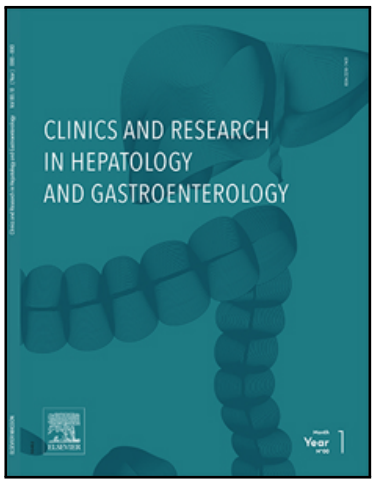


Individual and population screening of varices needing treatment by a simple, safe and accurate test

Federico Ravaioli , Arthur Berger , Oana Farcau , Antonio Colecchia , Horia Stefanescu , Camille Candillier , Pierre Nahon , Christophe Bureau , Nathalie Ganne-Carrié , Annalisa Berzigotti , Victor de Ledinghen , Salvatore Petta , Paul Calès , multicentric groups , Bologna , Cluj , Bondy , Toulouse , Bern , Bordeaux , Palermo

PII: S2210-7401(23)00048-7
DOI: https://doi.org/10.1016/j.clinre.2023.102123
Reference: CLINRE 102123

To appear in: *Clinics and Research in Hepatology and Gastroenterology*

Please cite this article as: Federico Ravaioli , Arthur Berger , Oana Farcau , Antonio Colecchia , Horia Stefanescu , Camille Candillier , Pierre Nahon , Christophe Bureau , Nathalie Ganne-Carrié , Annalisa Berzigotti , Victor de Ledinghen , Salvatore Petta , Paul Calès , multicentric groups , Bologna , Cluj , Bondy , Toulouse , Bern , Bordeaux , Palermo , Individual and population screening of varices needing treatment by a simple, safe and accurate test, *Clinics and Research in Hepatology and Gastroenterology* (2023), doi: https://doi.org/10.1016/j.clinre.2023.102123

Highlights

Endoscopy to screen varices needing treatment can be avoided by non-invasive markers in a minority of patients. A new LIP test, designed for mains etiologies, spared 33% of endoscopies in population screening and 54% in individual screening. A new BLIP test, designed for NAFLD, spared 41% of endoscopies in population screening and 75% in individual screening. This new strategy, adapted to clinical practice, should improve screening adherence which is currently poor.

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

Federico Ravaioli[1]*, Arthur Berger[2]*, Oana Farcau[3], Antonio Colecchia[1-4], Horia Stefanescu[3], Camille Candillier[2], Pierre Nahon[5], Christophe Bureau[6], Nathalie Ganne-Carrié[5], Annalisa Berzigotti[7], Victor de Ledinghen[8], Salvatore Petta[9], Paul Calès[2], and multicentric groups[10]: VO-VCO, ANRS CO12 CIRVIR, M116 and VEB6. Angers[2], Bologna[1], Cluj[3], Bondy[5], Toulouse[6], Bern[7], Bordeaux[8], Palermo[9]

[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]Department of Gastroenterology, University of Modena & Reggio Emilia and Azienda Ospedaliero-Universitaria di Modena, Italy.

[5]Hepatology Department, Avicenne Hospital, Bobigny, Paris hospitals; INSERM UMR-1162, Paris Sorbonne University, France

[6]Hepato-Gastroenterology Department, Purpan University Hospital, Toulouse, France

[7]Hepatology, Swiss Liver Center, Visceral Surgery and Medicine Clinic (UVCM), Inselspital, University of Bern, Switzerland

[8]Hepatology Department, Haut-Lévêque Hospital, Bordeaux, France

[9]Department of Gastroenterology and Hepatology, PROMISE, University of Palermo, Palermo, Italy

[10]All authors and centres, merged in the VEB6 group, are listed in acknowledgments.

*Authors sharing first authorship.

Funding statement: This research did not receive any specific funding from any funding agency in the commercial or non-profit sectors. The study was financed by the institutional funds of the individual authors involved.

Conflict of interest disclosure: none related to the topic

Ethics approval statement: obtained for all subpopulations (details in Supplemental Material)

Data Availability Statement:

1\Patient data: Data entry sheet is available on simple request; individual data are available for validation studies or meta-analyses with predefined strategy. Contact the corresponding author.

2\Test formulae: They are fully detailed in the Supplemental Material. Their calculation is available for individual patient at an URL stated in the manuscript.

3\Additional data: Several additional data are detailed in the Supplemental material and in the Complementary material whose URL is indicated in the Supplemental Material.

ADDITIONAL INFORMATION

Short Title: VNT diagnosis by LIP

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, 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],

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

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

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

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

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

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

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

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

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

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

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

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

Marseille; [18]Hôpital Claude Huriez, Service d'Hépatologie, Lille; [19]Hôpital St André, Service d'Hépatologie, Bordeaux; [20]Hôpital Hôtel Dieu, Service d'Hépatologie, Clermont-Ferrand; [21]AP-HP, Hôpital Saint-Antoine, Service d'Hépatologie, Paris; [22]AP-HP, Hôpital Henri Mondor, Service d'Hépatologie, Créteil; [23]AP-HP, Hôpital Tenon, Service d'Hépatologie, Paris; [24]AP-HP, Hôpital Paul Brousse, Service d'Hépatologie, Villejuif; [25]Hôpital Trousseau, Unité d'Hépatologie, CHRU de Tours; [26]Hôpital d'Aix-En-Provence, Service d'Hépatologie, Aix-En-Provence; [27]Hôpital de la Côte de Nacre, Service d'Hépatologie, Caen; [28]AP-HP, Groupe Hospitalier de La Pitié-Salpêtrière, Service d'Hépatologie, Paris; [29]CHU Le Mans, Service d'Hépatologie, Le Mans; [30]CHU de Poitiers, Service d'Hépatologie, Poitiers; [31]Institut Mutualiste Montsouris, Service d'Hépatologie, Paris; [32]Hôpital Amiens Nord, Service d'Hépatologie, Amiens; [33]Hôpital Robert Debré, Service d'Hépatologie, Reims; [34]Hôpital Foch, Service d'Hépatologie, Suresnes; [35]Hôpital Jean Minjoz, Service d'Hépatologie, Besançon. France.

We also thank Kevin L. Erwin for writing assistance (English proofreading) 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.

