TITLE: Noninvasive Prediction of Esophageal Varices by Stiffness and Platelet in Nonalcoholic Fatty Liver Disease Cirrhosis

AUTHORS: Salvatore Petta[1], Giada Sebastiani[2], Elisabetta Bugianesi[3], Mauro Viganò[4], Vincent Wai-Sun Wong[5], Annalisa Berzigotti[6], Anna Ludovica Fracanzani[7], Quentin M. Anstee[8], Fabio Marra[9], Marco Barbara[1], Vincenza Calvaruso[1], Calogero Cammà[1], Vito Di Marco[1], Antonio Craxì[1], Victor de Ledinghen[10].

INSTITUTIONS:

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

[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 & Liver Unit, Newcastle Upon Tyne Hospitals NHS Foundation Trust, Freeman Hospital, Newcastle upon Tyne, United Kingdom.

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

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

CORRESPONDING AUTHOR: Dr. Salvatore Petta, Sezione di Gastroenterologia e Epatologia, Di.Bi.M.I.S., Università di Palermo, Piazza delle Cliniche, 2, 90127 Palermo, Italy.Phone: +39 091 655 2170. Fax +39 091 655 2156. E-mail: salvatore.petta@unipa.it

Conflict of Interest: None to declare

Contributors

S. Petta designed the study, contributed to data acquisition, was responsible for writing the manuscript, and participated in statistical analysis. G. Sebastiani, E. Bugianesi, M. Viganò, V.W-S. Wong, A. Berzigotti, A.L. Fracanzani, Q.M. Anstee, F. Marra, M. Barbara, V. Calvaruso, C. Cammà, V. Di Marco, A. Craxì, V. de Ledinghen were responsible for the project and writing of the manuscript. All authors have seen and approved the final version of the manuscript.

Acknowledgements

QMA, EB & FM are members of the *EPoS (Elucidating Pathways of Steatohepatitis)* consortium funded by the Horizon 2020 Framework Program of the European Union under Grant Agreement 634413. Aspects of this work were supported by the Newcastle NIHR Biomedical Research Center.

ABSTRACT

Background/Aim: Baveno VI and expanded Baveno VI non-invasive criteria can avoid the need for esophagogastroduodenoscopy(EGD) to screen for varices needing treatment(VNT) in a substantial proportion of compensated patients with viral and/or alcoholic cirrhosis. This multicenter, cross-sectional study aims to validate these criteria in patients with compensated cirrhosis due to nonalcoholic fatty liver disease(NAFLD), accounting for possible differences in liver stiffness measurement(LSM) values between M and XL probes.

Materials/Methods: We assessed 790 patients with NAFLD-related compensated cirrhosis who had EGD within 6 months of a reliable LSM measured by FibroScan® using M and/or XL probe. Baveno VI and expanded Baveno VI criteria were tested. The main variable used to optimize criteria was the percentage of endoscopies spared, keeping the risk of missing large VNT below a <5% threshold.

Results: 314 patients had LSM by both M and XL probes(training set), while 338 and 138 by only M or XL probe, respectively(validation sets). In the training set use of Baveno VI and expanded Baveno VI criteria reduced by 33.3% and by 58% the number of EGD, missing 0.9% and 3.8% of large OV, respectively. The best thresholds to rule-out VNT were identified at PLT>110,000 and LSM<30 KPa for M probe, and PLT>110,000 and LSM<25 KPa for XL probe(NAFLD cirrhosis criteria). These criteria spared 68.5% and 65% of EGD, by missing 4.2% and 4.9% of VNT, respectively. Usage of NAFLD cirrhosis criteria would have thus led to an absolute reduction in the number of EGD screened patients of 34.7% and 10.5% with respect to BAVENO VI and expanded BAVENO VI criteria respectively

Conclusion: The new NAFLD cirrhosis criteria, established for the FibroScan probe, can reduce by more than half the use of EGD to screen for VNT in NAFLD cirrhosis, with a chance of missing VNT below 5%.

Keywords: Varices, NAFLD, Cirrhosis, Baveno, stiffness

Introduction

The pandemic spreading of obesity and diabetes makes nonalcoholic fatty liver disease (NAFLD) the growing most common cause of chronic liver disease and cirrhosis [1], an increasing risk factor for hepatocellular carcinoma [2], and the emerging indication for liver transplantation [3]. Consistent with these data the management of patients with NAFLD-related cirrhosis represents a challenge in terms of epidemiological, clinical and economical burden. In this complex picture, the diagnosis of esophageal varices (EV) and especially large (grade 2/3) EV requiring primary prophylaxis (varices needing treatment, VNT), is of paramount prognostic importance in all patients with cirrhosis including those with NAFLD [4,5]. However, VNT are not frequent in patients with compensated cirrhosis, and strategies to reduce the number of unnecessary esophagogastroduodenoscopy (EGD) screening have been proposed. Recently, the Baveno VI guidelines proposed that compensated cirrhotic patients with a liver stiffness measurement (LSM) <20kPa and a platelet count >150000/μL can avoid screening endoscopy [6], the specificity of this strategy for excluding VNT being validated in different studies [7]. Furthermore, expanded Baveno VI criteria, obtained by optimizing LSM and platelets (PLT) values (<25kPa and >110000/ μL, respectively), have also been proposed and demonstrated to spare a higher proportion of unnecessary EGD when compared to Baveno VI criteria [8].

