurgie générale	Médecine
HENNI Samir	Médecine Vasculaire	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
LEGENDRE Guillaume	Gynécologie-obstétrique	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
MAY-PANLOUP Pascale	Biologie et médecine du développement et de la reproduction	Médecine
MENEI Philippe	Neurochirurgie	Médecine
MERCAT Alain	Réanimation	Médecine
PAPON Nicolas	Parasitologie et mycologie médicale	Pharmacie
PASSIRANI Catherine	Chimie générale	Pharmacie
PELLIER Isabelle	Pédiatrie	Médecine
PETIT Audrey	Médecine et Santé au Travail	Médecine
PICQUET Jean	Chirurgie vasculaire ; médecine vasculaire	Médecine
PODEVIN Guillaume	Chirurgie infantile	Médecine
PROCACCIO Vincent	Génétique	Médecine
PRUNIER Delphine	Biochimie et Biologie Moléculaire	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
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
SAULNIER Patrick	Biophysique et Biostatistiques	Pharmacie
SERAPHIN Denis	Chimie organique	Pharmacie
SCHMIDT Aline	Hématologie ; transfusion	Médecine
TRZEPIZUR Wojciech	Pneumologie	Médecine
UGO Valérie	Hématologie ; transfusion	Médecine
URBAN Thierry	Pneumologie	Médecine
VAN BOGAERT Patrick	Pédiatrie	Médecine
VENIER-JULIENNE Marie-Claire	Pharmacotechnie	Pharmacie
VERNY Christophe	Neurologie	Médecine
WILLOTEAUX Serge	Radiologie et imagerie médicale	Médecine

MAÎTRES DE CONFÉRENCES

ANGOULVANT Cécile	Médecine Générale	Médecine
BAGLIN Isabelle	Chimie thérapeutique	Pharmacie
BASTIAT Guillaume	Biophysique et Biostatistiques	Pharmacie
BEAUVILLAIN Céline	Immunologie	Médecine
BEGUE Cyril	Médecine générale	Médecine
BELIZNA Cristina	Médecine interne	Médecine
BELONCLE François	Réanimation	Médecine
BENOIT Jacqueline	Pharmacologie	Pharmacie
BIERE Loïc	Cardiologie	Médecine
BLANCHET Odile	Hématologie ; transfusion	Médecine
BOISARD Séverine	Chimie analytique	Pharmacie
BRIET Claire	Endocrinologie, Diabète et maladies métaboliques	Médecine
BRIS Céline	Biochimie et biologie moléculaire	Pharmacie
CAPITAIN Olivier	Cancérologie ; radiothérapie	Médecine
CASSEREAU Julien	Neurologie	Médecine
CHEVALIER Sylvie	Biologie cellulaire	Médecine
CLERE Nicolas	Pharmacologie / physiologie	Pharmacie

COLIN Estelle	Génétique	Médecine
DERBRE Séverine	Pharmacognosie	Pharmacie
DESHAYES Caroline	Bactériologie virologie	Pharmacie
FERRE Marc	Biologie moléculaire	Médecine
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
JUDALET-ILLAND Ghislaine	Médecine générale	Médecine
KHIATI Salim	Biochimie et biologie moléculaire	Médecine
KUN-DARBOIS Daniel	Chirurgie maxillo-faciale et stomatologie	Médecine
LACOEUILLE Franck		Pharmacie
LANDREAU Anne	Botanique/ Mycologie	Pharmacie
LEBDAI Souhil	Urologie	Médecine
LEGEAY Samuel	Pharmacocinétique	Pharmacie
LEMEE Jean-Michel	Neurochirurgie	Médecine
LE RAY-RICHOMME Anne-Marie	Pharmacognosie	Pharmacie
LEPELTIER Elise	Chimie générale	Pharmacie
LETOURNEL Franck	Biologie cellulaire	Médecine
LIBOUBAN Hélène	Histologie	Médecine
LUQUE PAZ Damien	Hématologie biologique	Médecine
MABILLEAU Guillaume	Histologie, embryologie et cytogénétique	Médecine
MALLET Sabine	Chimie Analytique	Pharmacie
MAROT Agnès	Parasitologie et mycologie médicale	Pharmacie
MESLIER Nicole	Physiologie	Médecine
MIOT Charline	Immunologie	Médecine
MOUILLIE Jean-Marc	Philosophie	Médecine
NAIL BILLAUD Sandrine	Immunologie	Pharmacie
PAILHORIES Hélène	Bactériologie-virologie	Médecine
PAPON Xavier	Anatomie	Médecine
PASCO-PAPON Anne	Radiologie et imagerie médicale	Médecine
PECH Brigitte	Pharmacotechnie	Pharmacie
PENCHAUD Anne-Laurence	Sociologie	Médecine
PIHET Marc	Parasitologie et mycologie	Médecine
PY Thibaut	Médecine Générale	Médecine
RAMOND-ROQUIN Aline	Médecine Générale	Médecine
RINEAU Emmanuel	Anesthésiologie réanimation	Médecine
RIOU Jérémie	Biostatistiques	Pharmacie
ROGER Emilie	Pharmacotechnie	Pharmacie
SAVARY Camille	Pharmacologie-Toxicologie	Pharmacie
SAVARY Dominique	Médecine d'urgence	Médecine
SCHMITT Françoise	Chirurgie infantile	Médecine
SCHINKOWITZ Andréas	Pharmacognosie	Pharmacie

SPIESSER-ROBELET Laurence	Pharmacie Clinique et Education Thérapeutique	Pharmacie
TESSIER-CAZENEUVE Christine	Médecine Générale	Médecine
TEXIER-LEGENDRE Gaëlle	Médecine Générale	Médecine
VIAULT Guillaume	Chimie organique	Pharmacie

AUTRES ENSEIGNANTS

PRCE

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

PAST

CAVAILLON Pascal	Pharmacie Industrielle	Pharmacie
DILÉ Nathalie	Officine	Pharmacie
MOAL Frédéric	Pharmacie clinique	Pharmacie
PAPIN-PUREN Claire	Officine	Pharmacie
POIROUX Laurent	Soins Infirmiers	Médecine

ATER

BOUCHENAKI Hichem	Physiologie	Pharmacie
MESSAOUDI kHALED	Immunologie	Pharmacie
MOUHAJIR Abdelmounaim	Biotechnologie	Pharmacie

PLP

CHIKH Yamina	Economie-gestion	Médecine

AHU

IFRAH Amélie	Droit de la Santé	Pharmacie
LEBRETON Vincent	Pharmacotechnie	Pharmacie

Mise à jour 27/08/2020

REMERCIEMENTS

Au Professeur CALES de m'avoir encadrée et aidée au cours de cette thèse et de l'internat.

Au Professeur CAROLI BOSC de me faire l'honneur de présider ce jury ainsi que pour votre enseignement au cours de l'internat.

Au Professeur BOURSIER pour avoir accepté de participer et juger cette thèse ainsi que pour ton enseignement au cours de l'internat.

Au Docteur OBERTI, pour avoir accepté de participer et juger cette thèse. Mais également m'avoir enseigné l'endoscopie au cours des astreintes ainsi pour votre encadrement et enseignement.

A l'équipe médicale du CHU d'Angers et du CH de Cholet, pour le temps que vous m'avez accordé dans les services, en endoscopie ou en dehors et pour tout ce que vous m'avez appris. Et plus particulièrement merci à Adrien, Carole et Clémence pour votre soutien.

A l'équipe médicale et paramédicale du service de réanimation du Mans pour m'avoir permis d'élargir mes connaissances en me transmettant quelques clés de votre spécialité.

A l'équipe paramédicale du CHU d'Angers pour votre accompagnement et votre bonne humeur au quotidien, ce fut un réel plaisir.

A tous mes co-internes de gastro, et surtout Hélène, Isabelle, Norma, Samuel et Tiphaine pour tous les bons moments que nous avons partagés et pour votre aide. Vous avez rendu ces 4 années bien plus belles.
A Claire Boyer, pour ton aide mon premier jour au CHU, tout le reste de mon internat et pour longtemps encore je l'espère… et bien sûr pour les plantes et les jeux.

A mes co-internes de réanimation, de principe vous avez rendu ce semestre inoubliable et merci de votre soutien pour cette dernière ligne droite.

A Camille et Louise, Franck, Valentin, Fanny qui me supportent depuis tant d'années, pour les heures passées à travailler et surtout toutes les autres. A Flora pour ton soutien malgré la distance et le décalage horaire.
A Céline pour parmi tant de chose, tes « t'es une winneuse » avant chaque examen.

A Eloïse, pour toutes ces « réunions de concertation pluridisciplinaires » à 2, pour ton aide, pour le potassium, pour les fous rires, pour m'avoir fait prendre goût à la colocation et tout le reste.

A mes grands-parents qui ont toujours crus en moi, pour votre fierté dans les yeux.
A Mes parents, pour tout, je ne vous remercierai jamais assez.
A FX pour ton soutien tout le long de mes études et bien avant.
A Mathilde, qui a toujours été là pour moi, pour ton aide, ta présence, tes coups de téléphone, tes photos, tes chansons, tes livres, nos voyages…

A Charles, pour ton soutien infaillible, tes petits mots et tes conseils (que j'ai plus ou moins écoutés…). Pour ces années passées avec toi et celles à venir.

Abbreviations

ALD :	Alcoholic liver disease
AUROC :	Area under the receiver operating characteristic
cALD	Compensated advanced chronic liver disease
CI	Confidence interval
CLD	Chronic liver disease
EV	Oesophageal varices
INR	International normalized ratio
LR-	Negative likelihood ratio
LSM	Liver stiffness measurement
MELD	Model for end-stage liver disease
NAFLD	Non-alcoholic fatty liver disease
NPV	Negative predictive value
PI	Prothrombin index
PPV	Positive predictive value
VCTE	Vibration-controlled transient elastometry
VNT	Varices needing treatment

Plan

Individual and population screening of varices needing treatment by a simple, safe and accurate test : LIP.

Federico Ravaioli,1* Arthur Berger,2* Oana Farcau,3 Davide Festi,1 Horia Stefanescu,3 Camille Candillier,2 Pierre Nahon,4 Christophe Bureau,5 Nathalie Ganne-Carrià,4 Annalisa Berzigotti,6 Victor de Ledinghen,7 Salvatore Petta,8 Paul Calès,2 and multicentric groups9: VO-VCO, ANRS CO12 CIRVIR, M116 and VEB6. Angers,1 Bologna,2 Cluj,3 Bondy,4 Toulouse,5 Bern,6 Bordeaux,7 Palermo.8

Authors' affiliations:

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

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

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

4 Hepatology Department, Paris Seine-Saint-Denis University Hospitals, APHP, Avicenne Hospital, Bobigny; UFR SMBH, University of Paris 13, Sorbonne Paris Cité; Cordeliers Research Center, INSERM, Sorbonne University, University of Paris, FunGeST, F-75006 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.

Financial support statement:

Institutional support from participating centers, Echosens (M116), French PHRC (VO-VCO) and ANRS (FRENSH: France Recherche Nord & Sud Sida-HIV Hépatites, CO12 CIRVIR), which granted original subpopulations with no other role in the present study.

ABSTRACT

Background and Aims : Several tests have been developed to screen varices needing treatment (VNT) but for different screening strategies: individual or population. We aimed to develop simple estimators to quantify VNT risk and to spare endoscopy, while missing <5% of VNT, adapted to different screenings in the main aetiologies.

Methods : 2,368 patients with chronic liver disease were included in derivation and validation sets. Results, being similar therein, are given in the whole population. Patient characteristics were, age: 59±11 years; male sex: 63.5%; aetiologies: virus: 50.2%, NAFLD: 28.9%, alcohol: 20.9%; MELD: 9.5±3.0; liver stiffness measurement (LSM) ≥10kPa: 92.8%; VNT: 15.2%. The main independent VNT predictors were platelets, prothrombin index (PI) and LSM. Their interactions led to score construction, LIP: (LSM*45)/(PI*platelets), and BLIP: BMI-adjusted LIP in NAFLD. Scores were categorised either for population (VNT sensitivity ≥95%) or individual (negative predictive value ≥95%) VNT screening. The performance criterion was spared endoscopy rate.

Results : 1) Scores diagnosing VNT. AUROCs were, PLER: 0.767 Anticipate: 0.773 (p=0.059 vs previous), LIP: 0.779 (p=0.136), PLEASE: 0.789 (p=0.196). 2) Population screening performance was in increasing order (with missed VNT rates), Baveno$_6$ criteria: 23.9% (2.5%), Anticipate: 24.5%, p=0.367 vs previous (3.3%), PLER: 27.3%, p<0.001 (3.6%), LIP: 33.4%, p<0.001 (4.2%), PLEASE: 35.2%, p=0.006 (3.6%). In NAFLD, performance was, LIP: 38.6%, BLIP: 40.8%, p=0.038. 3) Individual screening performance was, LIP: 54.1% expanded Baveno$_6$ criteria: 42.7%, p<0.001. In NAFLD, performance was, BLIP: 74.6%, NAFLD-cirrhosis criteria: 66.7%, p<0.001.

Conclusion: LIP combined simplicity, performance and safety in each aetiology. In NAFLD, BLIP outperformed other tests.

Words : 250

Keywords : Portal hypertension ; œsophageal varices ; liver elastometry ; prothrombin time

INTRODUCTION

The original Baveno VI criteria enabled the wide clinical acceptance of non-invasive tests for varices needing treatment (VNT). The Baveno VI criteria aimed at ruling out VNT are based on platelets and liver stiffness measurement (LSM)[1] assessed by vibration-controlled transient elastography (VCTE). The missed VNT rate must be <5%. Baveno VI criteria have several limitations. First, although largely validated[2-6], their clinical impact has been judged modest, providing a spared esophagogastroduodenoscopy (endoscopy hereafter) rate of only about 20%[3, 7]. Therefore, several authors have proposed improvements[8-11]. Second, in the literature, two main definitions of missed VNT are used[12]. One is based on the probability of missing <5% VNT in patients left without endoscopy. This is a negative predictive value (NPV) adapted to a clinical decision in a patient, i.e. an individual screening strategy. Thus, recent tests, like the expanded Baveno VI criteria[9] and NAFLD cirrhosis criteria[13], were constructed only for individual screening. Another definition is based on the probability of missing <5% VNT in patients with VNT. This corresponds to VNT sensitivity and thus is adapted to a public health decision, i.e. a population screening strategy. Third, the Baveno VI criteria were originally applied to patients with compensated advanced chronic liver disease (cACLD) since the rate of unnecessary endoscopies is high in that setting. However, we have recently shown that this selective strategy restricted to cACLD spared less endoscopies than a global strategy unrestricted by liver severity[14]. Nonetheless, that strategy must be secured, that is, designed to avoid missing VNT in severe CLD. Fourth, Baveno VI criteria and their derivatives do not estimate VNT probability. Recently, we proposed two new scores quantifying VNT probability and categorized in tests to spare endoscopy [14]. First, the platelets/liver elastometry ratio, called PLER score, and the platelets/liver elastometry ratio adjusted on aetiology, sex and INR called PLEASE. PLER is a simple test that performs better than the Baveno VI criteria, with a 27% spared endoscopy rate. The PLEASE test performs better still, offering a 35% spared endoscopy rate, but its complex formula requires a specific calculator. However, neither of those tests explicitly quantify VNT risk. Yet, having knowledge on any increased risk of VNT is beneficial for the physician/patient decision making process, wherein patients are often somewhat reluctant to undergo endoscopy. Thus, risk prediction might help reduce

the significant deficit in VNT surveillance, which impacts mortality in cirrhosis[15]. Finally, our previous work showed that other published tests, like the Baveno VI criteria, Anticipate and the expanded Baveno VI criteria, had limits in non-viral aetiologies.

NAFLD is a noteworthy and increasingly predominant aetiology. However, previous VNT tests in NAFLD were limited: tests constructed for population screening performed poorly whereas NAFLD cirrhosis criteria provided better performance as they were constructed only for individual screening. We observed however a particular role of platelets and BMI for VNT prediction in a previous study in NAFLD[16].

With the present work, our main objective was to develop a simple, bedside test for all main aetiologies of CLD by combining the advantages of previous tests and avoiding their limits. Our secondary objectives were to a) design this test for individual and population screenings, b) quantify the VNT risk, and to refine safety criteria for missed VNT and c) adapt this test to NAFLD.

PATIENTS AND METHODS

1. Participants

In this post-hoc analysis of prospectively collected data, the clinical information of patients with CLD was collected from centres participating in several studies wherein VNT was usually the main outcome and VCTE the measurement outcome. The protocol conformed to the Declaration of Helsinki and received approval from the ethics review boards of all participating centres. All study participants gave informed consent. Patients included in previously recorded CLD subpopulations of any main aetiology (alcoholic CLD (ALD), NAFLD, hepatitis B or C virus) were eligible for inclusion if they had undergone an endoscopy to determine oesophageal varice (EV) size. A platelet count, successful LSM by VCTE using the M probe, known EV stage and a maximum delay of six months between endoscopy and LSM or platelets were the four minimum inclusion criteria. Interventional treatment for portal hypertension (PHT) complications (TIPS, band ligation or sclerotherapy of EV) and incomplete data were exclusion criteria. Also, patients were included irrespective of LSM values and liver severity (i.e. non limited to cACLD) to enable a less biased analysis of the VNT subset. Of the 4132 patients across 47 centres (details in[14]) eligible for the study, 2368 were finally included in the present core population (Figure S1 in Supplementary Material). The included patients were randomised in derivation (2/3) and validation (1/3) populations with stratification on VNT and aetiology.

2. Methods

2.1. Data collection

Clinical data - The main clinical data were age, sex, height, body weight and CLD aetiology. The main laboratory data were liver function tests, blood cell count and serum creatinine (measured in each centre). The model for end-stage liver disease (MELD) score included bilirubin, the international normalised ratio (INR) and creatinine[17].

Endoscopy - Experienced operators performed a standard endoscopy and EV grades were recorded.

LSM - All LSMs were performed by experienced operators using VCTE, specifically M probe-equipped Fibroscan devices (Echosens, Paris, France). Technical characteristics are detailed elsewhere[18].

2.2. Definitions

Objectives - The primary objective was to develop a simple, bedside test for all the main aetiologies of CLD. This new test needed to match the performance and safety of previously published tests, and especially match the secureness of our previous PLEASE test.