Abbreviations:

ALD: alcoholic liver disease

AUROC: area under the receiver operating characteristic

cACLD: 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

PLER: platelets/liver elastometry ratio

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

PPV: positive predictive value

VCTE: vibration-controlled transient elastometry

VNT: varices needing treatment

Correspondence:

Paul Calès, Hepatology, CHU, 49933 Angers Cedex 09, France. Tel: (33) 2 41 35 34 10, Fax: (33) 2 41 35 41 19, E-mail: paul.cales@univ-angers.fr

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

Item count:

4992 words (5,000 words maximum, including title page, abstract, lay summary and main text), number of figures and tables: 4 + 4 (≤8), number of references: 31 (≤50), number of supplementary files for online publication: 1 including 16 tables and 11 figures. Title: 101 characters (<130 characters including spaces). Short title: 20 characters.

Author contributions:

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

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

Oana Farcau: critical revision of the manuscript, data collection

Antonio Colecchia: study supervision, critical revision of the manuscript, data collection

Horia Stefanescu: critical revision of the manuscript, data collection

Camille Candillier: critical revision of the manuscript, data collection, literature review

Pierre Nahon: critical revision of the manuscript, data collection

Christophe Bureau: critical revision of the manuscript, data collection

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

Annalisa Berzigotti: critical revision of the manuscript, data collection

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

Salvatore Petta: critical revision of the manuscript, data collection

VEB6 group: data collection

Paul Calès: study concept and design, analysis and interpretation of data, drafting of the manuscript, data collection, statistics, guarantor of the article

Authors' e-mails:

f.ravaioli@unibo.it

arthur.berger@chu-bordeaux.fr

oana.farcau@gmail.com

antonio.colecchia@unimore.it

horia.stefanescu@irgh.ro

camille.candillier@yahoo.fr

pierre.nahon@aphp.fr

bureau.c@chu-toulouse.fr

nathalie.ganne@aphp.fr

Annalisa.Berzigotti@insel.ch

victor.deledinghen@chu-bordeaux.fr

salvatore.petta@unipa.it

paul.cales@univ-angers.fr

Writing assistance:

Kevin L. Erwin for English proofreading (institutional support)

Abstract

Background: Several tests have been developed to screen VNT in different screening settings. We aimed to develop simple estimators to quantify VNT risk and spare endoscopy while missing <5% of VNT, adapted to different screenings in the main etiologies.

Methods: 2,368 patients with chronic liver disease were included. The main 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 categorized either for population (VNT sensitivity ≥95%) or individual (negative predictive value ≥95%) VNT screening.

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 rate), 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, LIP 38.6%, BLIP 40.8%, p=0.038. 3) Individual screening performance was, expanded Baveno$_6$ criteria: 42.7%, LIP 54.1%, p<0.001. In NAFLD, performance was, NAFLD-cirrhosis criteria: 66.7%, BLIP 74.6%, p<0.001.

Conclusion: LIP combined simplicity, performance and safety in each etiology. In NAFLD, BMI-adjusted LIP outperformed other tests.

Words: 250 (≤250)

Keywords:

Portal hypertension; oesophageal varices; platelets; liver elastometry; prothrombin time

INTRODUCTION

The original Baveno VI criteria enabled the wide clinical acceptance of non-invasive tests for varices needing treatment (VNT) [1]. Baveno VI criteria have several limitations. First, although largely validated [2-6], their clinical impact has been judged modest, providing a spared endoscopy rate of only about 20% [3, 7]. Second, two main definitions of missed VNT are used [8]. One is based on the probability of missing <5% VNT in patients left without endoscopy. This negative predictive value (NPV) is adapted to an individual screening strategy. Thus, the expanded Baveno VI criteria [9] and non-alcoholic fatty liver disease (NAFLD) cirrhosis criteria [10] 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 adapted to a population screening strategy. Some tests were constructed for individual screening and others for population screening. However, the screening category needed to be explicitly stated and no study compared the two strategies. 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, this selective strategy restricted to cACLD spared less endoscopies than a global strategy unrestricted by liver severity [11]. Nonetheless, that global strategy must be secured by avoiding missed 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 them in tests to spare endoscopy [11]. First, the platelets/liver elastometry ratio, called PLER, and the platelets/liver elastometry ratio adjusted on etiology, 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 offered a 35% spared endoscopy rate, but its complex formula requires

a specific calculator. Finally, the Baveno VI criteria, Anticipate and the expanded Baveno VI criteria, had limits in non-viral etiologies [11].

NAFLD is an increasingly predominant etiology. 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 for individual screening [10]. We observed however a particular role of platelets and BMI for VNT prediction in a previous study in NAFLD [12].

Our main objective was to develop a simple bedside test for all main etiologies of chronic liver disease by combining the advantages of previous tests (performance and secureness) and avoiding their limits. Our secondary objectives were to a) design and compare this test for individual and population screenings, b) quantify the VNT risk, c) refine safety criteria for missed VNT and c) adapt this test to NAFLD.

PATIENTS AND METHODS

Participants

In this post-hoc analysis of prospectively collected data, patients' clinical information with CLD was collected from centres participating in several studies wherein VNT was usually the main outcome and liver stiffness measurement (LSM) by transient elastography (FibroScan) was 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 etiology (alcoholic CLD (ALD), NAFLD, hepatitis B or C virus) were eligible for inclusion if they had undergone an endoscopy to determine oesophageal varices (EV) size. The four minimum inclusion criteria were 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. The exclusion criteria were complications (ascites, gastrointestinal bleeding) and interventional treatment for portal hypertension (PHT) (transjugular intrahepatic portosystemic shunt, band ligation or sclerotherapy of EV) and incomplete data. 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 [11]) 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 etiology.

Methods

Data collection

Clinical data - The main clinical data were age, sex, height, body weight and etiology. The main laboratory data were liver function tests whose prothrombin index (PI), 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 [13].

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

Liver Stiffness Measurement (LSM) - All LSMs were performed by experienced operators using vibration-controlled transient elastography (VCTE), specifically M probe-equipped FibroScan devices (Echosens, Paris, France). Technical characteristics are detailed elsewhere [14].

Definitions

Objectives - The primary objective was to develop a simple bedside test for all the main etiologies of chronic liver disease. This new test needed to match the performance and safety of previously published tests, especially the secureness of our previous platelets/liver elastometry ratio (PLEASE) test.

Secondary goals included developing a test capable of quantifying VNT risk (probability), determining cut-offs for individual and population screenings, and adapting the test to the specifics of NAFLD. Ancillary goals included achieving 100% specificity for VNT and refining qualitative safety parameters for the missing VNT.

Outcome - The main outcome was varices needing treatment (VNT), defined as large EV (grade 2 or 3, i.e. a diameter ≥5 mm [15]).