Both Baveno VI and expanded Baveno VI criteria were proved reliable in large cohorts of patients with compensated cirrhosis mostly due to viral and/or alcoholic etiologies, while patients with NAFLD-related cirrhosis are absent or under represented. This is a significant limitation considering that in patients with NAFLD-related cirrhosis LSM could also account for severity of steatosis and obesity [9,10]. Furthermore, studies assessing the diagnostic accuracy of Baveno VI and expanded Baveno VI criteria considered LSM values obtained only by M probe. However, NAFLD patients are at high prevalence of obesity, this issue resulting in frequent need to measure LSM by XL probe in case of failure of M probe.

We aimed to validate the Baveno VI and expanded Baveno VI criteria in patients with compensated cirrhosis due to NAFLD, also taking into account potential differences in LSM values from M and XL probes.

Methods:

Patients selection:

Data from 790 patients full filling the above reported inclusion criteria and prospectively recruited at the first diagnosis of NAFLD-related compensated cirrhosis in 10 different centers were retrospectively reviewed and analyzed.

Inclusion criteria were presence of a reliable LSM and of an EGD within 6 month of LSM. LSM was obtained by FibroScan machine by using M and/or XL probe. In some centers M probe was the only available, in some others both M and XL probes were available. When both probes were available some centers used the XL probe only in case of failure of the M probe, some others measured LSM

by both probes. Cirrhosis was diagnosed by histology [11] and/or by LSM >11.5 KPa for M probe [12] and >11 KPa for XL probe [13]. In patients without histology, diagnosis of NAFLD required detection of ultrasonographic steatosis plus at least one criterion of the metabolic syndrome (obesity, diabetes, arterial hypertension, dydlipidemia). Other causes of liver disease were ruled out, including alcohol intake >20 g/day during the previous year (evaluated by interview of patients on amount, frequency and type, and confirmed by at least one family member), viral (HBsAg, anti-HCV and anti-HIV negativity) and autoimmune hepatitis, hereditary hemochromatosis, and alpha1-antitrypsin deficiency. Patients with advanced (Child Pugh B or C) cirrhosis, hepatocellular carcinoma, liver transplantation, EV banding, portal or splenic vein thrombosis and splenectomy were excluded. The study cohort finally included 372 patients from the Centre d'Investigation de la Fibrose Hépatique, Bordeaux University Hospital, 110 patients from the Division of Gastroenterology and Hepatology, McGill University Health Centre of Montreal QC, 104 patients for the Section of Gastroenterology and Hepatology, University of Palermo, 63 patients from the Hepatology Unit, Ospedale San Giuseppe University of Milan, 36 patients from the Division of Gastroenterology, Department of Medical Sciences, University of Torino, 35 patients from the Department of Medicine and Therapeutics, The Chinese University of Hong Kong, 25 patients from the Swiss Liver Center, 21 patients from the Department of Pathophysiology and Transplantation, Ca' Granda IRCCS Foundation, Policlinico Hospital, University of Milan, 12 patients from the Newcastle Upon Tyne Hospitals NHS Trust, Freeman Hospital, and 12 patients from the Dipartimento di Medicina Sperimentale e Clinica, University of Florence.

The study was carried out in accordance with the principles of the Helsinki Declaration, and with local and national laws. Approval was obtained from the hospital Internal Review Boards and their Ethics Committees, and written informed consent for the study was obtained from all patients.

Patients evaluation:

Clinical and anthropometric data, including BMI, the presence of arterial hypertension and type 2 diabetes, were collected at the time of enrollment. The same day a 12-hour overnight fasting blood sample was drawn to determine serum levels of AST, ALT, PLT, albumin, total bilirubin, INR, total and HDL cholesterol, triglycerides, and plasma glucose concentration.

Transient elastography was performed with the FibroScan (Echosens, Paris, France) medical device, using the M and/or XL probes. In each center, LSM was assessed after at least 4 hour fasting, by a trained operator who had previously performed at least 300 determinations in patients with chronic liver disease. Only patients with 10 valid measurements and with reliable results according to published criteria were enrolled [14].

EGD was performed by a small number of experienced operators at each Hospital. Patients were excluded from this study, if there was more than six month time interval between TE and EGD examinations. At endoscopy, high-risk EV which warrant primary

prophylaxis against EV bleeding were defined by medium or large size or the presence of high-risk stigmata findings (red wale marks, cherry red spots) [6].

Statistical analysis

Continuous variables were summarized as mean ± SD, and categorical variables as frequency and percentage.

The training set was defined as the subcohort of patients for which LSM were available by both M and XL probes, while the remaining patients with only one of M or XL probes acted as validation sets. Patients in the training set were used to find two different combined thresholds of LSM and Platelets count, one for the M and the other for the XL probe, which maximized the absolute number of spared endoscopies while keeping the risk of missed VNTs below 5%. (while constraining the negative predictive value to be at least 0.95). The 5% false negative rate of undetected VNT was agreed as a reasonable criterion by experts in the Baveno VI consensus conference and later adopted by the American Gastroenterological Association and several other authors [6]. Found thresholds' performances were subsequently evaluated in their respective validation sets.

Baveno VI (LSM <20 KPa and PLT > 150,000) and expanded Baveno VI (LSM <25 KPa and PLT > 110,000) criteria were also evaluated in both training and validation sets.

Performance evaluations were made in terms of number and percentage of spared endoscopies, number and percentage of undetected EV, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive and negative likelihood ratios.

All the analyses were performed using the R statistical computing environment.