The secondary objectives were: to dispose of a test able to quantify VNT risk, especially in the indeterminate test subset where missed VNT rates is ≥5%; to determine cut-offs for individual and population screenings; and to adapt the test to the particularities of NAFLD. Ancillary objectives were to reach 100% specificity for VNT and to develop several qualitative safety criteria for the missed VNT.

Outcome - The main outcome was VNT, defined as large EV (grade 2 or 3, i.e. a diameter ≥5 mm[19]).

Outcome measurements - The primary outcome measurements were the performance (spared endoscopy rate) and safety (missed VNT rate) of tests for VNT. The spared endoscopy rate was calculated as the ratio between the number of patients with a missed VNT rate <5% by test and the total number of patients. This corresponded to the negative test rate and was dependent on VNT prevalence. The missed VNT rate was the ratio of missed VNT on VNT in population screening or spared endoscopies in individual screening[12]. Considering the missed rate of <5%, the first definition corresponds to ≥95% sensitivity. The second definition, corresponding to ≥95% NPV, has been used for some published tests[9, 13]. The negative likelihood ratio (LR-) was a secondary outcome measurement (details in Supplementary Material). The LR- expresses how many times less likely a negative test result (spared endoscopy) is to be expected in patients with VNT as compared to patients without VNT. When <0.1, LR- is considered excellent[20]. However, LR- is not suited to individual screening. Metavir fibrosis stages were estimated by LSM[21].

Safety criteria - We evaluated three criteria. First, the classical *quantitative* criterion is a missed VNT rate <5%, the calculations and populations of which differ between individual and population screening. Second, we evaluated the level of liver dysfunction as a function of missed VNT

(*qualitative safety*). The principle of the two new criteria of *qualitative safety* is to privilege the test having the lowest liver dysfunction in missed VNT and to discard (or limit) a test inducing missed VNT in severe CLD. Indeed, the incidence and mortality of variceal bleeding grows with liver severity[22]. Thus, a test was called *secured* when the missed VNT rate was 0% in CLD with poor liver function (MELD score ≥10 or INR ≥1.24)[14] called severe CLD hereafter. Moreover, we evaluated the level of liver dysfunction in missed VNT called *functional safety*.

Comparators - The new tests were compared first for population screening. For this setting, they were compared to the main published tests: Baveno VI criteria[1] based on cut-offs of platelets and LSM; Anticipate (a logit function of platelets and LSM)[23]; and PLER (platelets/LSM ratio) and PLEASE (PLER adjusted to INR, sex and aetiology)[14]. Then, they were compared in the individual screening setting with the expanded Baveno VI criteria[9] and the NAFLD cirrhosis criteria[13].

VNT diagnostic estimators - An estimator was called a *score* when it provided a numerical score (i.e. a continuous variable quantifying the VNT risk). Scores were arithmetic combinations of variables. An estimator was called a *test* when it was categorised by cut-off(s), resulting in a qualitative variable indicating the VNT categories.

VNT screening strategies - The characteristics of these screening strategies are summarised in Table 1.

2.3. Score development

The predefined strategy resulted from our previous studies showing the contribution of interaction terms[14, 16]. Details are provided in the Supplemental Material.

All aetiologies - In multivariate analysis, PI, LSM and platelets were the most predictive of VNT (Table S1). There were several interactions between these markers. Therefore, an interaction term including these three markers was the most predictive. Thus, following the previous determination of the LSM/platelets ratio[14], we selected a simple ratio LSM/(PI*platelets). Indeed, this ratio provided the highest area under the receiver operating characteristic (AUROC) for VNT and largest subset with ≥95% sensitivity for VNT. The LSM/(PI*platelets) ratio correlated with VNT prevalence but its values were different from it. Therefore, this score was simply calibrated by adding a constant -45- in the

ratio numerator. Then, the upper values of the ratio were capped at 0.6 considering their far-off values beyond the maximum observed VNT probability. Thus, the final (LSM*45)/(PI*platelets) score, called LIP_{PI} and simply LIP hereafter, ranged from 0 to 0.6, with 0.6 expressing the maximum VNT probability. LIP distribution as a function of VNT is shown in Figure S2. We describe the corresponding LIP_{INR} score, using INR instead of PI, in the Supplemental Material.

NAFLD - This aetiology had several particularities (Table S2). First, BMI was significantly increased as expected. Second, platelet count was significantly increased. This led us to evaluate a relationship between those two variables. In fact, NAFLD was the only aetiology where BMI and platelets were positively correlated in patients without VNT (Figure S3A). Moreover, they were negatively correlated in patients with VNT. Therefore, BMI amplified VNT discrimination by platelets and LIP (Figure S3B). BMI significantly influenced the relationship between VNT and platelets or LSM but not INR in NAFLD (Figure S4). VNT discrimination by the platelets/LSM ratio was the most amplified by PI (or INR) in NAFLD (Figure 1A) and discrimination by LIP was amplified by BMI (Figure 1B). Therefore, we added BMI in the LIP formula to obtain the BLIP score: (LSM*45*30)/(PI*platelets*BMI). Finally, BLIP had a larger subset with ≥95% VNT sensitivity than LIP and reached 100% specificity (Figure S5).

Details on formulas of LIP and BLIP and their score cut-offs for individual and population screenings, adjusted on aetiology, are provided in the Supplemental Material.

2.4. Statistics

Quantitative variables were expressed as mean ± standard deviation and compared using the Student t-test or analysis of variance (ANOVA). Qualitative variables were expressed as proportions and compared using the Chi^2 test or Fisher test when unpaired and the Cochran or McNemar test when paired. Correlations were measured by the non-parametric Spearman correlation coefficient (r_s) and/or parametric Pearson correlation coefficient (r_p) when necessary. Independent VNT predictors were determined by forward binary logistic regression. In the next step, we systematically tested interactions between the three main predictors: platelets, LSM and PI. Models with variable collinearity ($r>0.8$) were excluded. Data were reported according to STARD[24] and Liver FibroSTARD[25] statements and analysed on a partial intention-to-diagnose basis. Thus, all patients were included

irrespective of reliability criteria of VCTE[26] (except in one NAFLD subpopulation[13]) but missing data were not replaced and patients with unsuccessful examinations (LSM and endoscopy) were not included. Test performance and safety were internally validated in the validation set and through a 95% confidence interval (CI) obtained by bootstrap on 1000 samples in the whole population. Thus, this was a TRIPOD 2a study[27]. The main statistical analyses were performed using SPSS version 18.0 (IBM, Armonk, NY, USA).

RESULTS

1. Patients

The study included 2368 CLD patients (Table S3). Because there were no significant differences between the characteristics of derivation and validation populations, the results that follow are those of the whole population. Nearly two-thirds of the patients were men. Viral CLD was the most frequent aetiology (50%), then NAFLD (29%) and ALD (21%). LSMs ≥10 kPa were observed in 92.8% of patients. Severe fibrosis (equivalent to Metavir F3 or F4) was observed in 89.7% of patients, 20.2% of whom had early cirrhosis and 40.6% definitive cirrhosis. Thus, 71.0% of VNT were observed in definitive cirrhosis, 15.6% in early cirrhosis and 11.1% in severe fibrosis as estimated by the LSM classification[21]. The characteristics as a function of aetiologies are reported in Table S2.

2. Scores estimating VNT probability

2.1. Score calibration

The LIP score was well calibrated to VNT prevalence as demonstrated by two criteria: the estimated mean prevalence was not significantly different from the VNT prevalence observed by endoscopy (14.8% vs 15.2%, p=0.566); and scatter plots of LIP and VNT as a function of LIP percentiles clearly showed the high correlation between estimated and observed prevalences (Figure 2).

2.2. Score discrimination

VNT discrimination by scores was evaluated with AUROC. Figure 3A clearly shows that the AUROC of the LIP score was significantly higher ($p<0.001$) than those of its composite markers (details in Table S4). In the derivation population, the AUROC of the LIP score was significantly higher than that of PLER ($p<0.001$) and lower than that of PLEASE ($p=0.015$) but not significantly different from Anticipate (Table 2).

3. Tests to spare endoscopy

3.1. New test

Figure 3A clearly shows that the LIP score provided a much larger subset of patients with VNT sensitivity ≥95%, which corresponds to the spared endoscopy rate, than its composite markers did. The performance and safety of LIP are described in Table S5. Briefly, considering population (VNT ≥95% sensitivity) and individual (NPV ≥95% sensitivity) screenings, the performance and safety of LIP were not significantly different between the derivation and validation populations. Therefore, the following results are presented in the whole population.

3.2. Comparison of tests

Figure 3B shows that the LIP score provided a larger subset of patients with VNT sensitivity >95% than the published tests did. This conferred an endoscopy sparing advantage for the LIP score compared to other scores. Indeed, this characteristic of LIP was more marked than those of other scores, in contrast with the lesser difference in AUROC, especially vs the PLEASE score.

Population screening - The spared endoscopy rates were, in increasing order: Baveno VI criteria: 23.9%, Anticipate: 24.5% (p=0.367 vs previous), PLER: 27.3% (p<0.001), LIP: 33.4% (p<0.001), PLEASE: 35.2% (p=0.006) (Table 3). Here, PLEASE test became superior to LIP because the former is an algorithm stratified on sex and aetiology contrary to the latter. Performance comparisons (p values) are detailed in Table S6. All tests were safe with missed VNT rates <5% (non-significant differences). LR- was excellent (around 0.1) for all tests (Table 3).

Individual screening - The spared endoscopy rates were, in increasing order: expanded Baveno VI criteria: 42.7%, LIP: 54.1% (p<0.001) (Table 4). The missed VNT rates were, respectively: 4.0 vs 5.2% (p=0.175). However, this putative missed VNT rate at the limit of safety for LIP, due to a higher rate in the validation population (Table S5), was eliminated by the secureness rule (see below).

3.3. Sensitivity analysis

Influence of liver dysfunction on performance and quantitative safety

cACLD - Tests were evaluated in the cACLD population for population screening (Table S7). The performance in the cACLD population was slightly lower (from -3.3% to -5.2%) than in the whole population while all tests remained safe. For example, LIP performance decreased from 33.4% to 28.7% ($p<0.001$).

Population screening - Test performance linearly decreased as a function of MELD deciles from around 70% to around 5% in the most severe CLD (Figure S6A). For all tests, a clear trend of decreasing missed VNT rates from around 12% to around 0% as a function of MELD deciles was observed (Figure S6B). Thus, the test safety increased as a function of growing liver dysfunction.

Individual screening – Here too, test performance linearly decreased as a function of MELD deciles from around 90% to around 5% in the most severe CLD (Figure S7A). One noteworthy difference with population screening was the growing missed VNT rate as a function of MELD deciles (Figure S7B), which was incompatible with the secureness rule.

Qualitative safety analysis

Secureness - In population screening, secureness, defined as no missed VNT in MELD score greater than or equal to ten[14], was excellent with Baveno VI criteria and LIP (Figure S6B). Secureness was almost excellent with PLER, PLEASE, and Anticipate, where there were one or two patients with missed VNT by MELD score ≥10. In individual screening, the expanded Baveno VI criteria and LIP were not secured. Indeed, the missed VNT rate was increased in MELD score ≥10, more markedly for the expanded Baveno VI criteria than for LIP (Figure S7B). Considering that MELD score ≥10 was a limit for the application of LIP to individual screening, we restricted the use of LIP to patients with MELD scores <10, where its performance increased to 60.4% vs 54.1% ($p<0.001$) in the whole population. Then, we evaluated LIP performance according to the intention to diagnose principle in the whole population. This means that LIP was replaced by endoscopy in MELD score ≥10. Consequently, its performance decreased from 54.1% in "extended" LIP use to 51.0% in the secured

"restricted" individual LIP use ($p<0.001$) whereas its missed VNT rate became safe at 4.2% (Table S8). The performance of the secured individual LIP was still superior to that of the secured population LIP (51.0% vs 33.4%, $p<0.001$) but at the expense of a greater number of missed VNT. Indeed, the missed VNT rate among VNT was 15.9% (n=57) vs 4.2% (n=15), respectively, $p<0.001$.

Functional safety – The comparison of liver dysfunction level between missed VNT and other VNT is presented in Table S9. INR was chosen as the reference for liver dysfunction since it was more predictive of VNT than the MELD score. As expected, INR level was significantly lower in missed VNT than in other VNT for all tests. INR levels in missed VNT were not significantly different between tests in population screening. However, INR levels in missed VNT were significantly higher in tests for individual screening than the reference test for population screening. Thus, the INR level of LIP in missed VNT was significantly higher in individual screening vs population screening.

Validation by bootstrapping

In population screening, the inferior limit of the 95% CI, obtained by bootstrap, for spared endoscopy rate was >30% only for PLEASE and LIP (Table 3). The superior limit for the missed VNT rate was slightly superior to 5% (range: 5.5-6.1%) for all tests except the Baveno VI criteria (4.3%). In individual screening, the superior limit of the 95% CI for the missed VNT rate was >5% for LIP and the expanded Baveno VI criteria.

VNT specificity

In the whole population, unlike LIP (Figure 3), only bilirubin >250 µmol/l provided a 100% specificity or positive predictive value (PPV) for VNT in four patients (2 in ALD and 1 each in NAFLD and viral CLD), i.e. 0.2% of patients and 1.1% of VNT. This cut-off derived in the derivation set was robust in the validation set (Figure S8).

In NAFLD, BLIP provided 100% specificity or PPV for VNT in five patients (Figure 4), i.e. 0.8% of patients and 5.4% of VNT. Thus, this represents another interest of considering BMI in NAFLD.

Indeed, this gain significantly increased the performance of BLIP vs LIP in population screening (see below).

Aetiology influence

Population screening - The performance and safety of tests are described as a function of aetiology in Table 3. The Baveno VI criteria and Anticipate were limited in NAFLD by a missed VNT rate at 7.4% and in ALD by a spared endoscopy rate ≤17.2%. ALD was a challenging aetiology with a reduced spared endoscopy rate ≤25.9% in each test. However, PLEASE and LIP had no further limits. LIP was the best-performing test in NAFLD (37.2%). LR- was excellent (around 0.1) for all tests in all aetiologies, except the Baveno VI criteria and Anticipate in NAFLD.

Individual screening - LIP performance was significantly higher than that of the expanded Baveno VI criteria in viral CLD and NAFLD (Table 4). In NAFLD, LIP performance was significantly higher than that of the NAFLD cirrhosis criteria (Table S10).

NAFLD - The BLIP score was well calibrated for VNT probability (Figure S9). It was more discriminant for VNT than the LIP score was, with respective AUROCs of 0.822 and 0.804, p<0.001. The BLIP test performed better and/or demonstrated better safety than LIP in population screening (Table S11) and individual screening (Table S12). Thus, in individual NAFLD screening, the missed VNT rate by LIP was 5.4% in MELD <10 vs 14.3% in MELD ≥10 (p=0.116). These respective rates were 5.0% vs 4.8% (p=1) by BLIP. Thus, BLIP was safe whatever the level of liver dysfunction (but not secure). Likewise, the respective rates for the NAFLD cirrhosis criteria were 4.3% vs 12.5% (p=0.099). Finally, BLIP offered a better safety profile than did LIP or the NAFLD cirrhosis criteria (Figure S10). However, the secured BLIP score restricted to MELD <10 increased performance from 74.6% in all NAFLD to 79.7% (p=0.038). Then, applying the intention to diagnose principle, BLIP performance decreased from 74.6% in "extended" LIP to 70.8% in the secured "restricted" individual BLIP (p=0.128) whereas its missed VNT rate remained safe at 4.8% (Table S13).

Others

The marginal influence of inflammation (Table S14) and LSM reliability (Table S15) is described in the Supplemental Material.

3.4. Clinical application

All aetiologies - The LIP test is used in clinical practice as follows (Figure 4A). First, in population screening, LIP can be applied to all CLD whatever the liver dysfunction. Endoscopy can be confidently avoided in 33.4% of patients under a LIP cut-off around 5%, i.e. a missed VNT rate <5%, corresponding to a VNT prevalence of 1.9%. Otherwise, endoscopy is required in the remaining 66.6% of patients having a VNT prevalence of 21.8%. In individual screening (Figure 4B), LIP must be limited to MELD scores <10 where it securely spared 60.4% of endoscopies. The VNT risk can be more precisely quantified by the LIP score itself. Thus, its PPV can reach a maximum of 46.1%. However, the last two applications should be applied to populations with estimated VNT prevalences ≤15.2%: this preserves safety as predictive values are dependent on prevalence. We note that this rather low prevalence corresponds to clinical practice.

NAFLD - BLIP, which adds BMI to LIP, is the preferred test since it offers better safety and performance, partially due to its ability to also rule in VNT. Thus, in population screening, BLIP spared 40.8% of endoscopies (Figure 4C). In individual screening, BLIP must be limited to MELD scores <10 where it securely spared 79.7% of endoscopies (Figure 4B). The VNT risk can be more precisely quantified by the BLIP score itself. Thus, its PPV can reach a maximum of 100% (Figure 4C).

Practice - Clinical use adapted to every setting is summarised in Figure 4D. A simple exportable calculator (Excel file) is available at https://uabox.univ-angers.fr/index.php/s/wvZ84PzjM7FVwD6

DISCUSSION AND CONCLUSION

Originalities - Our large population encompassed a substantial spectrum of characteristics for the evaluation of VNT tests in CLD, including a wide range of liver function and the three main CLD aetiologies. The inclusion of patients irrespective of liver severity had two advantages. First, a few patients had VNT below the cut-off of 10 kPa used in the original definition of cACLD[1]. Excluding these patients would have resulted in missing more than 5% of VNT, which is the limit conceded in the Baveno VI statement. Our inclusion criteria fit better with the recent cACLD criteria[28]. Second, we have recently shown that a global strategy of non-invasive VNT screening performs better than a strategy restricted to cACLD, provided the test is secured, i.e. no missed VNT in severe CLD[14]. We report two screening strategies, one intended for individual patient screening and another for population screening. Until now, safety has been based on a quantitative definition (missed VNT <5%). Here, we refine this characteristic. Quantitatively, two safety definitions were used (sensitivity and NPV) and qualitatively two criteria were described: secureness, with no missed VNT in severe liver dysfunction; and a requirement for a low level of liver dysfunction in missed VNT (*functional safety*). Finally, with this large population, we were able to further evaluate interactions and resultantly provide an improved ratio of VNT markers. In the present setting, LR- is particularly relevant since it reflects both performance and safety in a single descriptor, but its use is restricted to population screening. Finally, we extended the principle of sparing endoscopies by ruling in VNT thanks to 100% specificity.