Outcome measurements - The primary outcome measurements were the performance (spared endoscopy rate) and safety (missed VNT rate) of the tests developed. 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. The missed VNT rate was the ratio of missed VNT on VNT in population screening or spared endoscopies in individual screening [8]. 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, 10]. The negative likelihood ratio (LR-) was a secondary outcome measurement (details in Complementary Material).

Safety criteria - We evaluated three criteria. First, the classical *quantitative* criterion is a missed VNT rate <5%. Second, we evaluated the *qualitative* safety by the level of liver dysfunction as a function of VNT status. The principle of *qualitative* safety is to privilege the test having the lowest liver dysfunction in missed VNT and discard (or limit) a test inducing missed VNT in severe CLD. Indeed, the incidence and mortality of variceal bleeding grows with liver severity [16]. Thus, *secureness* defined a *secured* test without missed VNT in CLD with poor liver function (MELD score ≥10 or INR ≥1.24) [11] called *severe* CLD hereafter. The level of liver dysfunction in missed VNT was called *functional* safety.

Comparators - The new tests were compared first for population screening: Baveno VI criteria [1] based on cut-offs of platelets and LSM; Anticipate (a logit function of platelets and LSM) [17]; and PLER (platelets/LSM ratio) and PLEASE (PLER adjusted to etiology, sex and INR) [11]. Then, they were compared in the individual screening setting with the expanded Baveno VI criteria [9] and the NAFLD cirrhosis criteria [10].

VNT diagnostic estimators - An estimator was called a *score* when it provided a numerical variable quantifying precisely the VNT probability. An estimator was called a *test* when it was categorized 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.

Score development

Details on score construction in the derivation set are provided in the Supplemental Material.

All etiologies - The (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 score distribution as a function of VNT is shown in Figure S2. In the Supplemental Material, we describe the corresponding LIP_{INR} score using INR instead of PI.

NAFLD - We added BMI in the LIP formula to obtain the BLIP score for NAFLD: (LSM*45*30)/(PI*platelets*BMI). BLIP had a larger subset with ≥95% VNT sensitivity than LIP and reached 100% specificity (Figure S3).

Statistics

Quantitative variables were expressed as mean ± standard deviation and compared using the Student t-test or analysis of variance. 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 [18] and Liver FibroSTARD [19] statements and analysed on a partial intention-to-diagnose basis. Thus, all patients were included irrespective of reliability criteria of VCTE [20] (except in one NAFLD subpopulation [10]). 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 [21]. The main statistical analyses were performed using SPSS version 18.0 (IBM, Armonk, NY, USA).

RESULTS

Patients

The study included 2368 patients with CLD. Because there were no significant differences between the characteristics of derivation and validation populations (Table S1), the following results are those of the whole population. Nearly two-thirds of the patients were men. Viral-related chronic liver disease was the most frequent etiology (50%), then NAFLD (29%) and alcohol-related (21%). LSMs ≥10 kPa were observed in 92.8% of patients. The prevalence of cACLD was estimated at 43.8%. The characteristics as a function of etiologies are reported in Table S2.

LIP in all etiologies

Scores estimating VNT probability

Score calibration

The LIP [(LSM*45)/(PI*platelets)] score was well calibrated to VNT prevalence: the scatter plots of LIP and VNT as a function of LIP percentiles showed the high correlation between estimated and observed prevalences (Figure 1); and estimated mean prevalence was not significantly different from the VNT prevalence (14.8% vs 15.2%, p=0.566).

Score discrimination

VNT discrimination by scores was evaluated with AUROC. Figure 2A clearly shows that the AUROC of the LIP score was significantly higher ($p<0.001$) than those of its composite markers (details in Table S3). In the derivation population, the AUROC of the LIP score was significantly higher than that of PLER (platelets/liver elastometry ratio) ($p<0.001$) and lower

than that of PLEASE (PLER adjusted to etiology, sex and INR) score (p=0.015) but not significantly different from Anticipate (Table 2).

Tests sparing endoscopy

LIP

Figure 2A clearly shows that the LIP score [(LSM*45)/(PI*platelets)] 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 S4. 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.

Comparison of tests

Figure 2B shows that the LIP score provided a larger subset of patients with VNT sensitivity >95% than the published tests did, especially vs the PLEASE (PLER adjusted to etiology, sex and INR) score in contrast with the lesser AUROC. This conferred an endoscopy-sparing advantage for the LIP score compared to other scores.

Population screening - The spared endoscopy rates were, in increasing order: Baveno VI criteria: 23.9%, Anticipate: 24.5% (p=0.367 vs previous), PLER (platelets/liver elastometry ratio): 27.3% (p<0.001), LIP [(LSM*45)/(PI*platelets)]: 33.4% (p<0.001), PLEASE (PLER adjusted to etiology, sex and INR): 35.2% (p=0.006) (Table 3). Performance comparisons are detailed in Table S5. 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 4.0

vs 5.2%, respectively (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 S4), was eliminated by the secureness rule.

Sensitivity analysis

The influence of liver function, etiology, inflammation and LSM reliability are detailed in the Supplemental Material. Concerning secureness (defined as no missed VNT in MELD score ≥10 [11]) in population screening, it was excellent with Baveno VI criteria and LIP (Figure S4A). Secureness was almost excellent with PLER, PLEASE, and Anticipate, where there were few patients with missed VNT by MELD score ≥10. The expanded Baveno VI criteria and LIP were not secured in individual screening. Indeed, the missed VNT rate was increased in MELD score ≥10, more markedly for the expanded Baveno VI criteria than for LIP (Figure S5A). Considering that a MELD score ≥10 was a limit for applying 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 S6). 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.

BLIP in NAFLD

The BLIP score [(LSM*45*30)/(PI*platelets*BMI)] was well calibrated for VNT probability (Figure S6). It was more discriminant for VNT than the LIP score, 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 S7) and individual screening (Table S8). 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 secured). 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 LIP or the NAFLD cirrhosis criteria (Figure S7). 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 S9).

Clinical application

All etiologies - First, in population screening, LIP can be applied to all CLD whatever the liver dysfunction (except in ascites due to elastographic limitation) (Figure 3A). Endoscopy can be confidently avoided in 33.4% of patients under a LIP score cut-off of around 5%, i.e. a missed VNT rate <5%. Otherwise, endoscopy is required in the remaining 66.6% of patients. In individual screening (Figure 3B), LIP must be limited to MELD scores <10, which securely spared 60.4% of endoscopies. The LIP score can more precisely quantify the VNT risk. Thus, its PPV can reach a maximum of 46.1%. However, the last two applications should be applied to populations with estimated VNT prevalences ≤of 15.2%: this preserves safety as

predictive values depend on prevalence. We emphasise that this relatively low frequency is consistent with standard clinical practice.