RESULTS

Patient characteristics

The baseline characteristics of the 790 patients with NAFLD-related compensated cirrhosis (682 Child Pugh A5, and 108 Child Pugh A6) included in the study are shown in Table 1. Mean age was 62.3 years, with a 55% prevalence of males. Sixty-two percent of patients were obese, and diabetes and arterial hypertension were present in 62.8% and 61.1% of cases respectively. Supplemental Table 1 shows the characteristics of the population split for each centre. Other 56 patients with NAFLD-related compensated cirrhosis were not included in the study because LSM was not reliable [14]. This cohort, compared to the enrolled 790 patients, had similar age (61±12 years) and similar prevalence of male gender (60.7%), and of VNT (10.7%), while higher BMI (35.1±8.0 Kg/m^2), this last being the main cause of LSM unreliability.

Liver cirrhosis was diagnosed by histology in 34.4% of patients. EV were present in 31.5% of cases, and VNT in 11.5%. The prevalence of EV (55.5% vs 27.4%, p<0.001) and of VNT (29.6% vs 8.6%, p<0.001) was higher in Child Pugh A6 compared to A5 patients. LSM was similar in patients with histological diagnosis of cirrhosis compared to those without, for both M (24.3±16.5 vs 25.3±15.8, p=0.44) and XL (22.9±14.0 vs 22.2±15.9, p=0.68) probes. Among the 139 patients with LSM <13 KPa –at lower risk of liver damage- the diagnosis of cirrhosis was made by histology in 53.2%, and based on LSM values [12,13] in the remaining 65 patients.

Three-hundred fourteen patients had LSM by both M and XL probes, and this cohort was considered as the training set. In this cohort, LSM by M probe was significantly higher than that by XL probe (24.2±14.7 vs 19.8±14.2; p<0.001). Three-hundred thirty-eight and 138 patients had LSM by only M and XL probe, respectively, and these cohorts were considered as validation sets. When considering patients with LSM by M probe, those in the training cohort had similar LSM (24.2±14.7 KPa vs 25.6±17.2, p=0.26) and PLT (176.5±80.0 10^3/mmc vs 178.4±98.1, p=0.78) values and similar prevalence of VNT (10.2% vs 13.3%, p=0.21) compared to those in the validation cohort, while having a significantly higher prevalence of obesity (64.7% vs 49.8%, p<0.001) and a significantly lower prevalence of any grade EV (26.4% vs 36.3%, p=0.006) (Table 1). When looking at patients with LSM by XL probe, those in the training cohort had similar PLT values (176.5±80.0 10^3/mmc vs 176.3±81.6, p=0.98) and similar prevalence of both any grade EV (26.4% vs 29.7%, p=0.47) and VNT (10.2% vs 10.1%, p=0.98) compared to those in the validation cohort, while having a significantly lower prevalence of obesity (64.7% vs 88.5%, p<0.001) and significantly lower LSM values (19.8±14.2 KPa vs 27.7±16.9, p<0.001) (Table 1).

Diagnostic accuracy of LSM and PLT for VNT

Training Set

One hundred six out of 314 patients (33.8%) met Baveno VI criteria as a rule-out for VNT of whom 11 (10.4%) had any grade EV and 1 patient (0.9%) had VNT (Table 2A). When looking at expanded Baveno VI criteria 182 out of 314 patients (58%) met criteria of whom 30 (16.5%) had any grade EV and 7 patient (3.8%) had VNT (Table 2A). Consistently, NPV for VNT was 99.1% and 96.2% for Baveno VI and expanded Baveno VI criteria, respectively (Table 2A).

Furthermore, we tested for new criteria specific for NAFLD-related cirrhosis –hereinafter called NAFLD cirrhosis criteria- and discriminated for M or XL probe. We identified as the best thresholds for rule-out VNT, PLT>110,000 and LSM<30 KPa for M probe, and PLT>110,000 and LSM<25 KPa for XL probe. The use of these thresholds allowed sparing 68.5% and 65% of EGD, by missing 4.2% and 4.9% of VNT, and 17.7% and 18.1% of any grade EV, respectively (Table 2A and B). Negative predictive value for VNT was 95.8% and 95.1% for patients with LSM by M or XL probe, respectively (Table 2A and B). Consistent with these data, usage of NAFLD cirrhosis criteria in this cohort would have led 31.5% of patients to EGD screening (99 out of 314), which respectively represents 34.7% less EGDs than those required using Baveno VI criteria (208 out of 314), and 10.5% less EGDs required using expanded Baveno VI criteria (132 out of 314).

Validation Set

The diagnostic accuracy of Baveno VI, expanded Baveno VI and new NAFLD cirrhosis criteria was tested in the validation cohorts.

In the validation cohort of 338 patients with LSM by only M probe, PLT>110,000 and LSM<30 KPa were confirmed better than Baveno VI and expanded Baveno VI criteria in terms of spared endoscopy (61.8% VS. 33.4% vs 54.1%) while keeping a similar rate of missed VNT (4.8% vs. 4.4% vs 4.4%) and of missed any grade EV (22.5% vs. 18.6% vs 21.3%), and a similar NPV (95.2%, 95.6% and 95.6%) (Table 2A). Consistently, the number of needed EGDs with NAFLD cirrhosis criteria is respectively 28.4% and 7.7 % less than those needed using BAVENO VI and expanded BAVENO VI criteria.