Main results - The LIP test performed better than the Baveno VI criteria, Anticipate and PLER. The exception was PLEASE, which provided an additional 1.8% in the spared endoscopy rate. That weak difference was significant (p=0.006) knowing the power of a paired test in a large population. However, LIP performed as well as PLEASE in NAFLD and ALD, and BLIP was superior to PLEASE in NAFLD.

The combination of three strong VNT predictors (platelets, LSM, PI) in a single ratio has three notable advantages. First, a unique VNT risk score can be derived in contrast to rules like the Baveno VI criteria, where two markers are used separately. Second, the cut-off for a fixed missed VNT rate is objectively determined and efficiently maximised (Figure S2). Thus, there is only one possible cut-off, the value of which depends only on the population characteristics. However, the present exhaustive population, with respect to size, aetiologies and liver severity, favours the exportability of a test cut-off from an epidemiological aspect. In comparison, there is an unlimited choice of combinations of cut-offs with two or more VNT predictors. Third, the calculation offers greater simplicity, robustness and performance compared to a logistic score including the markers. That greater simplicity is seen in the simple arithmetic calculation necessitating neither a specific web-calculator as for PLEASE nor a nomogram as for the Anticipate score. The greater robustness is conferred by the unique formula without marker coefficients depending on the population. Furthermore, the greater performance is demonstrated by the VNT AUROC, similar to that of a logistic score using the three variables but with a significantly higher spared endoscopy rate than the latter. 100% specificity was obtained optionally with bilirubin to rule in VNT in a few patients and even with BMI (included in the BLIP test) in more patients with NAFLD. This specificity level has not been previously reported. BLIP, a test devoted to NAFLD, performed better than the NAFLD cirrhosis criteria and spared up to 79.7% of endoscopies in individual screening where it was secured by restricting its use to MELD scores <10. We have shown that a MELD score cut-off of 8.6 corresponded to the upper cut-off of cACLD[14].

Which test to use? This is best approached by a process of elimination. First, the expanded Baveno VI criteria can be eliminated due to their unsafe missed VNT rate of 11% in population screening[14]. This high rate is due to the NPV definition used for missed VNT as discussed elsewhere [12]. Even considering individual screening, LIP outperformed the expanded Baveno VI criteria. Second, the Baveno VI criteria can be set aside also, as they were associated with the lowest overall performance and a missed VNT rate of 7.4% in NAFLD (responsible for a LR- at 0.20). Third, Anticipate can be discarded due to significantly lower overall performance and a missed VNT rate of 7.4% in NAFLD,

responsible for a low efficiency with LR- at 0.23. Fourth, PLER can be excluded due to its significantly lower overall performance. This leaves two competitors: LIP and PLEASE included in the VariScreen algorithm. PLEASE performed slightly better than LIP only in viral CLD. The higher PLEASE test performance was attributed to its algorithm stratified on sex and ethology; the PLEASE score discriminated VNT less well than LIP score (Figure 3B). LIP however had four advantages. First among them was its simplicity, and thus bedside usability. Second was its precise determination of VNT risk, born of its superior calibration for directly expressing VNT probability. Third, LIP provided exhaustive secureness. And fourth, having the nature of a score, it is far better adapted to a personalised strategy: individual and population screenings.

Finally, we may speak of two tests categories. On one hand, there are the simple tests like the Baveno VI criteria, PLER and LIP, with the latter outperforming the others, and on the other hand, there are the tests necessitating a specific calculator like Anticipate (requiring at least a nomogram) and PLEASE/VariScreen. Thus, the choice of one or the other depends on the clinical context. We note however that in NAFLD, which is becoming the major CLD aetiology, BLIP had two advantages in population screening: first, its performance (41%) was higher than that of PLEASE (37%) and second it allowed for the ruling in of VNT.

Which strategy? - Decision making will depend on the clinical background. The respective advantages and limits of the two possible strategies are detailed in Table 1. Three limits of the individual screening strategy must be opposed to its high performance. First, the rate of patients with missed VNT among VNT was higher (LIP: 18.4% vs 4.2%). Second, the patients with missed VNT had greater liver dysfunction, necessitating the restriction of individual screening to MELD scores <10. Third, the performance of the individual strategy depended on population prevalence and thus was limited to a VNT prevalence ≤15%, although that rate should often be the case considering a global strategy. Therefore, we privilege the population screening[12]. However, individual strategy is a personalized option adapted to certain patients, e.g. particularly reluctant to screening endoscopy. A

cost-efficacy study would help to determine the best strategy. In the meantime, the choice remains open and depends on clinical management and preferences (Figure 4D).

Limits - The first category of limits, which have been discussed in depth elsewhere [14], comprises those inherent to the population. Briefly, these include: the multicentric nature of the population, implying variability in patient recruitment; the retrospective design but prospective recording of data; the non-inclusion of grade 1 EV with red signs in VNT determination; the use of the MELD score (and not clinical complications like in the cACLD definition) to estimate the influence of liver dysfunction; and the non-evaluation of treatments. The second category of limits are particular to the present study, i.e. new tests. First, in population screening, the upper limit of the 95% CI by bootstrap was slightly above 5% in most tests including LIP. Therefore, despite reproducible results in the validation set, LIP should be validated in independent populations. It should be noted that PLEASE, determined in the same derivation set, has been independently and externally validated[29]. We underline however that robust validation requires large populations (≥400) as previously discussed[12].

Three putative LIP limits merit discussion. First, LIP was adjusted to aetiology. This induced one inconvenience. Indeed, three cut-offs (value around 5%) were necessary. However, this aspect also allowed us to make the test safe in NAFLD and to increase performance in ALD compared to a test with a unique cut-off (data not shown). ALD nonetheless remained a challenging aetiology for the VNT tests. The lower performance of LIP in ALD was due to poor synergy between LIP markers, as clearly shown in Figure 1A.

Second, LIP safety was 5.2% in individual screening, although this potential limit was circumvented by restricting the use of LIP to MELD scores <10, thus meeting secureness requirements. This restriction induced a paradox: performance was increased (+6.3%) in the remaining subgroup but decreased (-3.1%) in the intention to diagnose principle in the whole population.

Finally, the range of the LIP score's PPV varied from 0 to 46.1%, which is somewhat narrow. However, the knowledge of a VNT risk of about 50% will more greatly motivate patients and

physicians for endoscopy than will the vague risk "≥5%" provided by the usual binary tests. Of note, BLIP PPV reached 100% in NAFLD (Figure 4C).

Clinical application - LIP can be used in any patient with stable CLD whatever the liver severity (in population screening) and in the main three aetiologies, especially in the growing setting of NAFLD. However, patients with superimposed acute hepatitis should be excluded, as LIP is secured in patients with ALT up to 300 IU/l. The LIP test includes two categories, i.e. missed VNT <5% and ≥5%. The VNT risk in that last indeterminate category is quantified by the LIP score (Figure 4). The clinical limit of LIP is the requirement for VCTE. Otherwise, there is no additional cost associated with LIP in centres where VCTE is available. For several reasons, the use of VNT tests will continue even if primary prevention by non-selective beta-blockers (NSBBs) is extended to all liver complications[30] [31]. These reasons include adapting motivation for drug compliance, which is an important clinical challenge, and managing the contraindications and side effects of those NSBBs.[31] Moreover, the non-invasive criteria of clinically significant portal hypertension are not yet validated[28, 32]. Furthermore, adherence to screening by patients and primary care providers is improved by precise information[33, 34] and a negative perception of the disease[35]. With that respect, knowledge of the precise VNT status would be more convincing than that of clinically significant PHT.

Finally, physicians can choose between two strategies for LIP: population screening or individual screening. The latter performs better at the individual level (all etiologies: 54.1% vs 33.4%, NAFLD: 72.6% vs 37.2%) but is less safe from an epidemiological perspective and should be restricted to patients with MELD scores <10 (70% of the present population).

Conclusion - LIP is a test combining simplicity, performance, safety and deployability in each main aetiology of CLD. LIP can be used irrespective of liver severity in population screening and when the ALT level is ≤300. In NAFLD, the good performance of LIP is amplified by its combination with BMI. LIP performance can be very high for individual patient screening, even in NAFLD with BLIP (80%).

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-Carriè, Nathalie Barget.

Multicenter subpopulations:

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

- Baveno VI criteria 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], Vito Di Marco,[1] Antonio Craxi,[1] and Victor de Ledinghen[12]

[1]Department of Gastroenterology and Hepatology, University of 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, 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]CHG 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) and the sponsors of the subpopulations (who had no other role in the present study): VO-VCO study: the PHRC (Health Ministry, France); M116 study: Echosens, Paris, France; CO12 CirVir cohort: ANRS, France.

REFERENCES

1. de Franchis R, Faculty BV. Expanding consensus in portal hypertension: Report of the Baveno VI Consensus Workshop: Stratifying risk and individualizing care for portal hypertension. *J Hepatol* 2015;63(3):743-52.

2. 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(5):707-716.

3. Roccarina D, Rosselli M, Genesca J, Tsochatzis EA. Elastography methods for the non-invasive assessment of portal hypertension. *Expert Rev Gastroenterol Hepatol* 2018;12(2):155-164.

4. Thabut D, Bureau C, Layese R, et al. Validation of Baveno VI Criteria for Screening and Surveillance of Esophageal Varices in Patients With Compensated Cirrhosis and a Sustained Response to Antiviral Therapy. *Gastroenterology* 2019;156(4):997-1009.e5.

5. Moctezuma-Velazquez C, Saffioti F, Tasayco-Huaman S, et al. Non-Invasive Prediction of High-Risk Varices in Patients with Primary Biliary Cholangitis and Primary Sclerosing Cholangitis. *Am J Gastroenterol* 2019;114(3):446-452.

6. Stafylidou M, Paschos P, Katsoula A, et al. Performance of Baveno VI and Expanded Baveno VI Criteria for Excluding High-Risk Varices in Patients With Chronic Liver Diseases: A Systematic Review and Meta-analysis. *Clin Gastroenterol Hepatol* 2019;17(9):1744-1755 e11.

7. 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.

8. Ding NS, Nguyen T, Iser DM, et al. Liver stiffness plus platelet count can be used to exclude high-risk oesophageal varices. *Liver Int* 2016;36(2):240-5.

9. 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(6):1980-1988.

10. Jangouk P, Turco L, De Oliveira A, Schepis F, Villa E, Garcia-Tsao G. Validating, deconstructing and refining Baveno criteria for ruling out high-risk varices in patients with compensated cirrhosis. *Liver Int* 2017;37(8):1177-1183.

11. 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.

12. Cales P, Buisson F, Ravaioli F, et al. How to clarify the Baveno VI criteria for ruling out varices needing treatment by non-invasive tests. *Liver Int* 2019;39(1):49-53.

13. 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.

14. Berger A, Ravaioli F, Farcau O, et al. Including Ratio of Platelets to Liver Stiffness Improves Accuracy of Screening for Esophageal Varices That Require Treatment. *Clin Gastroenterol Hepatol* 2021;19:777-787.

15. Serper M, Kaplan DE, Shults J, et al. Quality Measures, All-Cause Mortality, and Health Care Use in a National Cohort of Veterans With Cirrhosis. *Hepatology* 2019;70(6):2062-2074.

16. Berger A, Ravaioli F, Farcau O, et al. The prevalence of esophageal varices needing treatment depends on gender, etiology and BMI. *J Hepatol* 2020;73:S751-2.

17. Kamath PS, Kim WR, Advanced Liver Disease Study G. The model for end-stage liver disease (MELD). *Hepatology* 2007;45(3):797-805.

18. Boursier J, Konate A, Gorea G, et al. Reproducibility of liver stiffness measurement by ultrasonographic elastometry. *Clin Gastroenterol Hepatol* 2008;6(11):1263-9.

19. Cales P, Oberti F, Bernard-Chabert B, Payen JL. Evaluation of Baveno recommendations for grading esophageal varices. *J Hepatol* 2003;39(4):657-9.

20. 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(1):29-36.

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(6):512-9.

22. Singal AK, Kamath PS. Model for End-stage Liver Disease. *J Clin Exp Hepatol* 2013;3(1):50-60.

23. Abraldes JG, Bureau C, Stefanescu H, et al. Noninvasive tools and risk of clinically significant portal hypertension and varices in compensated cirrhosis: The "Anticipate" study. *Hepatology* 2016;64(6):2173-2184.

24. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;351:h5527.

25. Boursier J, de Ledinghen V, Poynard T, et al. An extension of STARD statements for reporting diagnostic accuracy studies on liver fibrosis tests: the Liver-FibroSTARD standards. *J Hepatol* 2015;62(4):807-15.

26. Boursier J, Zarski JP, de Ledinghen V, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. *Hepatology* 2013;57(3):1182-91.

27. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015;350:g7594.

28. Papatheodoridi M, Hiriart JB, Lupsor-Platon M, et al. Refining the Baveno VI elastography criteria for the definition of compensated advanced chronic liver disease. *J Hepatol* 2021;74(5):1109-116.

29. Hu Y, Wen Z. Validation and comparison of non-invasive prediction models based on liver stiffness measurement to identify patients who could avoid gastroscopy. *Sci Rep* 2021;11(1):150.

30. Villanueva C, Albillos A, Genesca J, et al. beta blockers to prevent decompensation of cirrhosis in patients with clinically significant portal hypertension (PREDESCI): a randomised, double-blind, placebo-controlled, multicentre trial. *Lancet* 2019;393:1597-1608.

31. Garcia-Tsao G, Abraldes JG. Non-selective beta-blockers in compensated cirrhosis: Preventing variceal hemorrhage or preventing decompensation? *Gastroenterology* 2021.

32. Pons M, Augustin S, Scheiner B, et al. Noninvasive Diagnosis of Portal Hypertension in Patients With Compensated Advanced Chronic Liver Disease. *Am J Gastroenterol* 2020.

33. Yakovchenko V, Bolton RE, Drainoni ML, Gifford AL. Primary care provider perceptions and experiences of implementing hepatitis C virus birth cohort testing: a qualitative formative evaluation. *BMC Health Serv Res* 2019;19(1):236.

34. Ispas S, So S, Toy M. Barriers to Disease Monitoring and Liver Cancer Surveillance Among Patients with Chronic Hepatitis B in the United States. *J Community Health* 2019;44(3):610-625.

35. Sovaila S, Purcarea A, Gheonea D, Ciurea T. Specific Factors That Influence Adherence to Beta Blocker Treatment in Primary Prevention of Variceal Bleeding in Cirrhotic Romanian Patients. a Proof of Concept Qualitative Study. *J Med Life* 2018;11(4):355-358.

TABLES

Table I. Main characteristics of the two strategies for VNT screening.

Strategy	Cut-offs for VNT ruled		Advantages	Limits
	out	in		
Individual patient	95% NPV	**100% PPV/ specificity**	High performance. Easier cut-off determination.	Increased number and liver dysfunction of missed VNT. Restricted to MELD <10. VNT prevalence dependence. Comparison of missed VNT rate is less powerful [a]. LR- is not applicable.
Population	95% sensitivity	**100% PPV/ specificity**	VNT prevalence independence. The lowest liver dysfunction in missed VNT. LR- is a unique diagnostic descriptor.	Cut-offs are less optimistic since the reference population is smaller.

NPV: negative predictive value, PPV: positive predictive value, LR-: LR-: negative likelihood ratio
New strategy characteristic developed in the present study is in bold.
[a] Since using an unpaired statistical test

Table II. VNT discrimination by scores as a function of populations.

	Anticipate	PLER	PLEASE	LIP_{PI}
Derivation set:				
AUROC (95%CI)	0.770 (0.740-0.801	0.761 (0.731-0.792)	0.798 (0.770-0.827)	0.776 (0.747-0.805)
Comparison (p [a]):				
Anticipate	-	*0.024*	*0.007*	*0.274*
PLER		-	*<0.001*	*<0.001*
PLEASE			-	*0.015*
LIP				-
Validation set:				
AUROC (95%CI)	0.779 (0.741-0.818)	0.779 (0.740-0.817)	0.771 (0.732-0.810)	0.786 (0.748-0.825)
Comparison (p [a]):				
Anticipate	-	*0.913*	*0.560*	*0.286*
PLER		-	*0.586*	*0.109*
PLEASE			-	*0.218*
LIP				-
Whole population:				
AUROC (95%CI)	0.773 (0.749-0.797)	0.767 (0.743-0.791)	0.789 (0.766-0.812)	0.779 (0.756-0.803)
Comparison (p [a]):				
Anticipate	-	*0.059*	*0.061*	*0.136*
PLER		-	*0.012*	*<0.001*
PLEASE			-	*0.196*
LIP				-

PLER: platelet / liver elastometry ratio, PLEASE: platelet / liver elastometry ratio adjusted on aetiology, sex, INR
[a] Paired Delong test

Table III. Missed VNT (%) and spared endoscopy (%) rates of tests to spare endoscopy according to population screening (sensitivity ≥95%) in the whole population and as a function of aetiology.