NAFLD – BLIP [(LSM*45*30)/(PI*platelets*BMI)] is the preferred test since it offers better safety and performance. Thus, in population screening, BLIP spared 40.8% of endoscopies (Figure 3C); and in individual screening, BLIP (must be limited to MELD scores <10) securely spared 79.7% of endoscopies (Figure 3B). The BLIP score can more precisely quantify the VNT risk. Thus, its PPV can reach a maximum of 100% (Figure 3C).

Practice - Clinical use adapted to every setting is summarised in Figure 4. A simple exportable calculator (Excel file) is available at

https://uabox.univ-angers.fr/index.php/s/wvZ84PzjM7FVwD6

DISCUSSION

Originalities **-** Our large population encompassed a substantial spectrum of characteristics for evaluating varices needing treatment (VNT) tests in chronic liver disease, including a wide range of liver function and the three main liver etiologies. Including patients irrespective of liver severity (except for ascites) had two advantages. First, a few patients had VNT below the cut-off of 10 kPa used in the original definition of compensated advanced chronic liver disease (cACLD) [1]. Excluding these patients would have missed more than 5% of VNT, the limit conceded in the Baveno VI statement. Our inclusion criteria fit better with the recent cACLD criteria [22]. 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 liver disease [11]. We report two screening strategies, one developed for individual patient screening and another for population screening. Until now, safety has been based on a quantitative definition (missed VNT <5%). Here, we would refine the safety definitions. Two *quantitative* safety definitions were used (sensitivity and NPV), and two *qualitative* criteria were described (no missed VNT in severe liver dysfunction [11], and a requirement for a low level of liver dysfunction in missed VNT. With regard to *functional safety*, we focus in detail on the Supplementary Material. Finally, with this large population, we could further examine interactions and, consequently, give a better ratio of VNT markers. LR- is particularly significant in the current context since it represents both performance and safety in a single descriptor, but its utility is limited to population screening because LR is dependent on sensitivity and specificity. Finally, we extended the principle of sparing endoscopies by ruling in VNT thanks to the 100% specificity cut-off.

Main results **-** The LIP test [(LSM*45)/(PI*platelets)] performed better than the Baveno VI criteria, Anticipate and PLER (platelets/liver elastometry ratio) in population screening. The exception was PLEASE (PLER adjusted to etiology, sex and INR), 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 alcoholic liver disease, and BLIP [(LSM*45*30)/(PI*platelets*BMI)] was superior to PLEASE in NAFLD. Furthermore, it should be mentioned that the PLEASE test was not designed for individual screening.

Combining 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 current comprehensive population in terms of size, etiologies, and liver severity favours the exportability of a test cut-off from an epidemiological point of view. In comparison, there is an infinite number of cut-off choices when two or more VNT predictors are applied independently. Third, the calculation offers greater simplicity, robustness and performance than a logistic score including the same variables. That increased simplicity can be observed in the straightforward arithmetic calculation, which does not require a particular web calculator like the PLEASE score or a nomogram like the Anticipate score. The unique formula conferred greater robustness without marker coefficients depending on the population. Furthermore, the AUROC demonstrates superior performance, similar to that of a logistic score utilising three factors, but with a substantially higher spared endoscopy rate than the latter (details in Complementary Material). 100% specificity was obtained optionally with bilirubin to rule in VNT in a few patients of the whole population (details in Complementary Material) and even with BMI, included in the BLIP test, in more patients with NAFLD. This specificity level has not been reported previously. 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 [11].

Which test to use? Due to many constraints, the Baveno VI criteria, the expanded Baveno VI criteria, and Anticipate must be rejected. This leaves two real competitors: LIP [(LSM*45)/(PI*platelets)] and PLEASE (platelets/liver elastometry ratio adjusted on etiology, sex and INR), with the latter outperforming LIP exclusively in viral chronic liver disease. The higher PLEASE test performance was attributed to its algorithm stratified on sex and etiology; the PLEASE score discriminated VNT less well than the LIP score (Figure 2B). LIP, on the other hand, had four benefits. The first was its simplicity and hence bedside usage. The second benefit of its improved calibration for directly expressing VNT probability was its exact estimate of VNT risk. Third, LIP gave complete security (in population screening). Fourth, because it is a score, it is considerably better suited to a personalised strategy than the PLEASE test, which offers individual and population tests.

Finally, we may discuss two types of tests. On the one hand, there are simple tests like the Baveno VI, PLER, and LIP criteria, which outperform the others. On the other hand, the two tests necessitate the use of a specialised calculator: Anticipate (requires at least one nomogram) and PLEASE/VariScreen. As a result, the decision between the two is determined by the clinical situation. However, in NAFLD, BLIP had two benefits in population screening: first, its performance (41%) was superior to PLEASE (37%), and second, it allowed the diagnosis of VNT.

Which strategy? - 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 number of patients with missed varices needing treatment was far higher (LIP: 18.4% vs 4.2% of VNT). Second, the patients with missed VNT had greater liver dysfunction (details in Supplemental Material). Third, the performance of the individual

strategy depends on population prevalence and thus is limited to a VNT prevalence ≤15%, although that rate should often be the case considering a global strategy. Therefore, we privilege population screening [8] as a public health strategy to prevent more deaths from variceal bleeding. However, individual screening is a personalised option adapted to certain patients, e.g. particularly reluctant to screening endoscopy (Figure 4). Consequently, its use, often privileged by clinicians, should be restricted. Finally, this choice is an individual clinician decision.

Limits **-** The first category of limits, which has been discussed in depth elsewhere [11], 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; LSMs obtained with the XL probe must not be used; and the non-evaluation of certain 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 despite its more complex formula [23]. We underline however that robust validation requires large populations (≥400) as previously discussed [8].

Three putative LIP [(LSM*45)/(PI*platelets)] limits merit discussion. First, LIP was adjusted to etiology. This induced one inconvenience. Indeed, three cut-offs (value around 5%) were necessary. However, this aspect also allowed us to make the test safe for NAFLD and increase performance in alcohol-related liver disease compared to a test with a unique cut-off (data not shown). Alcohol-related liver disease nonetheless remained a challenging etiology for the

VNT tests. The lower performance of LIP in the alcohol context was due to poor synergy between LIP variables (as clearly shown in Figure S8A).