Finally, in the validation cohort of 138 patients with LSM by only XL probe – those where in clinical practice XL probe would be used because of failure of M probe- we confirmed the good diagnostic accuracy of PLT>110,000 and LSM<25 KPa in terms of spared endoscopy (46.4%), missed VNT (1.6%), missed any grade EV (17.2%), and NPV (98.4%) (Table 2B).

Sensitivity analyses

Sensitivity analyses performed in each center considered separately, confirmed that when considering LSM by M probe the new elaborated NAFLD-cirrhosis criteria spared more EGD than Baveno VI and extendend Baveno VI criteria (range 50%-80.6% vs 25%-52.8% vs 41.7%-60%), while keeping a similar rate of missed VNT (range 0%-8.7% vs 0%-12.5% vs 0%-10%) and any grade EV (0%-38.6% vs 0%-32.4% vs 0%-41.3%) (Supplemental Table 2). Similar results were observed when considering the new elaborated NAFLD-cirrhosis criteria by using XL probe (range 33.3%-70.3% for spared EGD; 0%-22.2% for missed VNT; 0%-66.7% for missed any grade EV) (Supplemental Table 2).

When considering the entire cohort of 652 patients with LSM by M probe – those where in clinical practice M probe would be used - Baveno VI, expanded Baveno VI and NAFLD cirrhosis criteria spared 33.6%, 56% and 65% of unnecessary EGD, respectively, by missing 2.7%, 4.1% and 4.5% of VNT, respectively.

When excluding the 139 patients with LSM<13 kPa –at lower risk of liver disease severity-, the prevalence of any grade EV and of VNT was 35.6% and 14.1%, respectively; 412 patients had LSM by M and 344 by XL probe. In these patients NAFLD cirrhosis criteria spared 52.2% of unnecessary EGD by using M probe and 50.6% by using XL probe, also showing a good similar rate of missed VNT (5.6% and 4.6%, respectively) and missed any grade EV (23.3% and 20.1%, respectively). Notably, in the subgroup of patients with LSM by M probe, NAFLD cirrhosis criteria worked better than Baveno VI (18.7% spared EGD, 2.6% missed VNT and 14.3% missed any grade EV) and expanded Baveno VI (39.6% spared EGD, 4.9% missed VNT and 21.5% missed any grade EV) criteria.

Finally, when splitting the entire population according to obesity, 110 nonobese and 202 obese patients had LSM by both M and XL probe. In the obese group, NAFLD cirrhosis criteria spared 73.8% of unnecessary EGD by using M probe and 70.8% by using XL probe, also showing a good similar rate of missed VNT (1.3% and 2.8%, respectively) and missed any grade EV (15.4% for both) (Table 3). Notably, when looking at the group of nonobese patients, while NAFLD cirrhosis criteria by M or XL probes showed a similar performance, their accuracy was worse than that observed in obese patients. Specifically, in nonobese patients, NAFLD cirrhosis criteria spared 59.1% of unnecessary EGD by using M probe and 54.5% by using XL probe, providing a rate of missed VNT of 10.8% and 10.5%, respectively, and missed any grade EV of 23.1% and 25%, respectively (Table 3). To strength the robustness of these

results, we analyzed separately patients with clinical or histological diagnosis of NAFLD cirrhosis. Notably the worse performance of NAFLD cirrhosis criteria was confirmed in both groups. Specifically, in nonobese patients with histological (n=34) or clinical (n=76) diagnosis of NAFLD cirrhosis the rate of missed VNT was 10.5 and 11.1% for the M probe, and 5.5% and 11.9% for the XL probe, respectively.

Similar results by splitting patients according to obesity were observed for both M and XL probe-based NAFLD cirrhosis criteria in the entire cohorts of patients with LSM by M or XL probes (data not showed).

The different accuracy of NAFLD cirrhosis criteria according to obesity was also confirmed for Baveno VI and expanded Baveno VI criteria, even if Baveno VI criteria kept a rate of missed VNT <5% sparing only 20% of EGD (Table 3). Again, these results were confirmed in the entire cohorts of patients with LSM by M or XL probes, and in nonobese patients split for clinical or histological diagnosis of NAFLD cirrhosis (data not showed).

Characteristics of patients with missed VNT or any grade EV

In the entire cohort of 652 patients with LSM by M probe, NAFLD cirrhosis criteria missed 66 patients (15.5%) with small EV and 19 patients (4.4%) with VNT. Notably, when considering patients full filling NAFLD cirrhosis criteria, those with missed VNT, when compared to those without EV had significantly lower prevalence of obesity (26% vs 61%, p=0.007), and significantly lower PLT

(174±71 vs 213±82, p=0.03) and higher LSM (20±5.7 vs 16.9 vs 5.6, p=0.02) values (Table 4A). Similar results were observed for patients with missed small EV who had significantly lower PLT (182±68 vs 213±82, p=0.003) and higher LSM (18.5±5.8 vs 16.9±5.6, p=0.04) values compared to patients without EV (Table 4A).

When looking at the 452 patients with LSM by XL probe, NAFLD cirrhosis criteria missed 37 patients (13.8%) with small EV and 11 patients (4.1%) with VNT. Notably, among patients fulfilling NAFLD cirrhosis criteria, those with missed VNT when compared to those without EV had significantly lower prevalence of obesity (46% vs 78%, p=0.03), and significantly lower PLT (146±27 vs 217±77, p=0.002) and higher LSM (16.4±6.4 vs 13.7±5.0, p=0.08) values (Tables 4B). Similar results were observed for patients with missed small EV who had significantly lower PLT (186±63 vs 217±77, p=0.02) and a trend for higher LSM (15.2±5.6 vs 13.7±5.0, p=0.09) values compared to those without EV (Table 4B).