	B6C	Anticipate	PLER	PLEASE	LIP_{PI}	*p* [a]
Whole population:						
Missed VNT [b]	2.5 (0.9-4.3)	3.3 (1.5-5.5)	3.6 (1.7-5.6)	3.6 (1.7-5.6)	4.2 (2.2-6.1)	*0.469*
Spared endoscopy [c]	23.9 (22.0-25.5)	24.5 (22.6-26.2)	27.3 (25.5-29.0)	35.2 (33.1-37.0)	33.4 (31.6-35.3)	*<0.001*
LR-	0.045	0.118	0.115	0.089	0.114	-
Virus:						
Missed VNT	1.1 (0.0-2.8)	2.2 (0.5-4.5)	3.3 (1.1-6.1)	3.9 (1.2-6.7)	4.5 (1.7-8.0)	*0.102*
Spared endoscopy	21.6 (19.2-24.0)	25.0 (22.5-27.4)	25.9 (23.5-28.3)	38.0 (35.3-40.8)	34.9 (32.2-37.5)	*<0.001*
LR-	0.045	0.078	0.113	0.089	0.112	-
NAFLD:						
Missed VNT	7.4 (2.3-13.5)	7.4 (2.4-13.5)	4.9 (1.1-10.8)	3.7 (0.0-8.9)	4.9 (1.1-9.9)	*0.236*
Spared endoscopy	33.4 (29.7-36.9)	28.9 (25.6-32.5)	35.3 (31.6-38.9)	36.9 (33.2-40.3)	37.2 (33.6-40.6)	*<0.001*
LR-	0.201	0.233	0.125	0.089	0.118	-
ALD:						
Missed VNT	1.0 (0.0-3.3)	2.0 (0.0-5.0)	3.0 (0.0-6.9)	3.0 (0.0-7.0)	3.0 (0.0-7.0)	*0.573*
Spared endoscopy	16.0 (13.0-19.3)	17.2 (14.0-20.5)	19.6 (16.2-23.0)	25.9 (22.0-29.9)	24.7 (20.8-28.4)	*<0.001*
LR-	0.051	0.095	0.126	0.095	0.127	-
Comparison between aetiologies (*p* [d]):						
Missed VNT	*<0.001*	*<0.001*	*<0.001*	*<0.001*	*<0.001*	-
Spared endoscopy	*0.016*	*0.105*	*0.774*	*0.920*	*0.764*	-

B6C: Baveno VI criteria, VNT: varices needing treatment, LR-: negative likelihood ratio
Results in brackets are 95% CI obtained by bootstrapping based on 1000 samples stratified on aetiology and sex.
[a] Paired Cochran test
[b] Each pair comparison: p>0.05 by McNemar test
[c] Each pair comparison: p<0.001 except for B6C vs Anticipate: p=0.367 and PLEASE vs LIP_{PI}: p=0.006 by McNemar test. Other comparisons per aetiology in Table S6.
[d] Unpaired Chi^2 test for spared endoscopy and likelihood ratio test for missed VNT

Table IV. Missed VNT (%) and spared endoscopy (%) rates of tests to spare endoscopy according to individual screening (NPV ≥95%) in the whole population and as a function of aetiology.

	EB6C	LIP_{PI}	*p* [a]
Whole population:			
Missed VNT	4.0 [11.1]	5.2 [b] [18.4]	*0.175*
	(2.8-5.1)	(4.0-6.4)	
Spared endoscopy	42.7	54.1	*<0.001*
	(40.7-44.5)	(52.1-55.9)	
Virus:			
Missed VNT	3.6 [10.1]	5.2 [b] [19.1]	*0.185*
	(2.0-5.3)	(3.6-7.0)	
Spared endoscopy	42.1	54.7	*<0.001*
	(39.4-45.0)	(51.9-57.5)	
NAFLD:			
Missed VNT	4.2 [19.8]	5.4 [b] [33.3]	*0.397*
	(2.4-6.1)	(3.6-7.5)	
Spared endoscopy	55.8	72.6	*<0.001*
	(52.1-59.7)	(69.2-75.8)	
ALD:			
Missed VNT	4.7 [6.0]	3.8 [5.0]	*0.709*
	(1.5-8.8)	(0.8-7.1)	
Spared endoscopy	25.9	26.9	*0.511*
	(22.3-30.0)	(23.1-30.8)	
Comparison between aetiologies (*p* [c]):			
Missed VNT	*0.817*	*0.716*	-
Spared endoscopy	*<0.001*	*<0.001*	-

Figures in squared brackets are missed VNT among VNT (i.e. if cut-offs were applied to population screening). Results in brackets are 95% CI obtained by bootstrapping based on 1000 samples stratified on aetiology and sex.

[a] Unpaired Chi^2 test for spared endoscopy and likelihood ratio test for missed VNT

[b] The value is over the fixed cut-off at 5% but this drawback is circumvented by the secureness rule limiting individual screening to patients with MELD scores <10.

[c] Unpaired Chi^2 test for spared endoscopy and likelihood ratio test for missed VNT

FIGURES

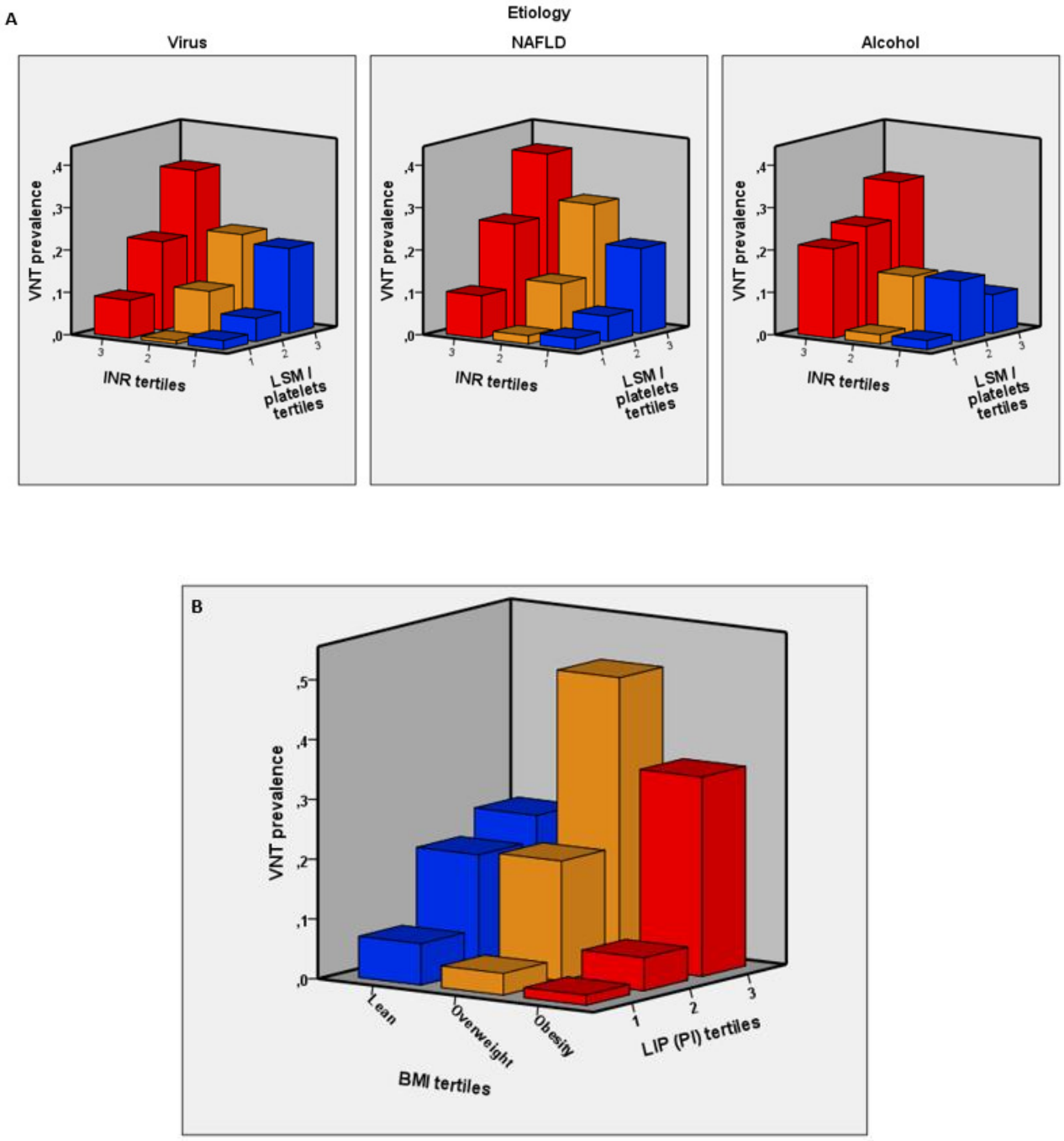


Fig. 1. Interaction and synergies of VNT markers. Panel A: INR with LSM / platelets ratio as a function of aetiology. Panel B: LIP with BMI in NAFLD.

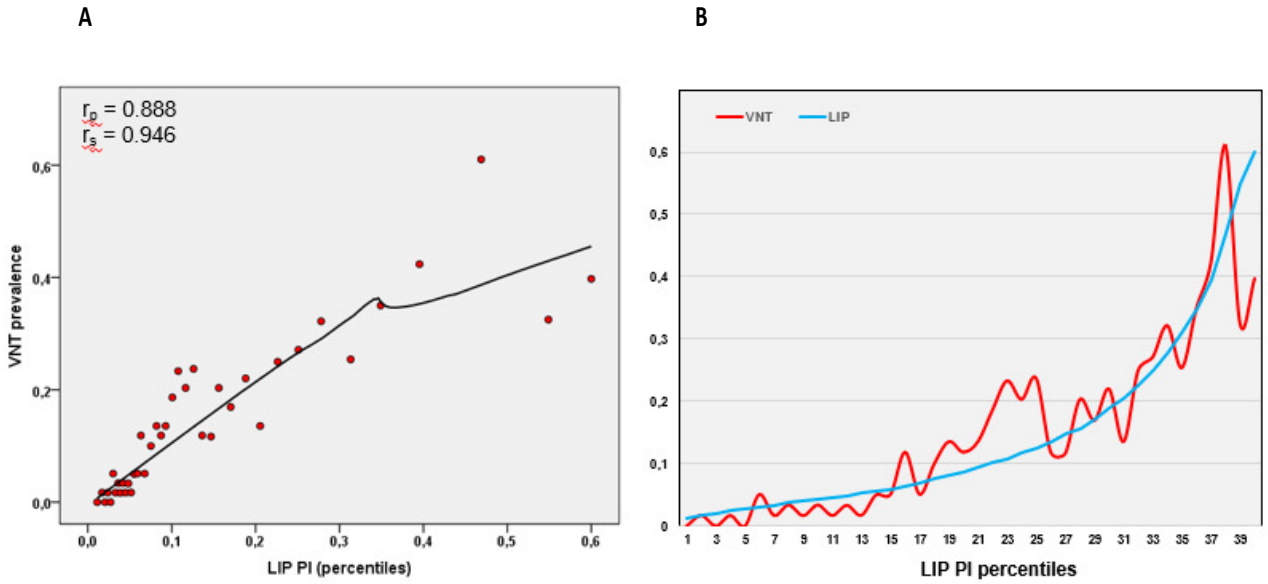


Fig. 2. Calibration of the LIP score for VNT risk. Panel A: curve from non-linear regression (LOWESS) with LIP score per percentile on X axis. Panel B: interpolation curve with LIP percentile rank (40) on X axis.

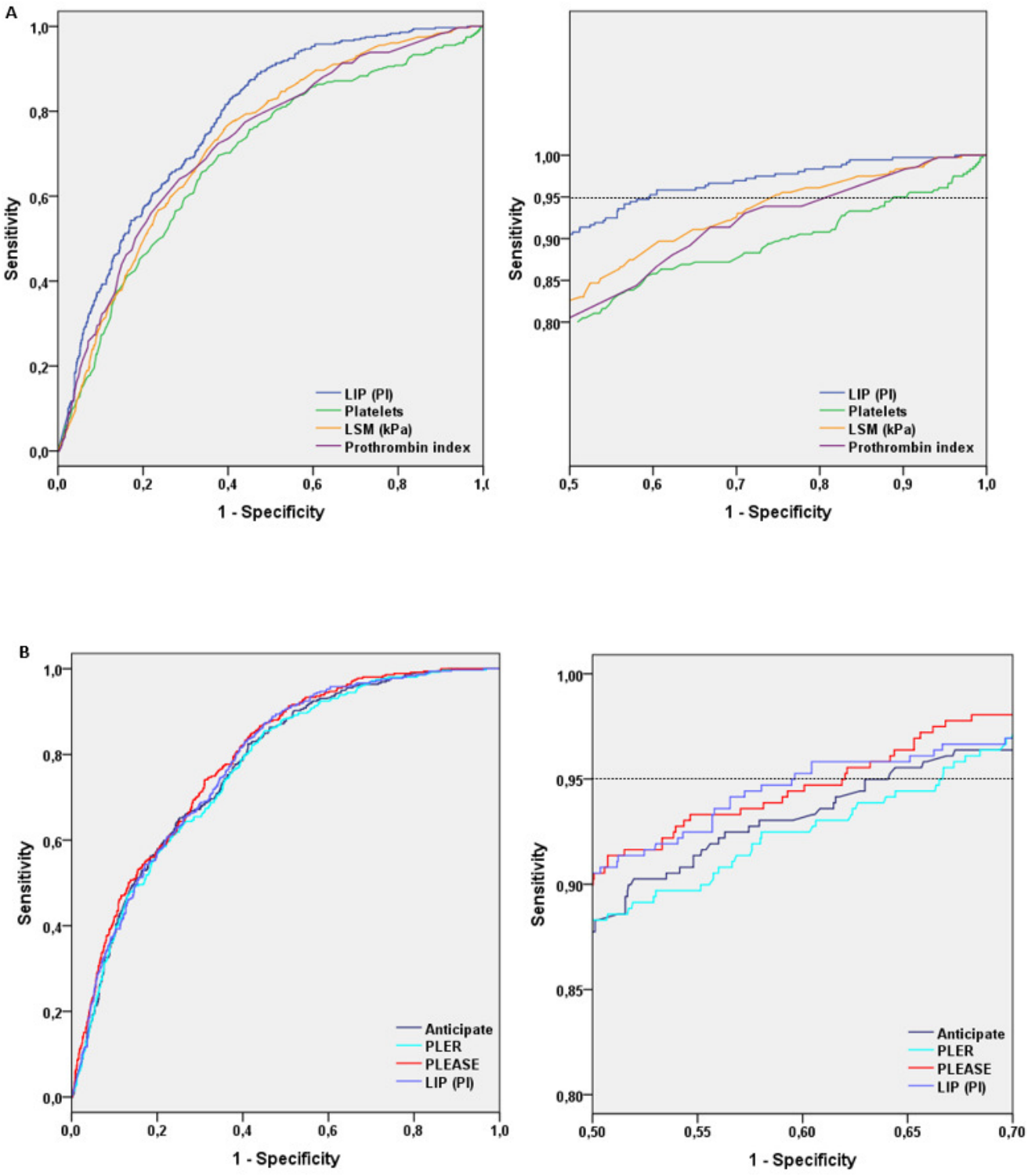


Fig. 3. Discrimination for VNT diagnosis (ROC curves). Panel A: LIP score and its composite markers with magnification showing the subset size with ≥95% sensitivity. Panel B: scores evaluated. In both of the figures on the right, the horizontal dashed lines show the superiority of LIP in sparing endoscopy with a missed VNT rate <5%.

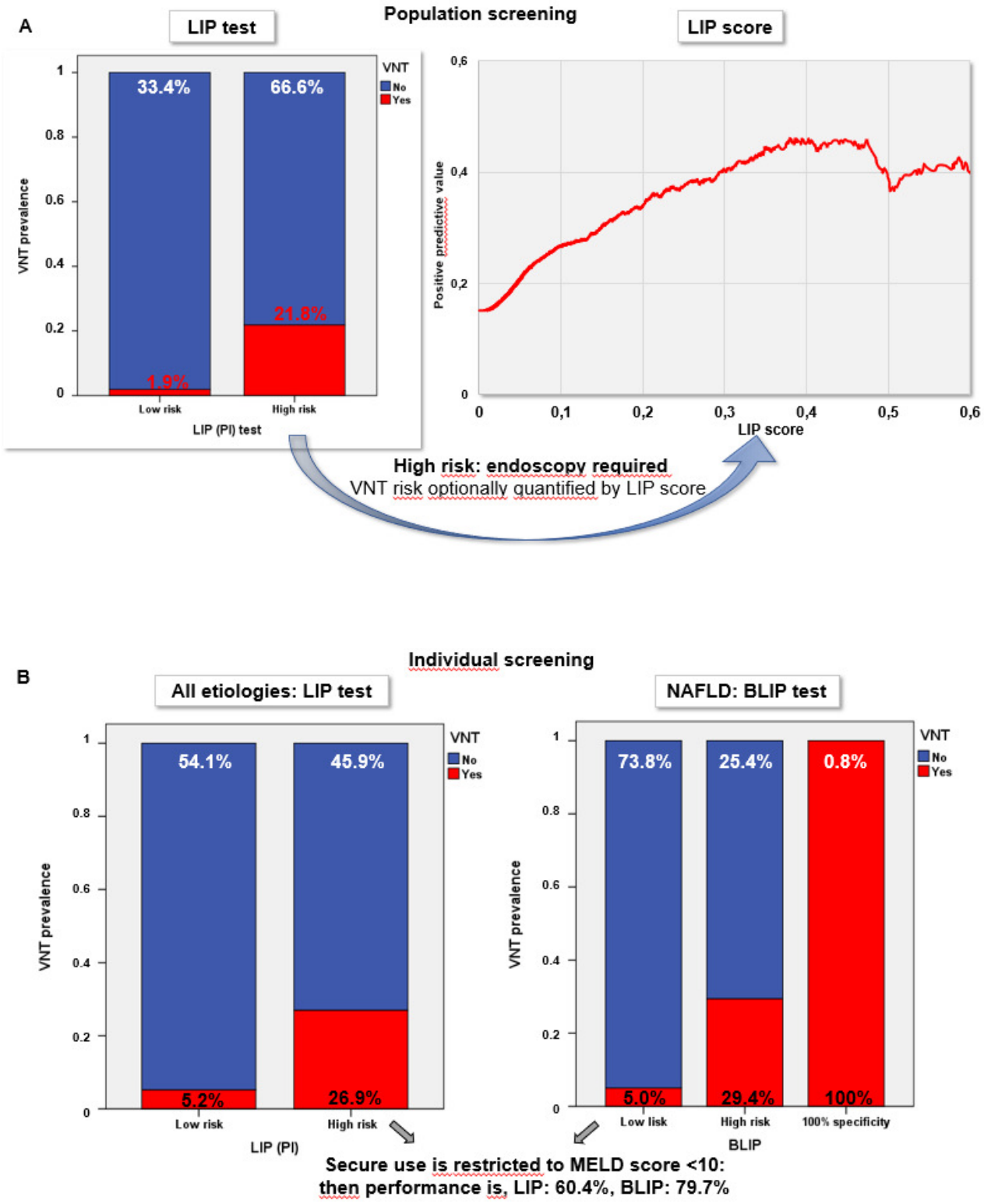
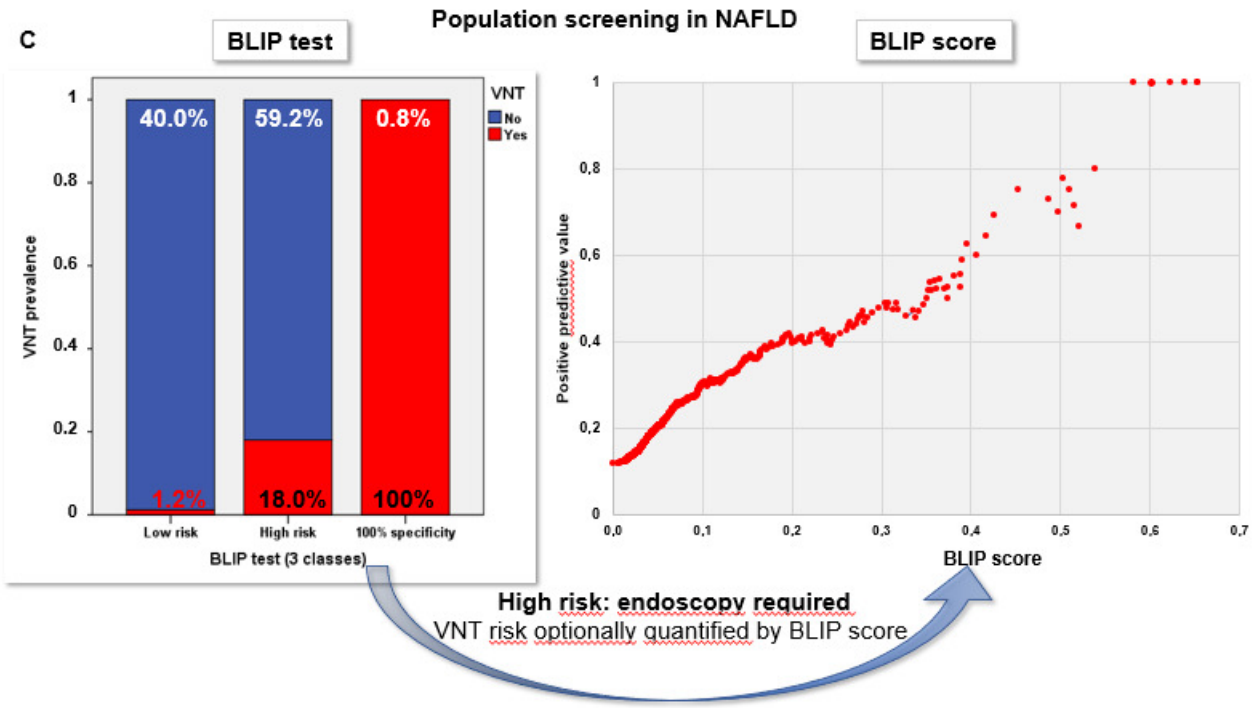