Second, in individual screening, LIP safety was 5.2%, albeit this potential restriction was avoided by limiting LIP usage to MELD scores <10, thereby fulfilling secureness standards. This restriction resulted in a paradox: performance improved (+6.3%) in the remaining subgroup but dropped (-3.1%) when the intention to diagnose concept was applied to the entire population. Finally, the PPV range of the LIP score was 0 to 46.1%, which is rather narrow. On the other hand, the awareness of a VNT risk of roughly 50% will push patients and physicians to undergo endoscopy more than the ambiguous risk (5%) offered by standard binary testing. Notably, in NAFLD, BLIP PPV reached 100% (Figure 3C).

Clinical application **-** LIP can be used in any patient with stable chronic liver disease, whatever the liver severity (in population screening) and in the main three etiologies, 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 (details in Supplemental material). 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 3). The clinical limit of LIP is the requirement for FibroScan apparatus, thus excluding ascites. Otherwise, there is no additional cost associated with LIP in centres where FibroScan 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 [24, 25]. This extension was endorsed in Baveno VII statements [26] which need to be refined [27]. These reasons include adapting motivation for drug compliance, which is an important clinical challenge, and managing the contraindications and side effects of those NSBBs [25]. Moreover, the non-invasive criteria of clinically significant portal hypertension still need to be validated [22, 28]. Furthermore, adherence to screening by patients and primary care

providers is improved by precise information [29, 30] and a negative perception of the disease [31]. With that respect, knowledge of the precise VNT risk would be more convincing than that of conceptual clinically significant portal hypertension.

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 [(LSM*45)/(PI*platelets)] is a test combining simplicity, performance, safety and deployability in each major liver disease etiologies. LIP can be used in population screening regardless of liver severity (except ascites). The combination of LIP and BMI, BLIP [(LSM*45*30)/(PI*platelets*BMI)] improves efficacy in NAFLD. LIP performance can be very high for individual patient screening, even in NAFLD with its BLIP version (80%).

Conflict of interest statement unrelated to the topic:

Federico Ravaioli, Oana Farcau, Horia Stefanescu, Camille Candillier, Annalisa Berzigotti, Antonio Colecchia, Paul Calès: no disclosures

Arthur Berger: consultant for Lilly until December 2018

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

Christophe Bureau: honoraria from Gore, Gilead, Abbvie

Nathalie Ganne-Carrié: grant from Echosens for a monocenter clinical study (M123); honoraria from Abbvie, Bayer, BMS and Gilead

Victor de Ledinghen: lectures for AbbVie, Gilead, Intercept Pharma, Echosens, Supersonic Imagine; consultant for Abbvie, Gilead, MSD, BMS, Siemens

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

Financial support statement:

Institutional support from participating centres for the present study, 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.

Supplemental Material

The Supplemental Material includes 16 tables and 11 figures.

REFERENCES

[1] De Franchis R, Faculty B V. Expanding consensus in portal hypertension: Report of the Baveno VI Consensus Workshop: Stratifying risk and individualizing care for portal hypertension. J Hepatol 2015; 63: 743-752.

[2] Marot A, Trepo E, Doerig C, Schoepfer A, Moreno C, Deltenre P. Liver stiffness and platelet count for identifying patients with compensated liver disease at low risk of variceal bleeding. Liver Int 2017; 37: 707-716.

[3] Roccarina D, Rosselli M, Genesca J, Tsochatzis E A. Elastography methods for the non-invasive assessment of portal hypertension. Expert Rev Gastroenterol Hepatol 2018; 12: 155-164.

[4] Thabut D, Bureau C, Layese R, Bourcier V, Hammouche M, Cagnot C, et al. Validation of Baveno VI Criteria for Screening and Surveillance of Esophageal Varices in Patients With Compensated Cirrhosis and a Sustained Response to Antiviral Therapy. Gastroenterology 2019; 156: 997-1009.e1005.

[5] Moctezuma-Velazquez C, Saffioti F, Tasayco-Huaman S, Casu S, Mason A, Roccarina D, et al. Non-Invasive Prediction of High-Risk Varices in Patients with Primary Biliary Cholangitis and Primary Sclerosing Cholangitis. Am J Gastroenterol 2019; 114: 446-452.

[6] Stafylidou M, Paschos P, Katsoula A, Malandris K, Ioakim K, Bekiari E, 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: 1744-1755 e1711.

[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] Cales P, Buisson F, Ravaioli F, Berger A, Carboni C, Marasco G, et al. How to clarify the Baveno VI criteria for ruling out varices needing treatment by non-invasive tests. Liver Int 2019; 39: 49-53.

[9] Augustin S, Pons M, Maurice J B, Bureau C, Stefanescu H, Ney M, et al. Expanding the Baveno VI criteria for the screening of varices in patients with compensated advanced chronic liver disease. Hepatology 2017; 66: 1980-1988.

[10] Petta S, Sebastiani G, Bugianesi E, Vigano M, Wong V W, Berzigotti A, et al. Noninvasive Prediction of Esophageal Varices by Stiffness and Platelet in Nonalcoholic Fatty Liver Disease Cirrhosis. J Hepatol 2018; 69: 878-885.

[11] Berger A, Ravaioli F, Farcau O, Festi D, Stefanescu H, Buisson F, 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.

[12] Berger A, Ravaioli F, Farcau O, Festi D, Stefanescu H, Buisson F, et al. The prevalence of esophageal varices needing treatment depends on gender, etiology and BMI. J Hepatol 2020; 73: S751-752.

[13] Kamath P S, Kim W R, Advanced Liver Disease Study G. The model for end-stage liver disease (MELD). Hepatology 2007; 45: 797-805.

[14] Boursier J, Konate A, Gorea G, Reaud S, Quemener E, Oberti F, et al. Reproducibility of liver stiffness measurement by ultrasonographic elastometry. Clin Gastroenterol Hepatol 2008; 6: 1263-1269.

[15] Cales P, Oberti F, Bernard-Chabert B, Payen J L. Evaluation of Baveno recommendations for grading esophageal varices. J Hepatol 2003; 39: 657-659.

[16] Singal A K, Kamath P S. Model for End-stage Liver Disease. J Clin Exp Hepatol 2013; 3: 50-60.

[17] Abraldes J G, Bureau C, Stefanescu H, Augustin S, Ney M, Blasco H, et al. Noninvasive tools and risk of clinically significant portal hypertension and varices in compensated cirrhosis: The "Anticipate" study. Hepatology 2016; 64: 2173-2184.

[18] Bossuyt P M, Reitsma J B, Bruns D E, Gatsonis C A, Glasziou P P, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015; 351: h5527.