Discussion

In this study on a large multicenter cohort of patients with NAFLD-related compensated cirrhosis we confirmed that expanded Baveno VI works better than Baveno VI criteria for ruling-out the presence of VNT, sparing more EGD even if slightly increasing the number of missed VNT. We also elaborated new more accurate criteria, optimized for M and XL FibroScan probes and always based on LSM and PLT values, that were able to spare more than 50% of unnecessary EGD.

Portal hypertension significantly affects the natural history of liver cirrhosis of any etiology, including NAFLD [4]. Along this line, the identification of EV by EGD or non-invasive tools is a key clinical need. Baveno VI expert recommendations suggest to avoid EGD screening for VNT in patients with LSM<20 KPa and PLT >150,000 mmc [6], and expanded Baveno VI criteria have been also suggested [8].

In the present study, we found that Baveno VI and expanded Baveno VI criteria spared 33.8% and 58% of unnecessary EGD by missing 0.9% and 3.8% of VNT, respectively, these data being replicated in the validation cohort. Our study firstly validated Baveno VI and expanded Baveno VI criteria in the clinical setting of NAFLD-related cirrhosis, overall confirming that the updated criteria work better than the original Baveno VI, even if slightly increasing the number of missed VNT. Notably, our data demonstrated a rate of spared EGD for Baveno VI criteria higher than the overall 20% reported in patients with cirrhosis due to viral and/or alcohol etiologies [7], also showing a higher rate of spared EGD for NAFLD by using the expanded Baveno VI criteria [8]. These data confirmed the similar trend observed by Augustin and colleagues in a small cohort of 90 patients with NAFLD-related advanced chronic liver disease [8]. This issue could be explained by the impact of viral infections and alcohol use on fluctuation of PLT values, and on hepatic inflammation.

Baveno VI and expanded Baveno VI criteria were proposed and extensively validated in patients with cirrhosis mostly due to viral infections and/or alcohol abuse, and by considering LSM values obtained only by using the M probe. For these reasons, we tested for new criteria directly elaborated in NAFLD-related cirrhosis and differentiated for M and XL probes. Notably, while keeping 110,000 as the best threshold for PLT, we identified in 30 KPa for M probe and 25 KPa for XL probe (the same threshold of LSM

applied in the expanded Baveno VI criteria for M probe) the best LSM thresholds, finally generating the new NAFLD cirrhosis criteria. When considering the use of M probe, the NAFLD cirrhosis criteria performed better than Baveno VI and expanded Baveno VI allowing to spare 68.5% of unnecessary EGD while keeping a similar rate of missed VNT, and replicating these results in the validation cohort. The better accuracy of a higher threshold of LSM compared to that of the expanded Baveno VI could be related to the fact that the LSM in NAFLD patients is directly related not only to fibrosis and –probably- portal hypertension but also to severity of steatosis [9], obesity [10] and higher skin-to-capsule length [15], all factors frequently observed in NAFLD patients. Furthermore, NAFLD cirrhosis criteria with modified LSM threshold for XL probe showed a similar good diagnostic accuracy for ruling out patients with VNT. These criteria used a lower LSM stiffness value compared to that identified by using the M probe. This is not surprising, because, as observed for fibrosis [13], also in a setting of cirrhotic patients, LSM values from XL probe were significantly lower than those obtained by using the standard M probe. Notably, the good and better –with respect to Baveno and Baveno VI- accuracy of NAFLD cirrhosis criteria for avoiding unnecessary EGD while keeping an acceptable rate of missed VNT, was largely validated in the at higher risk cohort of patients with LSM≥13 kPa and in all participating cohorts considered separately, some of these analyses being however limited by the small sample size.

In the present study we found a higher prevalence of VNT in nonobese compared to obese patients. This data, even if all included patients had compensated Child A cirrhosis, could suggest that nononbese patients, as expression of sarcopenia, had a more advanced disease – more severe portal hypertension, lower liver function, etc – unfortunately not measurable by Child Classification. Notably, from a diagnostic point of view, when splitting the population according to obesity, we confirmed that both M

and XL probe-based NAFLD cirrhosis criteria had a similar diagnostic accuracy in obese and nonobese patients. However, we also observed that these criteria –as also reported for expanded Baveno VI- had a worse performance in nonobese patients where the rate of missed VNT was of about 10%. This picture was further confirmed by the evidence that BMI was significantly lower in patients with missed VNT compared those without EV and full filling NAFLD cirrhosis criteria. Differences in skin-to-capsule distance and/or steatosis severity between obese and nonobese patients could explain discrepancies in the observed diagnostic performances. From a clinical point of view, these data, even if worthy of further validation, suggest to careful interpret non-invasive scores for ruling out VNT in nonobese NAFLD patients with cirrhosis, where Baveno VI criteria seems to be more conservative even if sparing only 20% of EGD. Anyway, the lack of a fully understanding of this phenomena suggests caution in the interpretation of this result.

