

EASL CLINICAL PRACTICE GUIDELINES (CPGs) ON NON-INVASIVE TESTS FOR EVALUATION OF LIVER DISEASE SEVERITY AND PROGNOSIS- 2020 Update

List of panel members (alphabetical order): Annalisa Berzigotti, Jerome Boursier, Laurent Castera, Nora Cazzagon, Mireen Friedrich-Rust, Salvatore Petta, Maja Thiele, Emmanouil Tsochatzis

Introduction

Liver fibrosis development marks the course of chronic liver disease and its presence and severity correlate with prognosis across etiologies [1-3]. The presence of cirrhosis identifies patients who are at risk of developing clinical decompensation and liver-related mortality [4], and who are at the highest risk of developing hepatocellular carcinoma, irrespective of the etiology of chronic liver disease. Liver biopsy is still the reference method for the assessment of liver fibrosis and allows a detailed evaluation of the localization and amount of fibrosis. The evidence supporting the use of liver biopsy has been reviewed in detail previously [5]. Although it provides extensive information and remains a key tool in hepatology, the liver biopsy specimen size has to be long enough and has to be interpreted by experts to provide reliable information [6]. In the field of NASH, variability among pathologists exists [7]. In addition to these technical considerations, liver biopsy is invasive, and as such it can lead, even if rarely, to severe complications. This, added to the relatively high cost, make non-invasive, repeatable and ideally cheaper alternative tools for the assessment of fibrosis highly desirable. Importantly, diagnostic methods for fibrosis in chronic liver disease should have low inter- and intra-rater variance in order to allow a comparison over time, since fibrosis is a dynamic process [8], which can regress. Non-invasive tests (NITs) should also hold prognostic information beyond fibrosis stage and allow monitoring of liver fibrosis and its complications.

The available NITs for the diagnosis and staging of liver fibrosis have been reviewed extensively elsewhere and in the previous EASL-ALEH CPGs [5], and a complete description is beyond the scope of the present CPGs update.

In brief, NITs for the assessment of chronic liver disease can be classified in:

- a) blood-based tests (serum markers of fibrosis; laboratory variables);
- b) methods assessing physical properties of the liver tissue (e.g.: liver stiffness; attenuation; viscosity);

c) imaging methods assessing the anatomy of the liver and other abdominal organs. These approaches can be considered complementary in several clinical scenarios. It should be underlined that NITs, liver biopsy/invasive diagnostic methods, and clinical acumen have to be integrated to achieve correct diagnoses and risk stratification in chronic liver diseases.

Considerations on diagnostic accuracy, advantages and limitations of NITs for the assessment of chronic liver disease

It has to be underlined that the sensitivity, specificity, PPV and NPV of a test depend on the prevalence of the condition under evaluation in the referral population [9]. The accuracy of diagnostic tests for fibrosis and steatosis in chronic liver disease is usually evaluated by comparing their sensitivity and specificity and area under the receiver operator characteristic curve (AUROC) with liver biopsy as the reference standard. However, liver biopsy is not a perfect reference standard (see above), and it has been shown that an AUROC >0.90 is not achievable even for a perfect biomarker [10].

A test should be able to correctly classify at least 80% of patients, and cut-offs with high sensitivity or high specificity should be chosen according to the clinical scenario (e.g. very sensitive cut-off avoiding false negatives if a given condition- e.g. cirrhosis- has to be ruled-out; Table 1). An AUROC below 0.80 is generally considered of too poor discriminatory accuracy to be of value in clinical practice. The calibration, or variance (goodness-of-fit, inter- or intra-operator variance) of a NIT is also important. Tests with poor reproducibility will result in imprecise measurements of little value for individual decision making. None of the existing NITs are ideal, and each of them has specific advantages and limitations.

There are several critical issues that should be considered when using NITs: availability, cost, and “context of use”. For instance, non-patented serum biomarkers, which are based on simple, inexpensive and widely available parameters, are well suited for use by non-specialists for testing for liver fibrosis in large populations in primary healthcare settings or diabetes clinics. On the contrary, sophisticated techniques like MRE, which are time consuming and costly with limited availability, are more suited for use by specialists in tertiary referral centers and for research purposes. In addition, when evaluating the performance of NITs, context of use in which they have been validated and applicability,

which is defined as the sum of reliability (the percentage of interpretable tests) plus failure rate (absence of test results), should be taken into account.