[19] Boursier J, De Ledinghen V, Poynard T, Guechot J, Carrat F, Leroy V, et al. An extension of STARD statements for reporting diagnostic accuracy studies on liver fibrosis tests: the Liver-FibroSTARD standards. J Hepatol 2015; 62: 807-815.

[20] Boursier J, Zarski J P, De Ledinghen V, Rousselet M C, Sturm N, Lebail B, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. Hepatology 2013; 57: 1182-1191.

[21] Collins G S, Reitsma J B, Altman D G, Moons K G. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 2015; 350: g7594.

[22] Papatheodoridi M, Hiriart J B, Lupsor-Platon M, Bronte F, Boursier J, Elshaarawy, et al. Refining the Baveno VI elastography criteria for the definition of compensated advanced chronic liver disease. J Hepatol 2021; 74: 1109-1116.

[23] 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: 150.

[24] Villanueva C, Albillos A, Genesca J, Garcia-Pagan J C, Calleja J L, Aracil C, 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.

[25] Garcia-Tsao G, Abraldes J G. Nonselective Beta-Blockers in Compensated Cirrhosis: Preventing Variceal Hemorrhage or Preventing Decompensation? Gastroenterology 2021; 161: 770-773.

[26] De Franchis R, Bosch J, Garcia-Tsao G, Reiberger T, Ripoll C, Baveno V I I F. Baveno VII - Renewing consensus in portal hypertension. J Hepatol 2022; 76: 959-974.

[27] Dajti E, Ravaioli F, Marasco G, Alemanni L V, Colecchia L, Ferrarese A, et al. A Combined Baveno VII and Spleen Stiffness Algorithm to Improve the Noninvasive Diagnosis of Clinically Significant Portal Hypertension in Patients With Compensated Advanced Chronic Liver Disease. Am J Gastroenterol 2022; 117: 1825-1833.

[28] Pons M, Augustin S, Scheiner B, Guillaume M, Rosselli M, Rodrigues S G, et al. Noninvasive Diagnosis of Portal Hypertension in Patients With Compensated Advanced Chronic Liver Disease. Am J Gastroenterol 2021; 116: 723-732.

[29] Yakovchenko V, Bolton R E, Drainoni M L, Gifford A L. Primary care provider perceptions and experiences of implementing hepatitis C virus birth cohort testing: a qualitative formative evaluation. BMC Health Serv Res 2019; 19: 236.

[30] 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: 610-625.

[31] 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: 355-358.

Table 1. 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 [b].
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-: negative likelihood ratio
New strategy characteristic developed in the present study is in bold.
[a] Since using an unpaired statistical test.
[b] Since LR are based on sensitivity and specificity.

Table 2. VNT discrimination by scores as a function of population sets.

	Anticipate	**PLER**	**PLEASE**	**LIP**
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 etiology, sex, INR

[a] Paired Delong test

Table 3. Missed VNT and spared endoscopy rates (%) of tests according to population screening (sensitivity ≥95%) in the whole population and as a function of etiology.

	B6C	Anticipate	PLER	PLEASE	LIP	*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 etiologies (*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 etiology 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: p=0.006 by McNemar test. Other comparisons per etiology in Table S5.

[d] Unpaired Chi² test for spared endoscopy and likelihood ratio test for missed VNT

Table 4. Missed VNT and spared endoscopy rates (%) of tests according to individual screening (NPV ≥95%) in the whole population and as a function of etiology.

	EB6C	**LIP**	***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 etiologies (*p* [a]):			
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 etiology and sex.

[a] Unpaired Chi² 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.

FIGURE LEGENDS

Fig. 1. 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. LIP: (LSM*45)/(PI*platelets); VNT: varices needing treatment.