Another issue from the present study is that by applying Baveno VI, expanded Baveno VI and NAFLD cirrhosis criteria, designed to rule-out VNT, we missed about 10%-15% of patients with EV not needing of treatment. This data is clinically relevant because the presence of EV significantly affect the prognosis of cirrhotic patients via a higher risk of liver decompensation and death [4,5]. Notably, when looking at M or XL probe-based NAFLD cirrhosis criteria, patients with missed EV not needing of treatment had significantly higher LSM and lower PLT values compared to patients without EV and considered not worthy of EGD. These data potentially suggest that some of these patients when retested at one year, as suggested by Baveno VI recommendations, could transit in the area of patients considered worthy of EGD screening.

This study has some limitations. First, the analysis of the data was retrospective and the evaluation of EV size has been performed by several endoscopists, not testing interobserver agreement. In addition lack of data about gastric varices and portal

hypertensive gastropathy could further affect the interpretation of our results. However, the same issues was present in other multicenter large studies assessing non-invasive prediction of VNT in cirrhosis [8], and both prevalence of VNT and performance of tested scores was similar among centers. Second, interobserver concordance of LSM examinations was also not assessed, this issue potentially affecting the interpretation of our results. However, all tests were performed by expert operators following the same protocol and fulfilling validity criteria. Moreover, different relevant studies assessing FibroScan in NAFLD were based on multicenter cohorts and/or on multiple operators [12,13] .Finally, many different studies reported good interobserver concordance for LSM and CAP [16,17]. Another limitation of our study lies in the fact that we cannot rule-out that some patients with reliable LSM by M probe could have a skin-to-capsule distance >2.5 cm, this issue potentially affecting the accuracy of our results. The maximum 6 month interval between LSM and EGD could further bias our results. A strength of our paper is the availability of both M and XL FibroScan probes in a large population of NAFLD patients. Next generation FibroScan devices directly drive the use of M and XL probe and devices under development will adjust LSM values eliminating differences related to probes. However available devices will be still used for many years so we are confident that our study adds relevant insights in this topic. A further methodological issue is the potentially limited external validity of the results for different populations and settings. Our study included a cohort of NAFLD-related compensated cirrhotic subjects, largely obese, who were referred to tertiary centers for liver disease, limiting the applicability of the results in different populations, and especially in Asiatic populations largely under-represented in this study. Finally, lack of follow-up data may limit the potency of our results.

In conclusion, we demonstrated that, in a large cohort of patients with compensated cirrhosis due to NAFLD, the new elaborated NAFLD cirrhosis criteria differentiated for FibroScan probe, can safely spare more than 50% of unnecessary EGD, working better than the already proposed Baveno VI and expanded Baveno VI criteria. Prospective studies are needed to further validate our results.

References

1. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global Epidemiology of Non-Alcoholic Fatty Liver Disease-Meta-Analytic Assessment of Prevalence, Incidence and Outcomes. Hepatology. 2015 Dec 28. doi: 10.1002/hep.28431.
2. Dyson J, Jaques B, Chattopadyhay D, et al. Hepatocellular cancer: the impact of obesity, type 2 diabetes and a multidisciplinary team. J Hepatol. 2014 Jan;60(1):110-7.
3. Wong RJ, Aguilar M, Cheung R, et al. Nonalcoholic steatohepatitis is the second leading etiology of liver disease among adults awaiting liver transplantation in the United States. Gastroenterology. 2015 Mar;148(3):547-55.
4. D'Amico, G., G. Garcia-Tsao, and L. Pagliaro. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. J Hepatol, 2006. 44(1): p. 217-31.
5. Di Marco, V. Calvaruso V, Ferraro D, Bavetta MG, Cabibbo G, Conte E et al., Effects of Viral Eradication in Patients With Hepatitis C Virus and Cirrhosis Differ With Stage of Portal Hypertension. Gastroenterology, 2016.

6. De Franchis, R. Expanding consensus in portal hypertension Report of the Baveno VI Consensus Workshop: Stratifying risk and individualizing care for portal hypertension. Journal of Hepatology, 2015. 63: p. 743–752.
7. Augustin S, Pons M, Genesca J. Validating the Baveno VI recommendations for screening varices. J Hepatol. 2017 Feb;66(2):459-460.
8. Augustin S, Pons M, Maurice JB, Bureau C, Stefanescu H, Ney M, et al. Expanding the Baveno VI criteria for the screening of varices in patients with compensated advanced chronic liver disease. Hepatology. 2017 Dec;66(6):1980-1988.
9. Petta S, Maida M, Macaluso FS, Di Marco V, Cammà C, Cabibi D, et al. The severity of steatosis influences liver stiffness measurement in patients with nonalcoholic fatty liver disease. Hepatology. 2015 Oct;62(4):1101-10.
10. Caussy C, Chen J, Alquiraish MH, Cepin S, Nguyen P, Hernandez C, et al. Association Between Obesity and Discordance in Fibrosis Stage Determination by Magnetic Resonance vs Transient Elastography in Patients with Non-alcoholic Liver Disease. Clin Gastroenterol Hepatol. 2017 Nov 2. pii: S1542-3565(17)31304-6. doi:10.1016/j.cgh.2017.10.037.
11. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;411:313-321.
12. Wong VW, Vergniol J, Wong GL, Foucher J, Chan HL, Le Bail B, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. Hepatology. 2010 Feb;51(2):454-62.
13. Wong VW, Vergniol J, Wong GL, Foucher J, Chan AW, Chermak F, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. Am J Gastroenterol. 2012 Dec;107(12):1862-71.
14. Boursier J, Zarski JP, de Ledinghen V, et al; Multicentric Group from ANRS/HC/EP23 FIBROSTAR Studies. Determination of reliability criteria for liver stiffness evaluation by transient elastography. Hepatology. 2013;57:1182-91.
15. Shen F, Zheng RD, Shi JP, Mi YQ, Chen GF, Hu X, et al. Impact of skin capsular distance on the performance of controlled attenuation parameter in patients with chronic liver disease. Liver Int. 2015 Nov;35(11):2392-400.