Fig. 4. Clinical application of new tests. Panel A: LIP for population screening in all aetiologies; the figures within bars indicate category prevalence (top) and VNT prevalence (bottom). Panel B: Individual screeningFigures within bars from the whole population indicate category prevalence (top) and VNT prevalence (bottom).



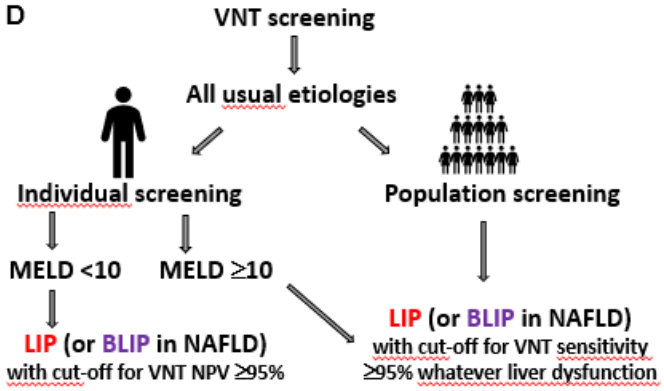


Fig. 4. Clinical application of new tests. Panel C: BLIP for population screening in NAFLD. Panel D: different screening strategies. Figures within bars from the whole population indicate category prevalence (top) and VNT prevalence (bottom).

TABLES LIST

FIGURES LIST

TABLE OF CONTENTS

ANNEXES

Individual and population screening of varices needing treatment by a simple, safe and accurate test

Supplemental Material

INTRODUCTION

Illustrations referenced in red characters refer to the main text and those in blue to the present Supplemental Material. Tables S1-S15 and Figures S1-S10 are referred to in the main text whereas Table S16 and Figures S11-S12 depict additional data. In the main text, LIP_{PI} was shortened to LIP. In the present Supplemental Material, the full LIP_{PI} designation will be used because we also describe LIP_{INR} here. Performance and safety of LIP_{INR} and $BLIP_{INR}$ were very close to those of LIP_{PI} and $BLIP_{PI}$, respectively. Therefore, they are not detailed.

METHODS

Predefined strategy

The planned strategy was discussed during two meetings in the presence of the co-authors during EASL meetings (Paris 2018 and Vienna 2019) and finalized with the main investigators (AB, FR, PC).

Score and test construction - The predefined strategy was finalized from three previous studies. The first of them clarified the methodological conditions of diagnostic studies based on non-invasive markers[1]. The second clarified the definitions used in the Baveno VI criteria[2]. The third selected the independent VNT predictors by using the biomarkers usually performed in CLD work-ups in a large series of patients unrestricted for liver severity and encompassing the three main CLD etiologies[3]. Having a large series with a VNT prevalence of 15% showing several new risk factors, the next step was to determine a good-performing predictive model for VNT enriched by interaction terms, the advantages of the development of which we have demonstrated. Indeed, we have demonstrated that the two markers used in the Baveno VI criteria interact and that the ratio platelets / LSM advantageously replaces both this interaction

term and the single markers in the model[4]. In our work, the advantages were increased accuracy and an automatic cut-off choice contrary to the Baveno VI criteria and their multiple derivatives distinguished by different cut-offs. Consequently, due to the increased number of markers tested here compared to our previous study[4], we planned to obtain several interaction terms between two markers with some markers common to several interaction terms. However, as our aim was to develop a simple test, we did not consider interactions with qualitative variables (aetiology, sex). Interactions with markers reflecting liver dysfunction were privileged. Indeed, based on a previous work on the PLEASE test, our hypothesis was that the adjustment on this kind of marker would fit the test with safety criteria, especially secureness i.e. no missed VNT in patients with marked liver dysfunction.

Then, we planned to derive different ratios by using a combination of LSM, platelets and the marker reflecting liver dysfunction the most predictive of VNT. Good-performing scores like Anticipate had the advantage of providing a logit score, by binary logistic regression, expressing VNT probability (from 0 to 1). But this logit function is also perceived as complex by physicians who appreciate simple ratios such as the FIB-4 test for liver fibrosis. Therefore, while privileging a simple arithmetic ratio rather than a logit function, we planned to compare the best-performing simple ratio with the logit score including the same markers. Nevertheless, the inconvenience of an arithmetic ratio is that it provides a number having a large range that does not directly express VNT probability. Therefore, we planned to adapt the ratio with two correcting factors so that it directly expressed this VNT probability. This was first a constant to calibrate the ratio to VNT probability and then to cap it to fit with the highest probability. It should be noted that even in a logit score, it is sometimes necessary to cap certain markers, e.g. platelets are capped to 150 G/l in the Anticipate score[5].

Although in a perspective of test simplicity, we also planned to derive a test to spare endoscopy from the previous score precisely quantifying the VNT risk. Consequently, we screened the ROC curves of obtained scores offering VNT sensitivities ≥95% (also 100% specificities: see below) in the largest patient subset. Then, we calculated the diagnostic indices (especially, sensitivity, NPV and specificity) as a function of growing score. The cut-off was determined by the first occurrence of the targeted values of diagnostic indices: sensitivity ≥95% then NPV ≥95% to obtain a missed VNT rate <5%. Thus, this technique maximized the spared endoscopy rate. The counterpart was the risk of exceeding the fixed limit of 5% for the missed VNT rate. We evaluated this risk through two characteristics, either the rate in the validation population or the confidence interval of the rate measured by bootstrapping in the whole population. We privileged this last measure in the whole population to get closer to that of the general population. We planned to obtain a safe test in each aetiology. Indeed, half of our patients had

viral CLD whereas NAFLD and ALD are growing aetiologies. Therefore, a single cut-off might penalize safety in other aetiologies due to the preeminent influence of viral CLD. Resultantly, cut-offs were determined in each aetiology.

We excluded LSM results obtained with the XL probe since this was only available in a subset of patients with NAFLD. This population had specific characteristics, especially risk factors for VNT in a preliminary study[3], which will necessitate specific further study.

Finally, we planned to derive and validate the final scores respectively in the same derivation and validation populations as in our previous study[4].

VNT specificity - We introduce here a new paradigm by hypothesizing that if tests could reach a 100% VNT specificity, then endoscopy could be made unnecessary also by ruling in VNT.

NAFLD - In the preliminary study on VNT risk factors[3], we observed BMI as a negative predictor in NAFLD. Therefore, we planned to explore its contribution in this setting all the more so in that this anthropomorphic marker is a clinical quantitative datum that is more easily handled in a ratio than other clinical variables are, especially qualitative variables such as sex, while aetiology was partially considered in cut-offs and age was not a VNT predictor[3].

Statistics

The cost-benefit ratio (or efficiency or profitability) of a diagnostic test is factually expressed by the LR- for VNT in this context. LR- is a conventional diagnostic index but sometimes considered as ancillary. The LR- is the ratio of 1 - sensitivity for VNT (i.e. the missed VNT rate) divided by specificity for missed VNT, and therefore theoretically independent of VNT prevalence. Thus, in the present setting, the LR- is particularly relevant since it reflects both performance and safety in a single descriptor. However, LR- calculation is pertinent to population but not individual screening.

Formulae

Scores to diagnose VNT

LIP_{PI} score: (LSM*45)/(PI*platelets)

LIP_{INR} score: (LSM*INR)/(platelets*2)

Both scores are capped at 0.6.

$BLIP_{PI}$ score: (LSM*1350)/(PI*platelets*BMI)

$BLIP_{INR}$ score: (LSM*INR*15)/(platelets*BMI)

Units: LSM (kPa, VCTE using the M probe (Fibroscan)), prothrombin index (PI: %), platelets (G/l), BMI: kg/m^2.

Tests to spare endoscopy

The cut-offs for scores to obtain tests missing <5% of VNT are listed in Table S16.

A simple exportable calculator (Excel file) is available at https://uabox.univ-angers.fr/index.php/s/wvZ84PzjM7FVwD6

RESULTS

Sensitivity analysis

Inflammation influence - In population screening, test performance (Figure S6C) and safety (Figure S6D) did not globally vary as a function of ALT deciles. However, and interestingly, the missed VNT rate was 0% for all tests in the last decile including 237 patients whose ALT distribution (IU/l) was, range: 134-615, median: 179, quartiles 1 and 3: 155-217, 90th percentile: 291, 95th percentile: 346. This indicates that tests were also secured in patients with mild hepatitis. The same trend was observed in individual screening (Figure S11). In normal transaminases, test performance was significantly higher, but safety did not significantly change vs increased transaminases (Table S14).

LSM reliability - Counterintuitively, the performance was significantly higher in poorly reliable vs. very reliable LSM[6] while safety was not significantly influenced by LSM reliability (Table S15). Thus, LSM reliability criteria, adapted to Metavir fibrosis staging[6], have no influence on VNT diagnosis.

Additional results

BLIP performance (40.8%) was not significantly (p=0.113) different from PLEASE performance (38.3%) in NAFLD.

The logistic score including platelets, LSM and PI had an AUROC of 0.777 vs 0.776 for LIP_{PI} (p=0.835) in the derivation set, but the subset with VNT sensitivity ≥95% was larger with LIP_{PI} (Figure S12A), i.e. 29.3% vs 26.6% (p=0.001 by McNemar test). The logit score validates the limit of LIP_{PI} at 0.6 (Figure S12B). The logit score and LIP_{PI} were well correlated (r_p: 0.935, r_s: 0.949) (Figure S12C), which validated the accuracy of the LIP_{PI} formula.

DISCUSSION

Safety interpretation in individual screening

We underline that there is a potential misinterpretation for the comparison of missed VNT rates between tests here. Indeed, contrary to population screening, the rates are not calculated in the patient group with VNT but in the group with spared endoscopy whose size depends on test performance. The spared endoscopy rates were 42.7% for the extended Baveno VI criteria and 54.1% ($p<0.001$) for LIP_{PI}. The respective missed VNT rates were 4.0% and 5.2% ($p=0.175$) but the number of patients was 1.65 times higher with LIP_{PI} (n=66) than with the extended Baveno VI criteria (n=40). However, in severe CLD (MELD score $\geq$10), the respective VNT rates were 12.2% and 8.6% ($p=0.595$) so that the number of patients was 0.8 times lower with LIP_{PI} (n=8) than with the extended Baveno VI criteria (n=10) here.

REFERENCES

1. Boursier J, de Ledinghen V, Poynard T, et al. An extension of STARD statements for reporting diagnostic accuracy studies on liver fibrosis tests: the Liver-FibroSTARD standards. *J Hepatol* 2015;62(4):807-15.
2. Cales P, Buisson F, Ravaioli F, et al. How to clarify the Baveno VI criteria for ruling out varices needing treatment by non-invasive tests. *Liver Int* 2019;39(1):49-53.
3. Berger A, Ravaioli F, Farcau O, et al. The prevalence of esophageal varices needing treatment depends on gender, etiology and BMI. *J Hepatol* 2020;73:S751-2.
4. Berger A, Ravaioli F, Farcau O, et al. Including Ratio of Platelets to Liver Stiffness Improves Accuracy of Screening for Esophageal Varices That Require Treatment. *Clin Gastroenterol Hepatol* 2021;19:777-787.
5. Abraldes JG, Bureau C, Stefanescu H, et al. Noninvasive tools and risk of clinically significant portal hypertension and varices in compensated cirrhosis: The "Anticipate" study. *Hepatology* 2016;64(6):2173-2184.
6. Boursier J, Zarski JP, de Ledinghen V, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. *Hepatology* 2013;57(3):1182-91.

Table S1. Main analyses for VNT by forward logistic regression in the derivation population. p values [a] and influence [b] of variables are simultaneously reported. AUROC are also reported.

	AUROC	Multivariate analysis		
Interactions tested		**No [c]**	**Yes**	**Yes**
R^2	-	0.225	0.252	0.241
Aetiology (reference: virus):	-	***0.016***	***0.039***	*0.064*
NAFLD	-	*0.430*	*0.859*	*0.224*
ALD	-	***0.018***	***0.021***	***0.022***
Platelets	0.684	***<0.001***	*0.921*	*0.677*
ALT	0.538	***0.052***	***0.026***	***0.035***
Bilirubin	0.673	***0.020***	***0.001***	***0.003***
LSM	0.726	***<0.001***	***0.031***	*0.608*
PI	0.734	***<0.001***	***<0.001***	***0.032***
Albumin	0.670	***0.006***	***0.010***	***0.013***
Sex (male)	-	***0.001***	***<0.002***	***0.001***
Age	0.509	*0.379*	*0.681*	*0.567*
BMI	0.512	*0.971*	*0.852*	*0.852*
Platelets / LSM ratio [d]	0.761	-	***<0.001***	***0.924***
Platelets * PI	-	-	*0.722*	*0.720*
PI * LSM	-	-	***0.009***	*0.316*
PLER * PI	-	-	-	***<0.001***

LSM: liver stiffness measurement, PI: prothrombin index (%), PLER: platelets / LSM ratio. Bold p: significant difference. Units: see Table S2.

[a] -: not tested.

[b] Influence on VNT: positive, negative β coefficient

[c] MELD score was not an independent predictor in a subpopulation (details not shown). Variables excluded owing to collinearity: AST (with ALT), INR (with PI)

[d] Ratio to circumvent classical interaction terms. Platelets/LSM ratio is detailed elsewhere[4].

Table S2. Patient characteristics as a function of aetiology in the whole population [a].

	Viral	**NAFLD**	**Alcoholic**	***p*** [b]			
	(V)	**(N)**	**(A)**	**All**	**V vs N**	**V vs A**	**N vs A**
Patients (n)	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*
Obesity (%) [c]	16.9	**56.6**	22.4	*<0.001*	*<0.001*	*0.026*	*<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: IQR/median (%)	19±13	19±12	17±30	*0.465*	*0.982*	*0.221*	*0.604*
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 cut-off for VNT (kPa)	9.0	10.3 [c]	10.0	-	-	-	-

LSM: liver stiffness measurement, VNT: varices needing treatment. Specific characteristics of each aetiology in bold.

[a] These results were analysed in the whole population due to the similarity of population sets.

[b] Student's t test / ANOVA or Chi² test / Fisher test

[c] BMI ≥ 30 kg/m²

[d] One outlier value at 6.3 excluded

Table S3. Patient characteristics as a function of population sets.

	All	Derivation	Validation	*p* [ab]
Patients (n)	2368	1579	789	-
Age (years)	59.2±11.0	59.3±10.9	59.0±11.2	*0.436*
Sex (% male)	63.5	62.6	65.1	*0.239*
Aetiology (%):	-	-	-	*0.999*
Viral	50.2	50.2	50.2	*1*
NAFLD	28.9	28.9	28.9	*1*
Alcoholic	20.9	20.8	20.9	*1*
VNT (%)	15.2	15.1	15.2	*1*
BMI (kg/m²)	28.4±5.8	28.5±5.8	28.1±5.6	*0.204*
AST (IU/l)	72±55	71±51	74±62	*0.427*
ALT (IU/l)	66±58	66±58	68±59	*0.277*
Normal transaminases (%)	15.0	15.0	14.9	*0.949*
Albumin (g/l)	39.8±5.6	39.9±5.6	39.6±5.6	*0.262*
Bilirubin (µmol/l)	19±21	20±23	19±17	*0.488*
INR	1.2±0.2	1.2±0.2	1.2±0.2	*0.953*
Platelets (G/l)	149±75	149±74	149±76	*0.940*
Creatinine (µmol/l)	72±32	72±31	71±35	*0.531*
MELD score	9.5±3.0	9.5±3.0	9.4±3.0	*0.485*
LSM (kPa)	27±18	27±18	28±19	*0.459*
LSM ≥10 kPa (%)	92.8	93.0	92.4	*0.612*
LSM reliability (%):	-	-	-	*0.717*
Very reliable	26.7	26.1	27.9	-
Reliable	63.6	63.8	63.0	-
Poorly reliable	9.8	10.1	9.0	-
LSM cutoff for VNT (kPa)	-	-	10.6 [c]	-
Metavir F stage by LSM (%):	-	-	-	*0.743*
<F3 (≤10.8 kPa)	10.3	10.1	10.8	*0.596*
F3 (10.9-17.6 kPa)	28.9	29.3	28.1	*0.549*
F3/4 (17.7-25.7 kPa)	20.2	20.8	19.4	*0.496*
F4 (≥25.8 kPa)	40.6	40.0	41.7	*0.434*

LSM: liver stiffness measurement, VNT: varices needing treatment

[a] Student's t test or Chi² test / Fisher test. [b] Derivation vs validation set. [c] One outlier value at 6.3 excluded

Table S4. VNT AUROC: LIP_{PI} score vs its composite markers in the whole population (2368 patients).