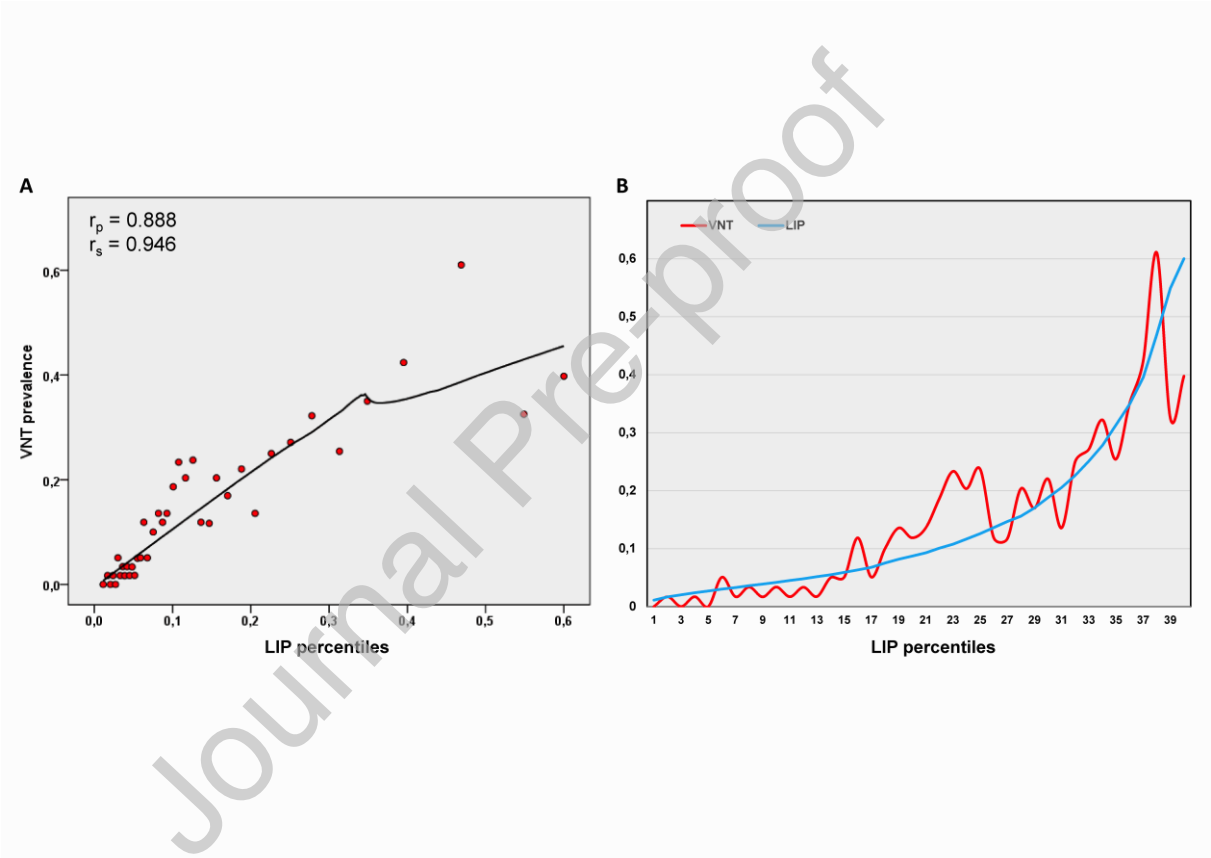


Fig. 2. Discrimination for VNT diagnosis (ROC curves). Panel A: LIP score and its composite markers with magnification showing the subset sizes with ≥95% sensitivity. Panel B: scores evaluated. In both of the figures on the right, the horizontal dashed lines show the superiority of LIP in sparing endoscopy with a missed VNT rate <5% (population screening). LIP: (LSM*45)/(PI*platelets); LSM: liver stiffness measurement; VNT: varices needing treatment; PLER: platelets/liver elastometry ratio; PLEASE: platelets/liver elastometry ratio adjusted on etiology, sex and INR.

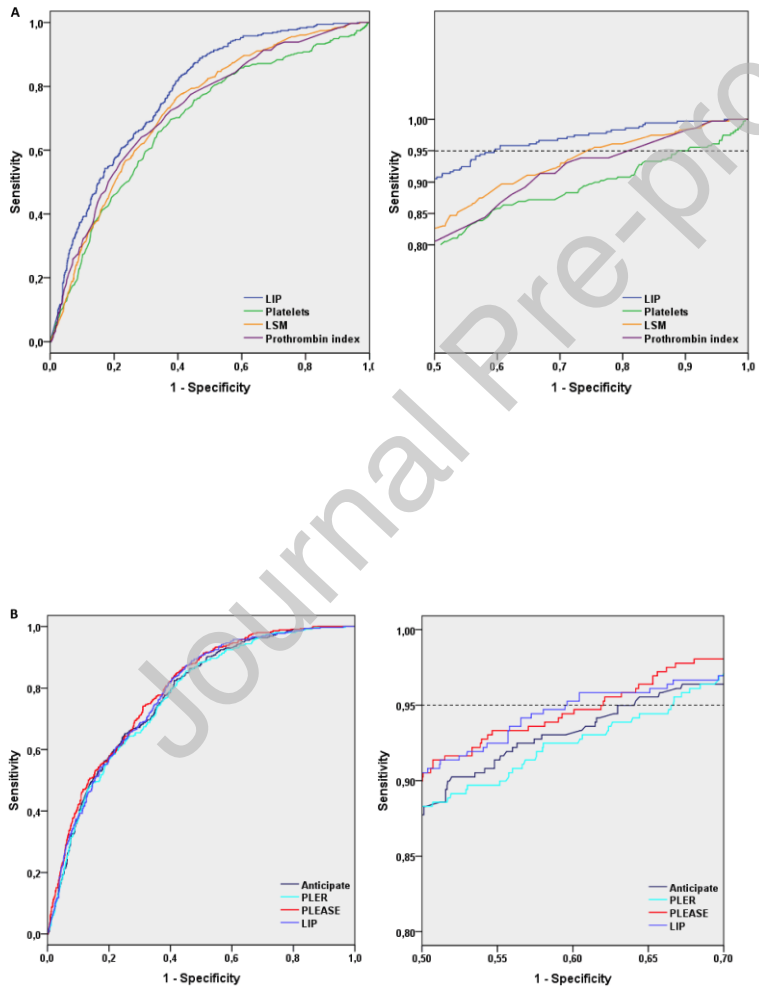
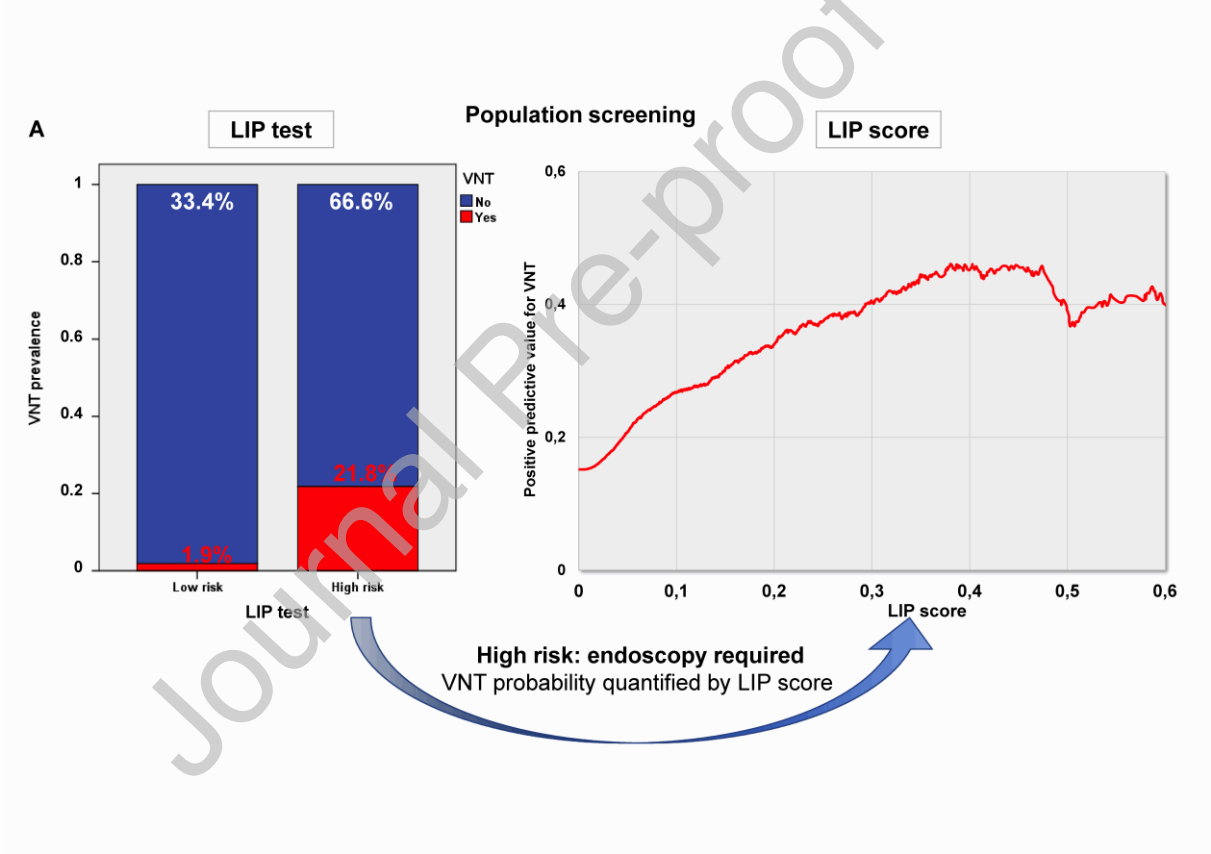
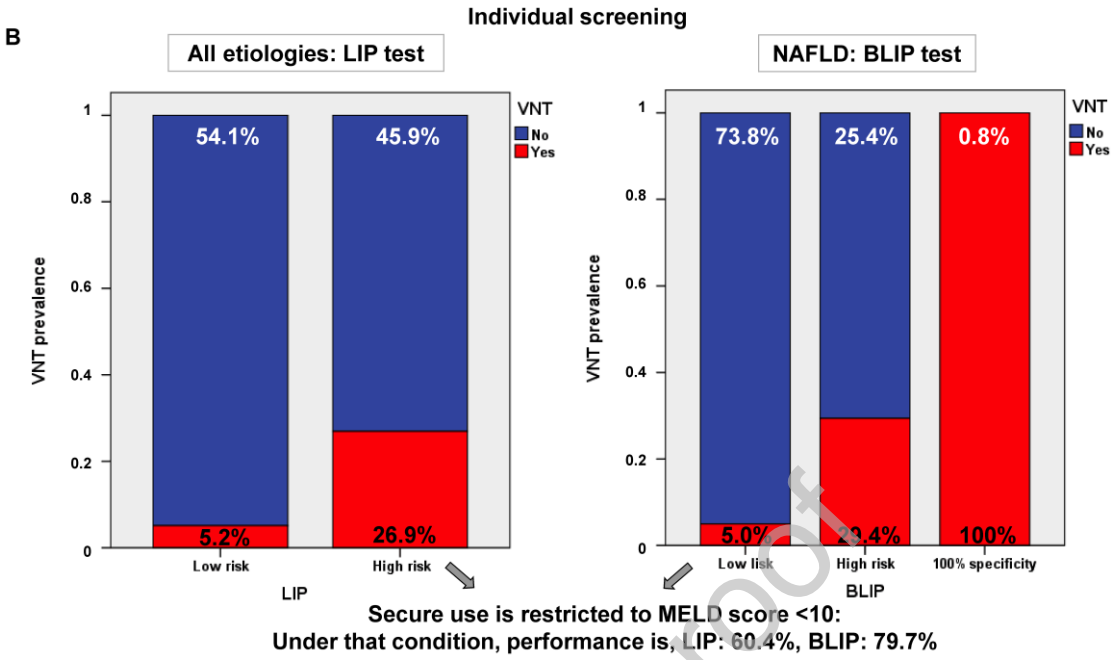