16. Afdhal NH, Bacon BR, Patel K, Lawitz EJ, Gordon SC, Nelson DR, et al. Accuracy of fibroscan, compared with histology, in analysis of liver fibrosis in patients with hepatitis B or C: a United States multicenter study. Clin Gastroenterol Hepatol. 2015 Apr;13(4):772-9.e1-3.
17. Boursier J, Konate A, Guilluy M, Gorea G, Sawadogo A, Quemener E, et al. Learning curve and interobserver reproducibility evaluation of liver stiffness measurement by transient elastography. Eur J Gastroenterol Hepatol. 2008 Jul;20(7):693-701.

Table 1. Baseline Demographic, Metabolic, Laboratory and Instrumental Features of patients with NAFLD-related cirrhosis.

	NAFLD Cirrhosis Entire Cohort N=790	NAFLD Cirrhosis Training Cohort N=314	NAFLD Cirrhosis Validation Cohort M Probe N=338	NAFLD Cirrhosis Validation Cohort XL Probe N=138	P value°	P value^{ç}
Mean Age – years	62 ± 10	63 ± 11	62 ± 10	2 ± 11	0.593	0.406
Male Gender	55 %	60 %	55 %	44 %	0.239	0.001
Mean BMI – Kg/m^2	32.6 ± 6.7	33.3 ± 6.7	30.3±5.0	37.5 ± 7.6	<0.001	<0.001
ALT – IU/L	55 ± 42	54 ± 41	59 ± 44	46 ± 36	0.104	0.070
PLT – 10^3/mmc	177 ± 85	177 ± 80	178 ± 90	176 ± 82	0.785	0.980
Total Bilirubin – mg/dL	1.0±0.5	1.0±0.4	1.0±0.5	1.0±0.6	0.981	0.942
INR	1.0±0.2	1.0±0.4	1.0±0.2	1.0±0.2	0.954	0.896
Albumin – g/L	4.1 ± 0.5	4.1 ± 0.4	4.2 ± 0.5	3.9 ± 0.6	0.312	<0.001
Blood Glucose – mg/dL	129 ± 54	133 ± 61	125 ± 48	127 ± 47	0.126	0.430
Total Cholesterol – mg/dL	178 ± 49	183 ± 46	178 ± 51	164 ± 48	0.186	<0.001
Triglycerides – mg/dL	161± 123	183 ± 127	141 ± 120	162 ± 109	<0.001	0.136
Type 2 Diabetes	62.8 %	64.9%	62.0 %	59.8%	0.480	0.352
Arterial Hypertension	61.2%	60.8 %	63.2%	56.7%	0.558	0.480
LSM by M probe– Kpa*	25 ± 16	24 ± 15	26 ± 17	-	0.264	-
LSM by XL probe- KPa$^{\#}$	22 ± 16	20 ± 14	-	28 ± 17	-	<0.001
Esophageal Varices						
Presence	31.3 %	26.4 %	36.4 %	29.7 %	0.008	0.545
VNT	11.5 %	10.2 %	13.3 %	10.1 %	0.266	1.000

Abbreviations: BMI, body mass index; PLT, platelet; ALT, alaninoaminotransferase; LSM, liver stiffness measurement; VNT, Varices needed of treatment. Data are given as mean ± standard deviation, or as percentage of cases (%).

*Data relative to 652 patients; $^{\#}$Data relative to 452 patients; ° comparison between Training Cohort and Validation Cohort M Probe; ^{ç} comparison between Training Cohort and Validation Cohort XL Probe.

Table 2. Diagnostic Accuracy for VNT of Baveno VI, Extended Baveno VI and NAFLD Cirrhosis Criteria in Training and Validation Sets of Patients with NAFLD-related Compensated Cirrhosis.

A		Training Set M Probe	Validation Set M Probe	B		Training Set XL Probe	Validation Set XL Probe
		N = 314	N = 338			N = 314	N = 138
	VNT (%)→	32 (10)	45 (13)		VNT (%)→	32 (10)	14 (10)
Baveno VI Criteria PLT> 150 X10³ and LSM < 20 KPa	Spared Endoscopies (%)	106(33.8)	113(33.4)			-	–
	Missed VNT n. (%)	1(0.9)	5(4.4)				
	Missed EV (%)	11(10.4)	21(18.6)				
	sensitivity	0.97	0.89				
	specificity	0.37	0.37				
	Negative Predictive Value	0.99	0.96				
	Positive Predictive Value	0.15	0.18				
	Positive Likelihood Ratio	1.54	1.41				
	Negative Likelihood Ratio	0.08	0.30				
Extended Baveno VI Criteria PLT> 110 X10³ and LSM < 25 KPa	Spared Endoscopies (%)	182(58)	183(54.1)			-	-
	Missed VNT n. (%)	7(3.8)	8(4.4)				
	Missed EV (%)	30(16.5)	39(21.3)				
	sensitivity	0.78	0.82				
	specificity	0.62	0.60				
	Negative Predictive Value	0.96	0.96				
	Positive Predictive Value	0.19	0.24				
	Positive Likelihood Ratio	2.06	2.04				
	Negative Likelihood Ratio	0.35	0.30				
NAFLD Cirrhosis Criteria PLT> 110 X10³ and LSM < 30 KPa	Spared Endoscopies (%)	215(68.5)	209(61.8)	**NAFLD Cirrhosis Criteria PLT> 110 X10³ and LSM < 25 KPa**		204(65)	64(46.4)
	Missed VNT n. (%)	9(4.2)	10(4.8)			10(4.9)	1(1.6)
	Missed EV (%)	38(1	47(22.5)			37(18.1)	11(17.2)
	sensitivity	7.7)	0.78			0.69	0.93
	specificity	0.72	0.68			0.69	0.51
	Negative Predictive Value	0.73	0.95			0.95	0.98
	Positive Predictive Value	0.96	0.27			0.20	0.18
	Positive Likelihood Ratio	0.23	2.42			2.20	1.89
	Negative Likelihood Ratio	2.67	0.33			0.45	0.14
		0.39					