	AUROC	p [a] vs LIP_{PI}
LIP_{PI}	0.779	-
Platelets	0.692	*<0.001*
LSM	0.729	*<0.001*
PI	0.728	*<0.001*
INR	0.728	*<0.001*

LSM: liver stiffness measurement, PI: prothrombin index (%)

[a] Paired Delong test

Table S5. Missed VNT and spared endoscopy rates (%) of LIP_{PI} to spare endoscopy as a function of population set.

	All	Derivation	Validation	*p* [a]
Population screening:				
All				
Missed VNT [b]	4.2	4.6	3.3	*0.781*
Spared endoscopy	33.4	33.1	34.1	*0.644*
Virus				
Missed VNT [b]	4.5	4.2	**5.1**	*0.721*
Spared endoscopy	34.9	34.3	36.1	*0.561*
NAFLD				
Missed VNT [b]	4.9	**5.6** [c]	3.7	*1*
Spared endoscopy	37.2	36.8	38.2	*0.738*
ALD				
Missed VNT [b]	3.0	4.5	0	*0.549*
Spared endoscopy	24.7	25.2	23.6	*0.741*
Individual screening:				
All				
Missed VNT [d]	**5.2** [e]	4.8	**5.8** [e]	*0.504*
Spared endoscopy	54.1	53.8	54.5	*0.760*
Virus				
Missed VNT [d]	**5.2** [e]	4.8	**6.0** [e]	*0.577*
Spared endoscopy	54.7	54.6	54.8	*1*
NAFLD				
Missed VNT [d]	**5.4** [e]	4.9	**6.4** [e]	*0.533*
Spared endoscopy	72.6	71.3	75.0	*0.319*
ALD				
Missed VNT [d]	3.8	4.4	2.4	*1*
Spared endoscopy	26.9	27.7	25.5	*0.667*

[a] Fisher test between derivation and validation sets
[b] In patients with VNT
[c] The limit of cut-off >5% is deleted with BLIP test (Table S12).
[d] In patients with spared endoscopy
[e] The limit of cut-off >5% is circumvented by the secureness rule limiting individual screening to patients with MELD scores <10.

Table S6. Test comparison (p) [a] for spared endoscopy rate reported in Table 3 in the whole population (n=2368) and as a function of etiology.

	B6C	Anticipate	PLER	PLEASE	LIP_{PI}
All					
B6C	-	*0.367*	*<0.001*	*<0.001*	*<0.001*
Anticipate		-	*<0.001*	*<0.001*	*<0.001*
PLER			-	*<0.001*	*<0.001*
PLEASE				-	*0.006*
LIP_{PI}					-
Virus					
B6C	-	*0.001*	*<0.001*	*<0.001*	*<0.001*
Anticipate		-	*0.207*	*<0.001*	*<0.001*
PLER			-	*<0.001*	*<0.001*
PLEASE				-	*0.001*
LIP_{PI}					-
NAFLD					
B6C	-	*0.001*	*0.142*	*0.037*	*0.003*
Anticipate		-	*<0.001*	*<0.001*	*<0.001*
PLER			-	*0.242*	*0.035*
PLEASE				-	*0.909*
LIP_{PI}					-
ALD					
B6C	-	*0.327*	*0.001*	*<0.001*	*<0.001*
Anticipate		-	*0.008*	*<0.001*	*<0.001*
PLER			-	*<0.001*	*<0.001*
PLEASE				-	*0.263*
LIP_{PI}					-

B6C: Baveno VI criteria

[a] Paired McNemar test

Table S7. Missed VNT (%) and spared endoscopy (%) rates by tests to spare endoscopy according to population screening in the whole population with cACLD [a] (2198 patients).

	B6C	Anticipate	PLER	PLEASE	LIP_{PI}	*p* [b]
Missed VNT	2.5	3.1	3.4	3.7	3.7	*0.683*
Comparison (p) [c]						
B6C	-	*0.727*	*0.508*	*0.344*	*0.344*	-
Anticipate		-	*1*	*0.791*	*0.774*	-
PLER			-	*1*	*1*	-
PLEASE				-	*1*	-
LIP_{PI}					-	-
Spared endoscopy	20.6	19.3	22.2	30.9	28.7	*<0.001*
Comparison (p) [c]						
B6C	-	*0.046*	*0.006*	*<0.001*	*<0.001*	-
Anticipate		-	*<0.001*	*<0.001*	*<0.001*	-
PLER			-	*<0.001*	*<0.001*	-
PLEASE				-	*0.001*	-
LIP_{PI}					-	-

B6C: Baveno VI criteria, VNT: varices needing treatment, cACLD: compensated advanced chronic liver disease
Colour legend: missed VNT (% in VNT): <5: safe, >10: unsafe, intermediate; spared endoscopy (%): <20: unsatisfactory, ≥20 & <30: fair, ≥30 & <40: satisfactory, >40: excellent
[a] cACLD definition: LSM ≥ 10 kPa and MELD score ≤ 8.6 [4]
[b] Paired Cochran test
[c] by McNemar test

Table S8. Missed VNT (%) and spared endoscopy (%) rates by tests to spare endoscopy according to individual screening as a function of secureness in the whole population.

	EB6C	LIP_{PI}	*p*
Missed VNT [b]			*p* [a]
All			
MELD <10	2.6	4.2	*0.146*
MELD ≥10	12.2	8.6	*0.595*
p [a]	*<0.001*	*0.069*	-
Virus			
MELD <10	*2.7*	*4.7*	*0.110*
MELD ≥10	*10.5*	*9.2*	*0.810*
p [a]	*0.011*	*0.133*	
NAFLD			
MELD <10	*0.0*	*2.3*	*0.894*
MELD ≥10	*8.3*	*6.7*	*0.565*
p [a]	*0.267*	*0.454*	
ALD			
MELD <10	*2.7*	*2.6*	*0.727*
MELD ≥10	*23.1*	*7.7*	*0.587*
p [a]	*0.015*	*0.347*	
Spared endoscopy [c]			*p* [b]
MELD <10	47.8	60.4	*<0.001*
MELD ≥10	16.3	18.1	*0.108*
p [a]	*<0.001*	*<0.001*	-

EB6C: expanded Baveno VI criteria, VNT: varices needing treatment
Colour legend: missed VNT (% in VNT): <5: safe, >10: unsafe, intermediate; spared endoscopy (%): <20: unsatisfactory, ≥20 & <30: fair, ≥30 & <40: satisfactory, >40: excellent
[a] Unpaired Fisher test or Chi^2
[b] Paired McNemar test

Table S9. INR distribution as a function of VNT status according to tests in patients with VNT of the whole population.

	p vs PLEASE [a]	**VNT ruled out**	***p*** [b]	***p*** [c]	**Indeterminate**
Population screening					
B6C	*0.083*	1.16±0.19	*0.085*	***0.021***	1.30±0.25
Anticipate	*0.124*	1.15±0.25	***0.038***	***0.002***	1.30±0.24
PLER	*0.135*	1.13±0.21	***0.013***	***0.001***	1.30±0.24
PLEASE	-	1.03±0.07	***<0.001***	***<0.001***	1.30±0.24
LIP_{PI}	*0.448*	1.06±0.09	***<0.001***	***<0.001***	1.30±0.24
Patient screening					
EB6C	***<0.001***	1.20±0.25	***0.007***	***<0.001***	1.31±0.24
LIP_{PI}	***<0.001***	1.13±0.13 [d]	***<0.001***	***<0.001***	1.33±0.25

B6C: Baveno VI criteria, EB6C: expanded Baveno VI criteria, Se: sensitivity, NPV: negative predictive value. Bold: significant p value. Indeterminate category between VNT ruled in or out, i.e. endoscopy required.

[a] Comparison of INR (in VNT ruled out) between test and PLEASE (reference as having the lowest INR level) by parametric Student *t* test

[b] Comparison between adjacent categories for each test by parametric Student *t* test

[c] Comparison between adjacent categories for each test by non-parametric Mann-Whitney test

[d] p=0.115 vs EB6C, p=0.012 vs LIP_{PI} for population screening

Table S10. Comparison of NAFLD cirrhosis criteria with LIP_{PI} and $BLIP_{PI}$ for individual screening in NAFLD of the core population.

	LIP_{PI}	***p*** [a]	**NCC**	***p*** [a]	**$BLIP_{PI}$**
Spared endoscopy rate (%)	73.8	*<0.001*	66.7	*<0.001*	74.6
Missed VNT rate (%)	5.8	*0.589*	4.8	*0.976*	4.97

NCC: NAFLD cirrhosis criteria

[a] McNemar test for spared endoscopy rate or Chi^2 for missed VNT rate

Table S11. Comparison of LIP_{PI} and $BLIP_{PI}$ for population screening in NAFLD of the whole population.

	LIP_{PI}	**$BLIP_{PI}$**	***p*** [a]
Derivation set			
Spared endoscopy rate (%)	37.9	40.3	*0.064*
Missed VNT rate (%)	6.1	4.1	*1*
Validation set			
Spared endoscopy rate (%)	40.0	39.5	*1*
Missed VNT rate (%)	4.0	4.0	*1*
Set comparison (p[b]):			
Spared endoscopy rate	*0.664*	*0.863*	-
Missed VNT rate	*1.0*	*1.0*	-
All			
VNT AUROC	0.804	0.822	*<0.001*
Spared endoscopy rate (%)			
LIP_{PI} vs $BLIP_{PI}$ without 100% specificity	38.6	40.0	*0.175*
LIP_{PI} vs $BLIP_{PI}$ with 100% specificity	38.6	40.8 [c]	*0.038*
Missed VNT rate (%)	5.4	4.1	*1*

[a] McNemar test or Delong test
[b] Fisher test
[c] vs PLEASE: 38.3% (p=0.113)

Table S12. Comparison of LIP_{PI} and $BLIP_{PI}$ for individual screening in NAFLD of the whole population.

	LIP_{PI}	**$BLIP_{PI}$**	***p a***
Derivation set			
Spared endoscopy rate (%)	72.2	73.6	*0.238*
Missed VNT rate (%)	5.3	4.9	*0.810*
Validation set			
Spared endoscopy rate (%)	77.1	74.3	*0.109*
Missed VNT rate (%)	6.8	5.1	*0.532*
Set comparison (p^b):			
Spared endoscopy rate	*0.211*	*0.923*	-
Missed VNT rate	*0.518*	*0.910*	-
All			
Spared endoscopy rate (%)			
LIP_{PI} vs $BLIP_{PI}$ without 100% specificity	73.8	73.8	*1*
LIP_{PI} vs $BLIP_{PI}$ with 100% specificity	73.8	74.6	*0.487*
Missed VNT rate (%)	5.8	4.97	*0.561*

[a] McNemar test for spared endoscopy rate or Chi^2 for missed VNT rate
[b] Fisher test for spared endoscopy rate or Chi^2 for missed VNT rate

Table S13. Comparison of $BLIP_{PI}$ performance and safety as a function of liver dysfunction for individual screening in NAFLD.

Patients	**All ITD**	***p*** [a]	**All**	***p*** [a]	**MELD score <10**	***p*** [b]	**MELD score ≥10**
Missed VNT (%)	4.75	*1*	4.97	*1*	4.98	*1*	4.76
Spared endoscopy (%)	70.8	*0.128*	74.6	*0.038*	79.7	*<0.001*	34.3

ITD: intention to diagnose
[a] Unpaired Chi^2 test
[b] Fisher test

Table S14. Missed VNT (%) and spared endoscopy (%) rates of tests to spare endoscopy according to population screening as a function of transaminase normality in the whole population.

	B6C	Anticipate	PLER	PLEASE	LIP_{PI}
Missed VNT:					
Normal transaminases	2.5	2.5	5.0	7.5	10.0
Increased transaminases	2.1	3.4	3.1	2.7	3.1
p [a]	*0.596*	*1*	*0.629*	*0.135*	*0.058*
Spared endoscopy:					
Normal transaminases	29.5	34.2	37.0	46.6	43.8
Increased transaminases	22.7	23.0	25.7	33.6	32.3
p [a]	*0.010*	*<0.001*	*<0.001*	*<0.001*	*<0.001*

B6C: Baveno VI criteria, VNT: varices needing treatment
Colour legend: missed VNT (% in VNT): <5: safe, >10: unsafe, intermediate; spared endoscopy (%): <20: unsatisfactory, ≥20 & <30: fair, ≥30 & <40: satisfactory, >40: excellent
[a] Fisher test

Table S15. Sensitivity analysis of VNT tests for population screening in the whole population [a]: influence of LSM reliability criteria [6].

	B6C	Anticipate	PLER	PLEASE	LIP_{PI}
Missed VNT rate (%):					
All (n=198)	1.5	2.5	3.5	3.0	5.1
Very reliable (n=67)	0.0	0.0	1.5	0.0	1.5
Reliable (n=106)	2.8	3.8	4.7	3.8	7.5
Poorly reliable (=25)	0.0	4.0	4.0	8.0	4.0
p [b]	*0.267*	*0.269*	*0.530*	*0.111*	*0.201*
p (linearity)	*0.569*	*0.151*	*0.379*	*0.037*	*0.289*
Spared endoscopy rate (%):					
All (n=1147)	19.0	21.7	22.7	31.3	30.3
Very reliable (n=306)	14.4	16.7	19.0	23.5	23.9
Reliable (n=729)	21.0	24.0	24.1	34.6	33.1
Poorly reliable (=112)	18.8	20.5	23.2	31.3	29.5
p [b]	*0.047*	*0.031*	*0.189*	*0.002*	*0.013*
p (linearity)	*0.072*	*0.081*	*0.145*	*0.009*	*0.038*

VNT: varices needing treatment, LSM: liver stiffness measurement

[a] These results were able to be analysed in the whole population due to the similarity of populations.

[b] Unpaired Chi^2 test for difference

Table S16. Cut-offs of LIP and BLIP scores for different missed VNT rates <5% according to population or individual screening.

	$BLIP_{PI}$			LIP_{PI} <		LIP_{INR} <	
Cut-offs	**Sensitivity**	**NPV**	**Spe**	**Sensitivity**	**NPV**	**Sensitivity**	**NPV**
Viral CLD	-	-	-	0.0603	0.1056	0.0607	0.1102
NAFLD [a]	<0.044	<0.11	>0.56	0.04505	0.1104	0.0364	0.1142
	(<0.048)	(<0.1125)	(>0.53)				
ALD	-	-	-	0.063	0.0742	0.066	0.076
All aetiology	-	-	-	0.05016	0.0985	0.0564	0.1007

NPV: negative predictive value, Spe: 100% specificity
[a] cut-off for $BLIP_{INR}$ in brackets

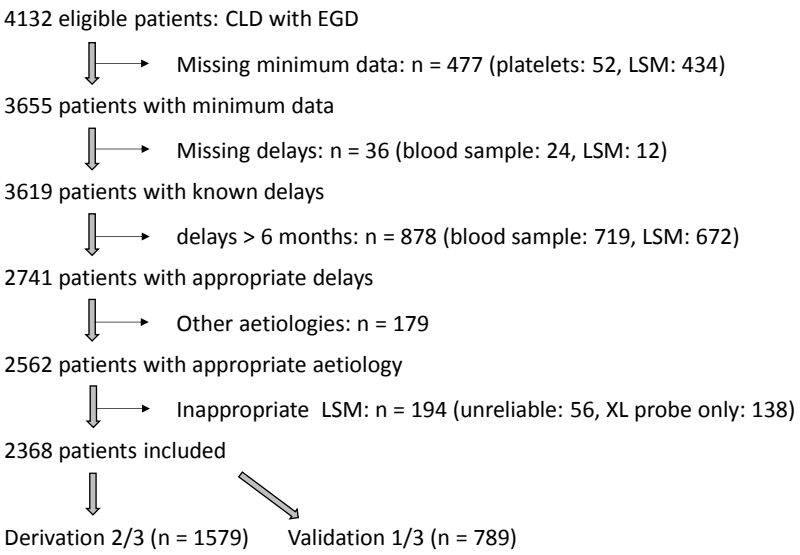


Figure S1. Patient flowchart.