Fig. 3. LIP and BLIP: performance and safety of tests, and PPV of scores as a function of screenings. Panel A: LIP for population screening in all etiologies. Panel B: LIP and BLIP for individual screening. Panel C: BLIP for population screening in NAFLD. Figures within bars (from the whole population) indicate category prevalence (in white at top) and VNT prevalence (in black at bottom). BLIP: (LSM*45*30)/(PI*platelets*BMI); LIP: (LSM*45)/(PI*platelets); MELD: model of end-stage liver disease; NAFLD: non-alcoholic fatty liver disease; VNT: varices needing treatment; PPV: positive predictive value.





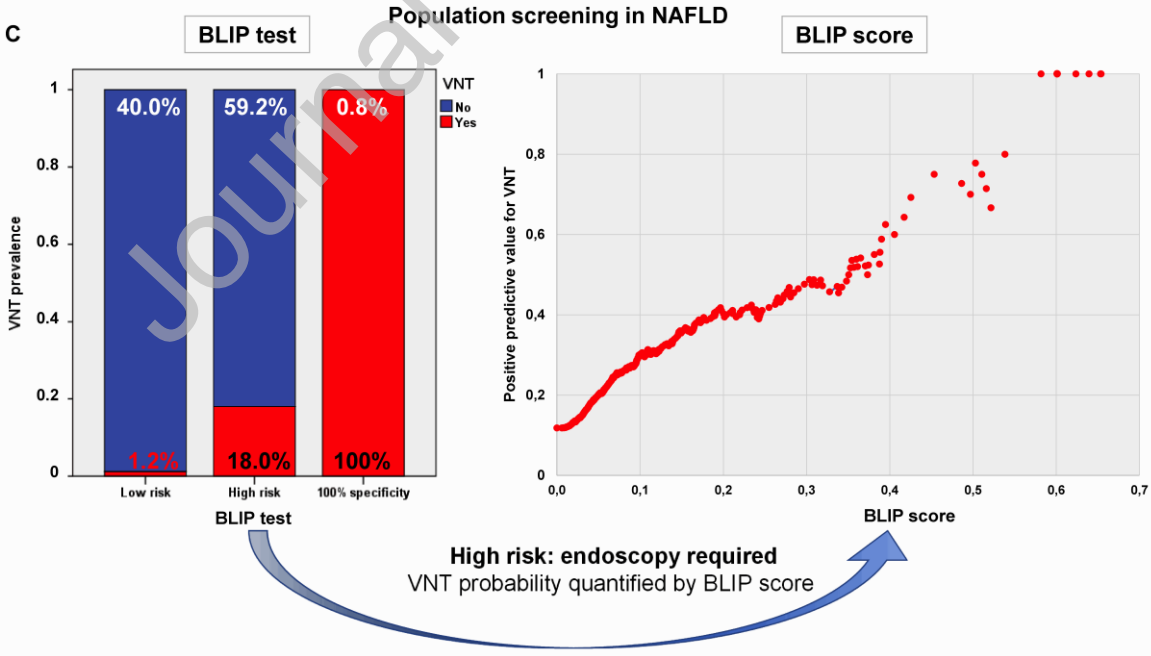


Fig. 4. Clinical application of new tests. Population screening is the first line option; individual screening is a second line option for certain patients (e.g. reluctant to endoscopy). BLIP is the preferred option in NAFLD.

BLIP: (LSM*45*30)/(PI*platelets*BMI); LIP: (LSM*45)/(PI*platelets); MELD: model of end-stage liver disease; NAFLD: non-alcoholic fatty liver disease; VNT: varices needing treatment; NPV: negative predictive value.