Table 3. Diagnostic Accuracy for VNT of Baveno VI, Extended Baveno VI and NAFLD Cirrhosis Criteria in Obese and Nonobese Patients with NAFLD-related Compensated Cirrhosis and with LSM by both M and XL Probes.

A		BMI >= 30 M Probe N = 202	BMI < 30 M Probe N = 110	B		BMI >= 30 XL Probe N = 202	BMI < 30 XL Probe N = 110
	EVNT (%)→	**12 (6)**	**19 (17)**		**VNT (%)→**	**12 (6)**	**19 (17)**
NAFLD Cirrhosis Criteria PLT> 110 X10³ and LSM < 30 KPa	Spared Endoscopies (%)	149(73.8)	65(59.1)	**NAFLD Cirrhosis Criteria PLT> 110 X10³ and LSM < 25 KPa**		143(70.8)	60(54.5)
	Missed VNT n. (%)	2(1.3)	7(10.8)			4(2.8)	6(10)
	Missed EV (%)	23(15.4)	15(23.1)			22(15.4)	15(25)
	sensitivity	0.83	0.63			0.67	0.68
	specificity	0.77	0.64			0.73	0.59
	Negative Predictive Value	0.99	0.89			0.97	0.90
	Positive Predictive Value	0.19	0.27			0.14	0.26
	Positive Likelihood Ratio	3.68	1.74			2.49	1.68
	Negative Likelihood Ratio	0.22	0.58			0.46	0.53
Baveno VI Criteria PLT> 150 X10³ and LSM < 20 KPa	Spared Endoscopies (%)	82(40.6)	23(20.9)				
	Missed VNT n. (%)	0	1(4.3)				
	Missed EV (%)	7(8.5)	4(17.4)				
	sensitivity	1	0.95				
	specificity	0.43	0.24			-	-
	Negative Predictive Value	1	0.96				
	Positive Predictive Value	0.1	0.21				
	Positive Likelihood Ratio	1.76	1.25				
	Negative Likelihood Ratio	0	0.22				
Extended Baveno VI Criteria PLT> 110 X10³ and LSM < 25 KPa	Spared Endoscopies (%)	127(62.9)	54(49.1)				
	Missed VNT n. (%)	0	7(13)				
	Missed EV (%)	16(12.6)	14(25.9)				
	sensitivity	1	0.63				
	specificity	0.67	0.52			-	-
	Negative Predictive Value	1	0.87				
	Positive Predictive Value	0.16	0.21				
	Positive Likelihood Ratio	3.02	1.31				
	Negative Likelihood Ratio	0	0.71				

Table 4. Differences in Baseline Demographic, Metabolic, Laboratory and Instrumental Features of patients with NAFLD-related cirrhosis full filling NAFLD Cirrhosis Criteria and split according to absence of EV or presence of Small EV or VNT.

(A) NAFLD Cirrhosis Criteria by M Probe

	Small EV	VNT	No EV	P value*	P value°
	N=66	N=19	N=339		
Male Gender (%)	42 (64)	13 (68)	178 (53)	0.12	0.26
Age - years	64.3±9.8	66.5±9.1	60.8±10.8	0.01	0.02
PLT – 10^3/mmc	182±68	174±71	213±82	0.003	0.03
BMI – Kg/m^2	32.8±6.3	27.9±3.9	32.5±6.5	0.79	0.002
BMI≥30 Khìg/m^2 (%)	41 (66)	5 (26)	194 (61)	0.50	0.007
LSM - KPa	18.5±5.8	20.0±5.7	16.9±5.6	0.04	0.02
ALT - IU	60±47	57±33	61±47	0.83	0.72

(B) NAFLD Cirrhosis Criteria by XL Probe

	Small EV	VNT	No EV	P value*	P value*
	N=37	N=11	N=220		
Male Gender (%)	26 (70)	8 (73)	120 (55)	0.10	0.38
Age - years	63.3±9.2	66.6±11.3	60.1±11.4	0.10	0.06
PLT – 10^3/mmc	186±63	146±27	217±77	0.02	0.002
BMI – Kg/m^2	33.6±6.4	30.0±4.6	35.8±7.4	0.09	0.01
BMI≥30 Khìg/m^2 (%)	23 (66)	5 (46)	159 (78)	0.19	0.03
LSM - KPa	15.2±5.6	16.4±6.4	13.7±5.0	0.09	0.08
ALT - IU	57±48	50±32	57±46	0.99	0.62

***Comparison between patients with small EV and those without EV; °Comparison between patients with VNT and those without EV.**