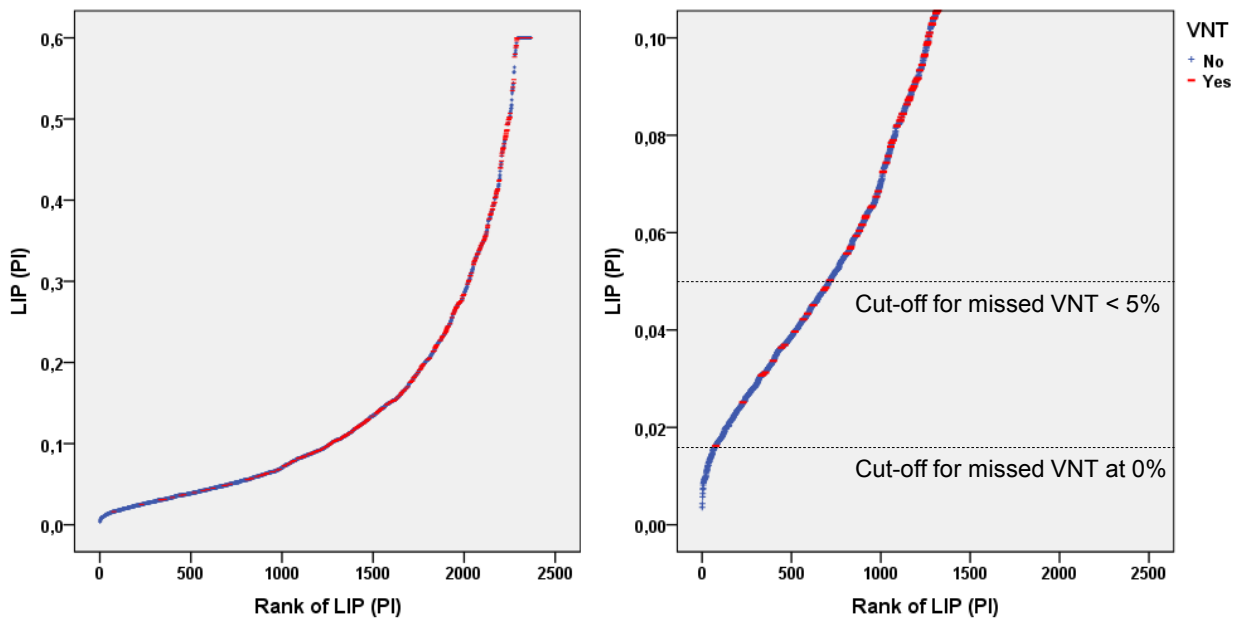


Figure S2. Distribution of LIP_{PI} score in the whole population with a magnification in the right panel showing the unique cut-off for predefined missed VNT rates.

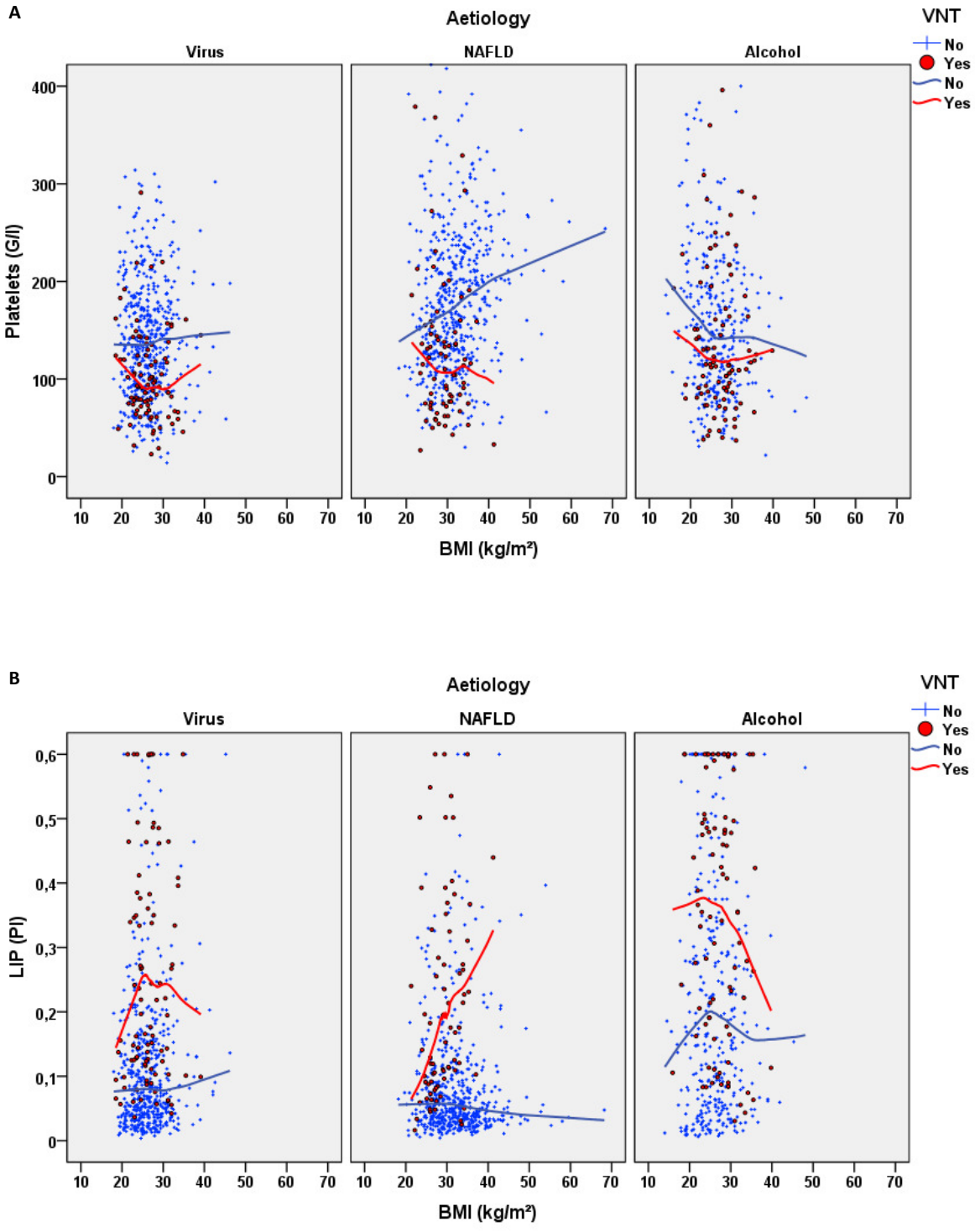


Figure S3. BMI influence on platelets and LIP as a function of VNT according to aetiology in the whole population. Curves from non-linear regression (LOWESS).

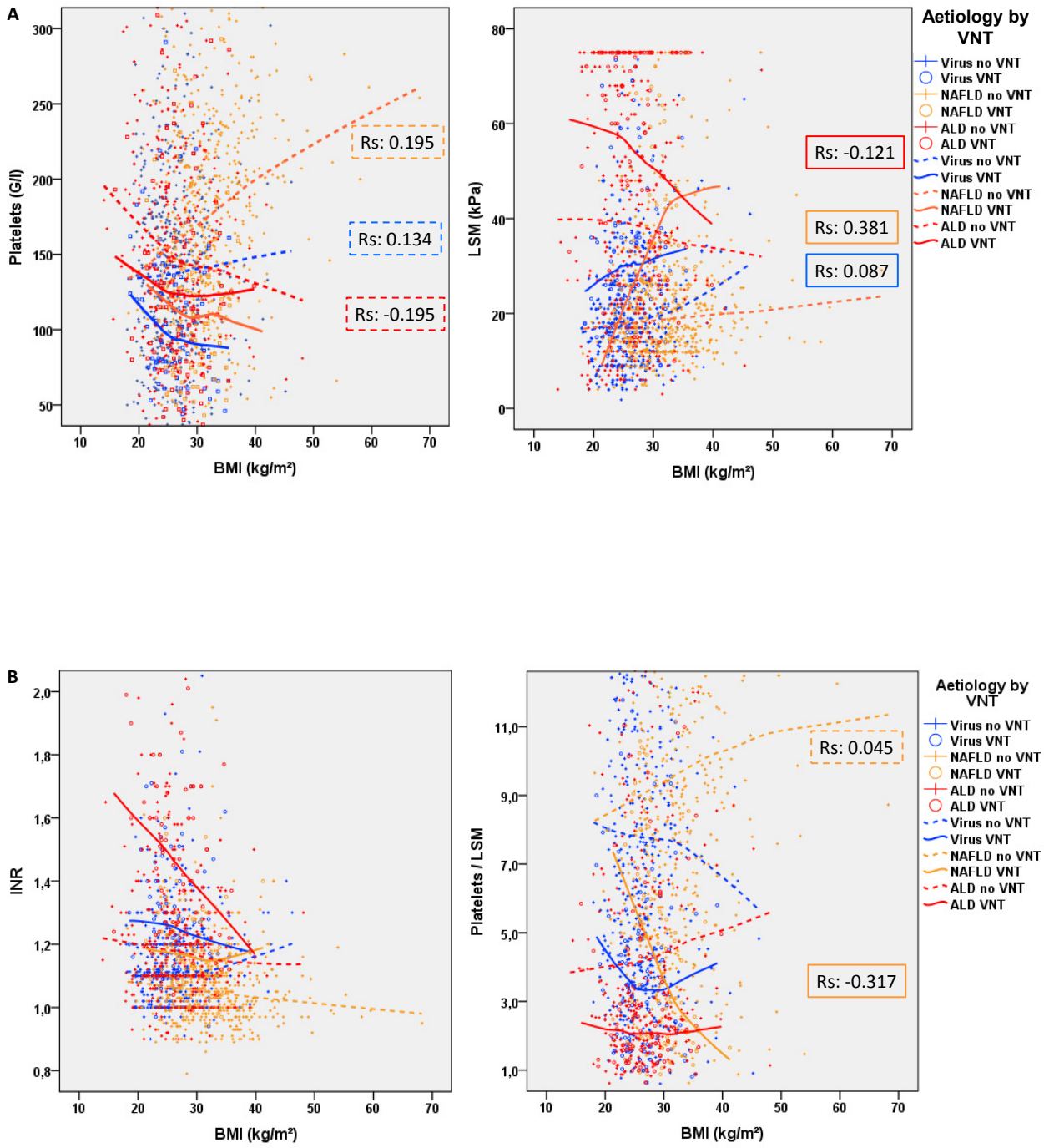


Figure S4. BMI influence on markers included in LIP_{PI} as a function of VNT status and aetiology in the whole population. Curves from non-linear regression (LOWESS).

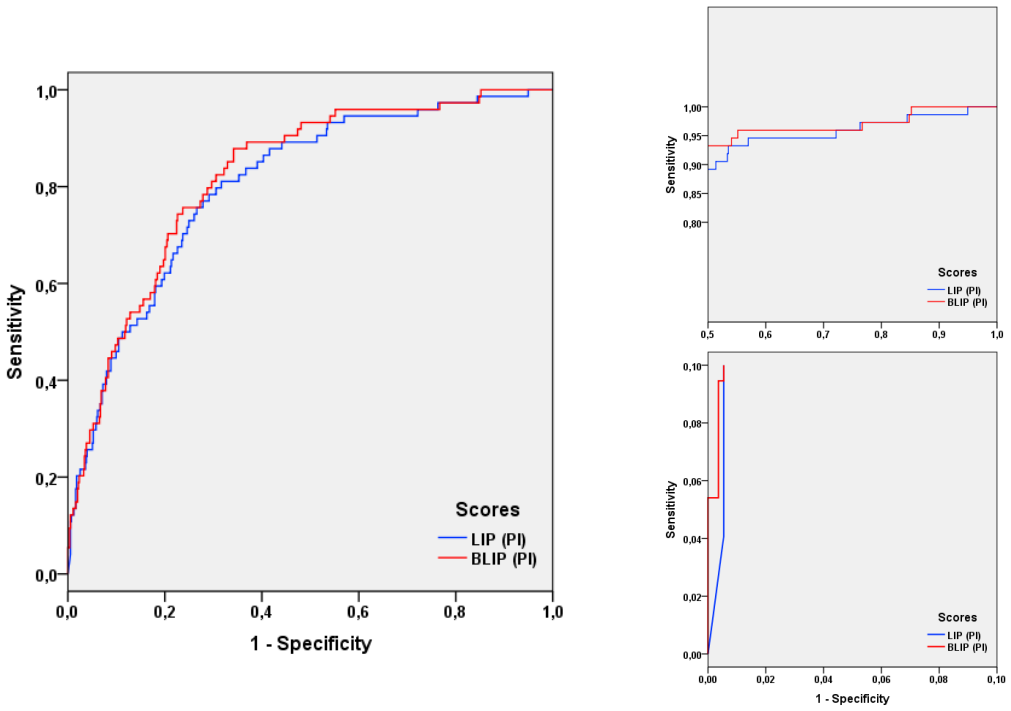


Figure S5. Comparison of ROC curves between LIP_{PI} and $BLIP_{PI}$ in the whole population. The right panels are magnifications showing patient subsets with sensitivities ≥95% and specificities at 100% for VNT.

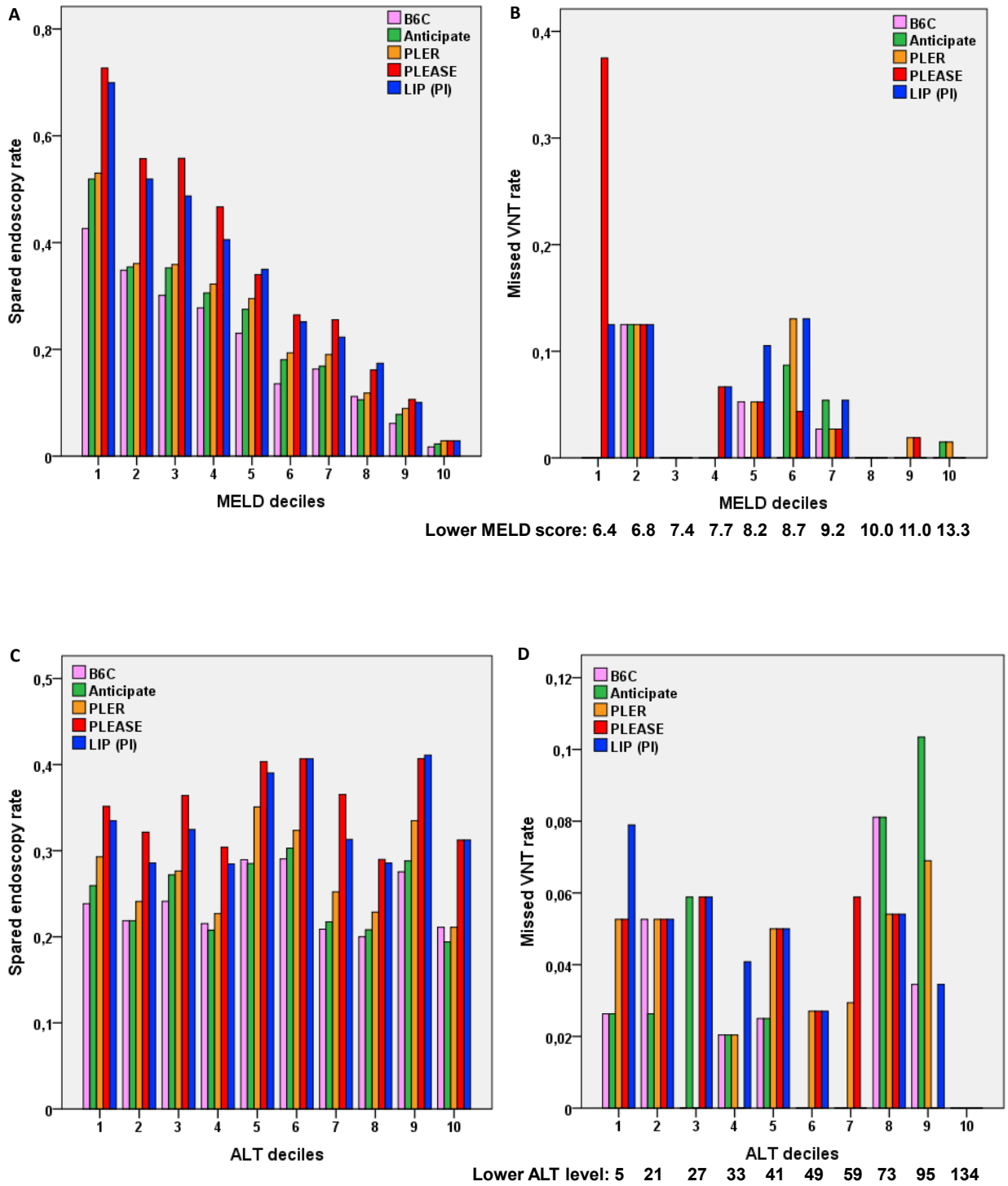
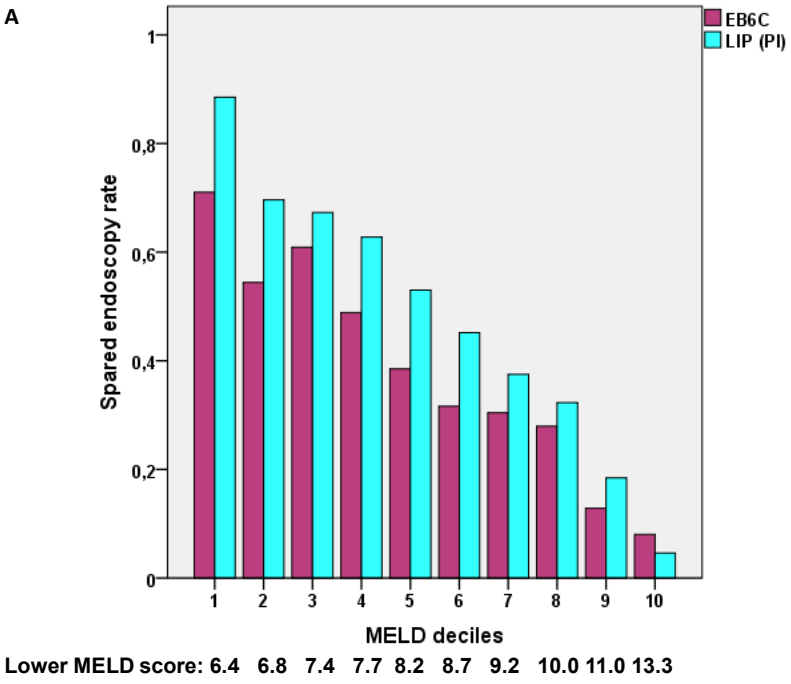


Figure S6. Sensitivity analysis of tests. Panel A: spared endoscopy rate as a function of MELD deciles. Panel B: missed VNT rate as a function of MELD deciles. Panel C: spared endoscopy rate as a function of ALT deciles. Panel D: missed VNT rate as a function of ALT deciles.



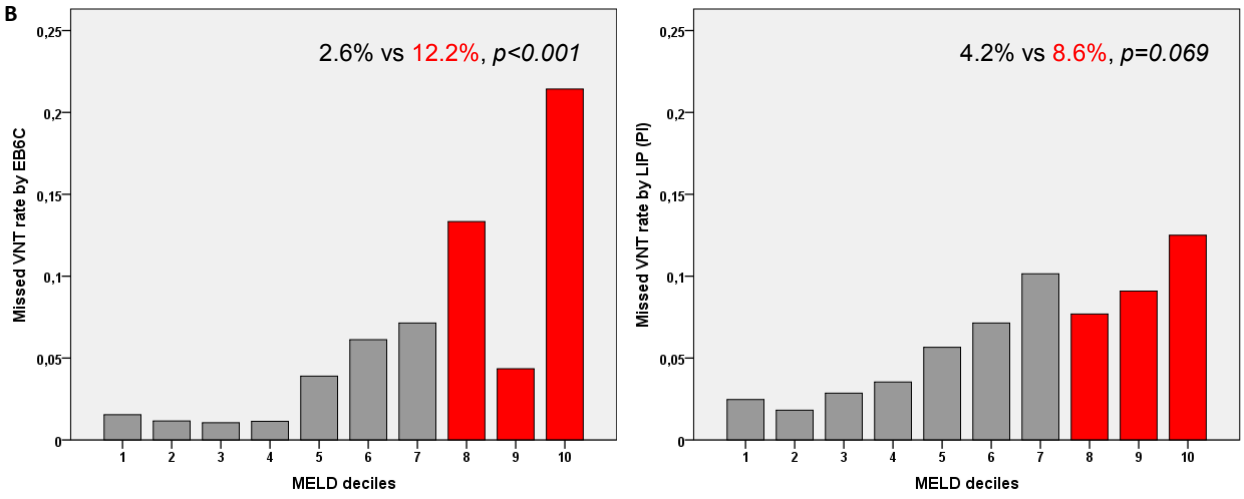


Figure S7. Comparison of performance (panel A) or safety (panel B) between LIP_{PI} and the expanded Baveno VI criteria (EB6C) for individual screening as a function of MELD deciles in the whole population.

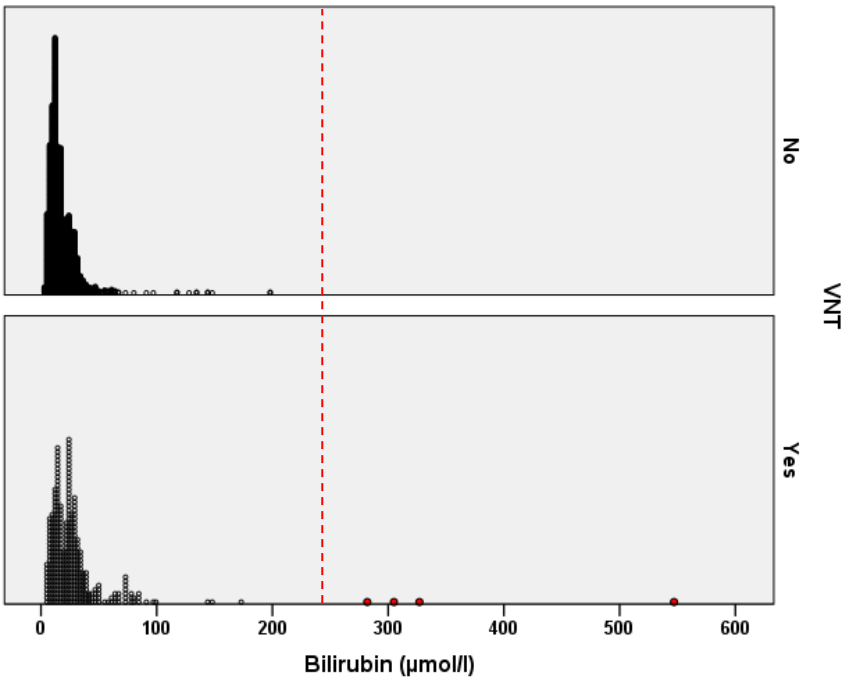


Figure S8. Specificity at 100% for VNT (red dots) of bilirubin in the whole population.

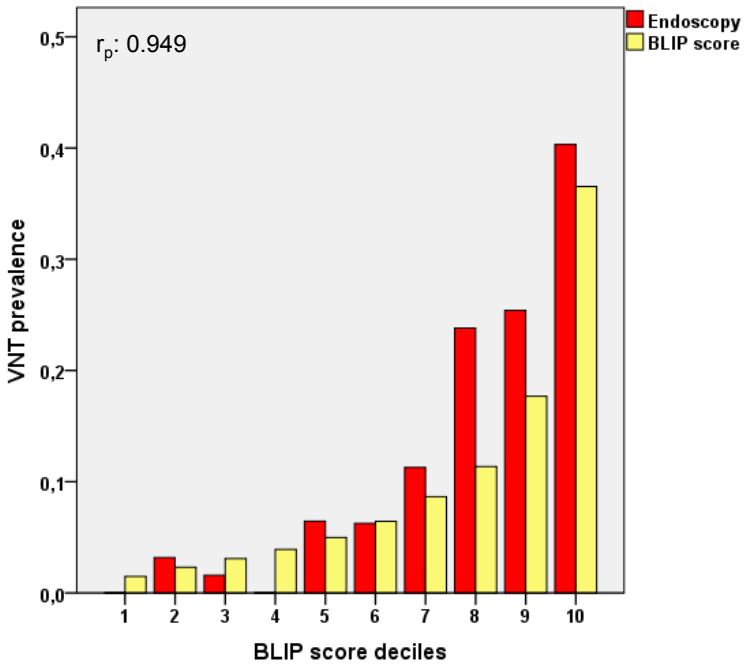


Figure S9. Calibration of $BLIP_{PI}$ score to quantify the VNT probability in NAFLD of the whole population.

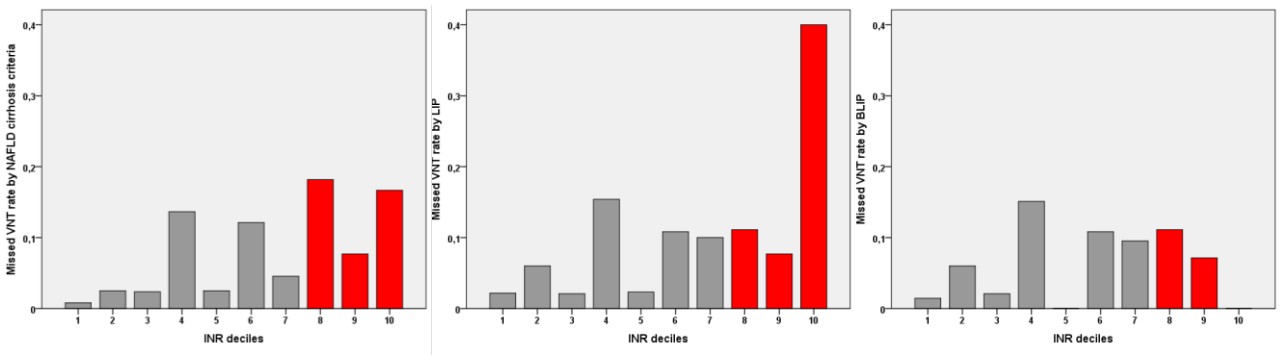


Figure S10. Safety of tests for individual screening as a function of INR deciles in NAFLD of the whole population. Severe CLD is indicated by red bars. INR is used here due to some missing MELD score values (missing creatinine) in NAFLD.

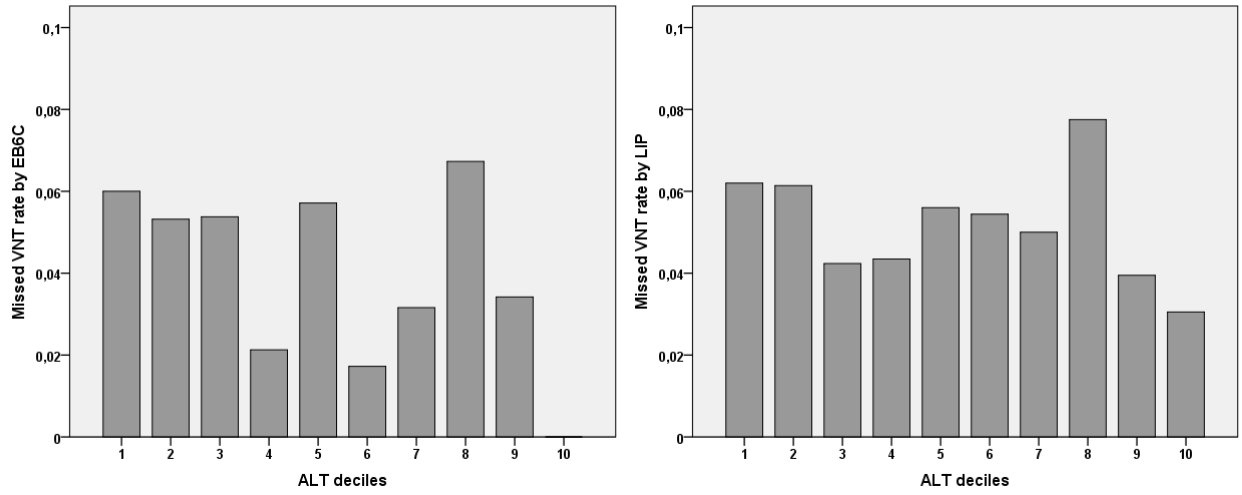


Figure S11. Sensitivity analysis of tests: missed VNT rate as a function of ALT deciles of tests for individual screening in the whole population. Expanded Baveno VI criteria (EB6C) and LIP_{PI}.

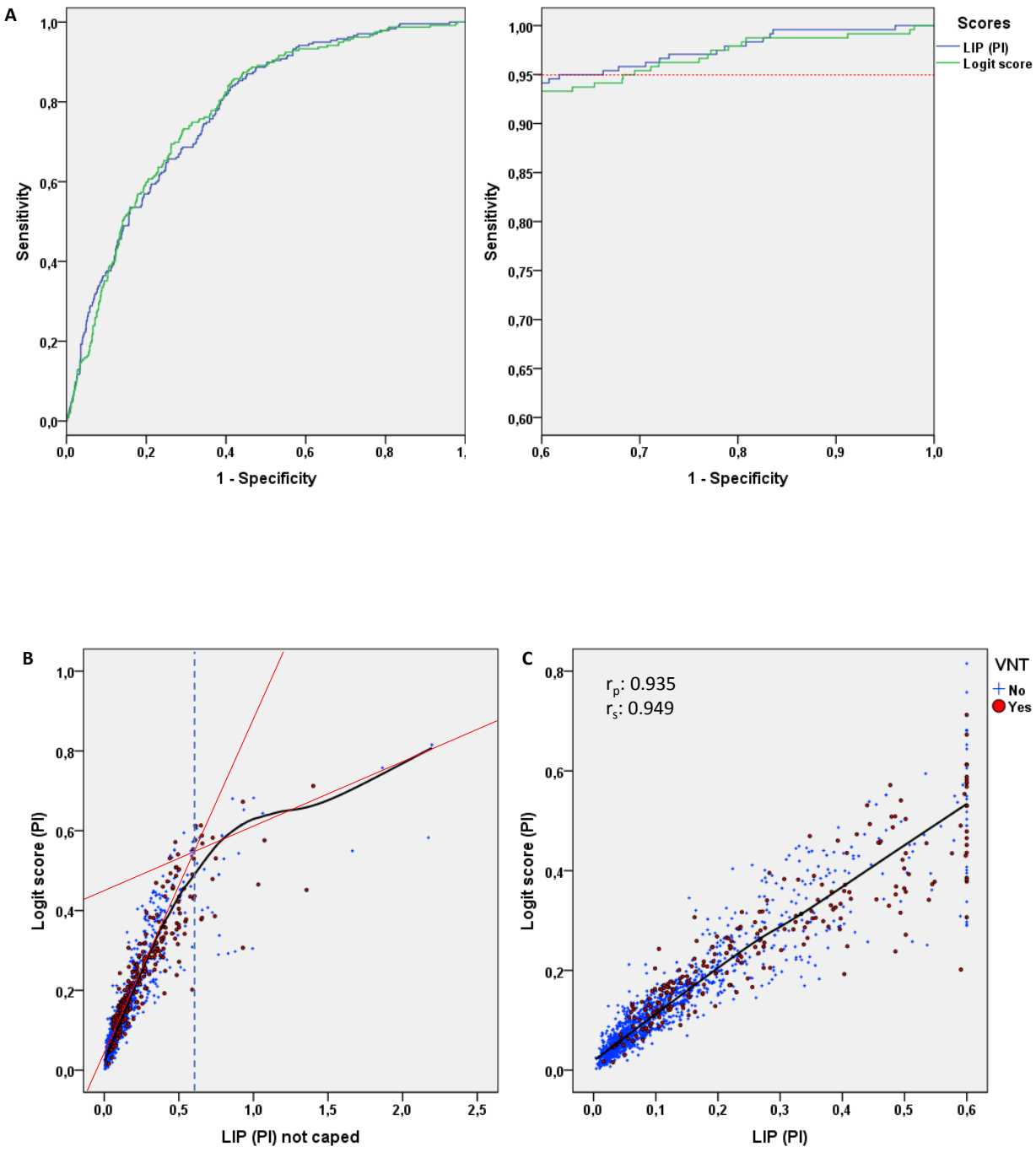


Figure S12. Comparison between LIP_{PI} score and the logit score including the LIP_{PI} markers in the whole population. Panel A: ROC curves. Panels B and C: scatter plots. Panel B validates LIP_{PI} capped to 0.6.

CANDILLIER Camille

Dépistage individuel et en population des varices nécessitant un traitement par un test simple, sûr et performant : LIP

RÉSUMÉ

Introduction : Plusieurs tests ont été développés pour le dépistage des varices nécessitant un traitement (VNT) mais pour différentes stratégies de dépistage : individuelle ou en population. Nous avons cherché à développer un estimateur simple permettant de quantifier le risque de VNT et d'éviter l'endoscopie, tout en manquant <5% de VNT, adapté aux 2 stratégies de dépistage dans les principales étiologies.

Méthodes : 2 368 patients atteints d'hépatopathie chronique fibrosante ont été inclus dans une population de dérivation et de validation. Les résultats étant similaires, sont donnés pour l'ensemble de la population. Les caractéristiques des patients étaient les suivantes : âge : 59±11 ans ; sexe masculin : 63,5% ; étiologies : virus : 50,2%, NAFLD : 28,9%, alcool : 20,9% ; MELD : 9,5±3,0 ; élastométrie hépatique (LSM) ≥10kPa : 92,8% ; VNT : 15,2%. Les principaux prédicteurs indépendants des VNT étaient les plaquettes, le taux de prothrombine (TP) et l'élastométrie hépatique. Leurs interactions ont conduit à la construction des scores, LIP : (LSM*45)/(PI*plaquettes), et BLIP : LIP ajusté à l'IMC dans la NAFLD. Les scores ont été définis pour le dépistage des VNT en population (sensibilité ≥ 95%) ou le dépistage individuel (valeur prédictive negative ≥ 95%). Le critère de performance était le taux d'endoscopie évité.

Résultats : 1) Scores diagnostiquant les VNT. Les AUROCs étaient, PLER : 0.767 Anticipate : 0.773 (p=0.059 vs précédent), LIP : 0.779 (p=0.136), PLEASE : 0.789 (p=0.196). 2) La performance du dépistage en population était (par ordre croissant, (avec le taux de VNT manqué), critères de Baveno 6 : 23,9 % (2,5 %), Anticipate : 24,5 %, p=0,367 vs précédent (3,3 %), PLER : 27,3 %, p<0,001 (3,6 %), LIP : 33,4 %, p<0,001 (4,2 %), PLEASE : 35,2 %, p=0,006 (3,6 %). Dans le cas de la NAFLD, la performance était, LIP : 38,6%, BLIP : 40,8%, p=0,038. 3) La performance du dépistage individuel était : LIP : 54,1%, critères de Baveno6 étendus : 42.7%, p<0.001. Dans la NAFLD, la performance était, BLIP : 74,6%, NAFLD-cirrhosis criteria : 66.7%, p<0.001.

Conclusion : LIP combine simplicité, performance et sécurité dans chaque étiologie. Dans la NAFLD, le BLIP est plus performant que les autres tests.

Mots-clés : Hypertension portale ; varices œsophagiennes ; plaquettes ; élastométrie hépatique ; taux de prothrombine

Individual and population screening of varices needing treatment by a simple, safe and accurate test : LIP

ABSTRACT

Background and Aims: Several tests have been developed to screen varices needing treatment (VNT) but for different screening strategies: individual or population. We aimed to develop simple estimators to quantify VNT risk and to spare endoscopy, while missing <5% of VNT, adapted to different screenings in the main aetiologies.

Methods: 2,368 patients with chronic liver disease were included in derivation and validation sets. Results, being similar therein, are given in the whole population. Patient characteristics were, age: 59±11 years; male sex: 63.5%; aetiologies: virus: 50.2%, NAFLD: 28.9%, alcohol: 20.9%; MELD: 9.5±3.0; liver stiffness measurement (LSM) ≥10kPa: 92.8%; VNT: 15.2%. The main independent VNT predictors were platelets, prothrombin index (PI) and LSM. Their interactions led to score construction, LIP: (LSM*45)/(PI*platelets), and BLIP: BMI-adjusted LIP in NAFLD. Scores were categorised either for population (VNT sensitivity ≥95%) or individual (negative predictive value ≥95%) VNT screening. The performance criterion was spared endoscopy rate.

Results: 1) Scores diagnosing VNT. AUROCs were, PLER: 0.767 Anticipate: 0.773 (p=0.059 vs previous), LIP: 0.779 (p=0.136), PLEASE: 0.789 (p=0.196). 2) Population screening performance was in increasing order (with missed VNT rates), Baveno$_6$ criteria: 23.9% (2.5%), Anticipate: 24.5%, p=0.367 vs previous (3.3%), PLER: 27.3%, p<0.001 (3.6%), LIP: 33.4%, p<0.001 (4.2%), PLEASE: 35.2%, p=0.006 (3.6%). In NAFLD, performance was, LIP: 38.6%, BLIP: 40.8%, p=0.038. 3) Individual screening performance was, LIP: 54.1% expanded Baveno$_6$ criteria: 42.7%, p<0.001. In NAFLD, performance was, BLIP: 74.6%, NAFLD-cirrhosis criteria: 66.7%, p<0.001.

Conclusion: LIP combined simplicity, performance and safety in each aetiology. In NAFLD, BLIP outperformed other tests.

Keywords : portal hypertension; esophageal varices; platelets; liver elastometry; prothrombin time

FACULTÉ DE SANTÉ
UNIVERSITÉ D'ANGERS