Generally, non-patented serum-based tests are highly applicable (>95%) and reproducible among different centers, but their results can be influenced by extrahepatic chronic diseases. Further, most NITs were developed and validated in secondary or tertiary settings, not tested for a context of use in primary care or the general population.

Liver stiffness measurement can be obtained by different methods (stand-alone bedside device: vibration controlled transient elastography-TE; techniques integrated in ultrasound devices: point-shear wave elastography (pSWE), bidimensional shear wave elastography (2D-SWE); and magnetic resonance elastography (MRE). TE is the most widely validated and available. TE using the appropriate probe and the other US-based measurement have an applicability of >95% (in patients who are not morbidly obese), provide results in real time and only require few minutes to be performed. In addition, they require a relatively short training. However, liver stiffness is a physical property of the tissue, which depends not only on the amount of liver fibrosis but also by several other factors. Therefore, results of liver stiffness measurement can overestimate fibrosis in case of inflammation, obstructive cholestasis, food ingestion, exercise, or venous congestion. These should be carefully excluded to avoid misdiagnosis. Meal ingestion increases liver stiffness values irrespective of the method used for its measurement. A minimum of 2 hours fasting was previously recommended [5]. However, several studies have shown since that return to normal values required at least 3 hours [11-13]. Therefore a minimum of 3 hours fasting is required for a correct measurement and interpretation. Additional details and recommendations can be found in the previous version of the EASL CPGs on non-invasive tests for chronic liver disease [5]. They are added as Supplementary material.

Imaging methods routinely used in chronic liver disease include ultrasound-based techniques, computerized tomography-based techniques and magnetic resonance-based techniques. They require specific devices and training and suffer from technique-specific limitations (in brief: ultrasound: operator-dependent; abdominal air and obesity limit the exploration; CT scan: exposure to ionising radiation; MR: impossible in case of old metal prosthesis; cost still elevated, limited availability). Standard imaging methods did not prove accurate to diagnose initial stages of fibrosis.

General limitations of NITs include a suboptimal accuracy to diagnose mild and moderate fibrosis, and to adequately discriminate between adjacent stages of fibrosis [5]; further, we still lack NITs to diagnose subclinical hepatic inflammation and ballooning, and to mirror the exact severity of portal hypertension in compensated advanced chronic liver disease (cACLD) (Text box). Specific advantages and limitations of the individual tests are described extensively elsewhere [5], and are summarized in Table 2. Finally, test-retest reliability of NITs and potential impact on its use remain incompletely studied and this should be object of future research.

Text Box: Definition of compensated advanced chronic liver disease (cACLD)

The term cACLD has been proposed as an alternative term for patients with chronic liver disease at risk of developing clinically significant portal hypertension, to better reflect that the spectrum of severe fibrosis and cirrhosis is a continuum in asymptomatic patients, and that distinguishing between the two is often not possible on clinical grounds.

According to the Baveno VI consensus conference, LSM ≥ 10 kPa is suggestive of cACLD and ≥ 15 kPa is highly suggestive of cACLD.

Methodology used for the development of the present Clinical Practice Guidelines (CPGs).

Given the numerous recent publications reporting on the accuracy of existing and novel non-invasive tests to assess liver disease, the European Association for the study of Liver Disease (EASL) decided to update the previous clinical practice guidelines [5]. The EASL Governing Board has involved a panel of experts in this field to elaborate the present CPGs according to the new format recently adopted, based on PICO (**P** Patient, Population, or Problem; **I** Intervention, Prognostic Factor, or Exposure; **C** Comparison or Intervention (if appropriate), **O** Outcome) questions [14]. These CPGs are directed at consultant hepatologists, specialists in training, and general practitioners and refer specifically to adult patients. Their purpose is to provide guidance on the best available evidence on the use of NITs to assess chronic liver disease.

The panel has initially established the most relevant topics that needed to be addressed and updated taking into account the content of the previous EASL guidelines on this topic [5] and the evidence that has been published since their publication (April 2015) until

October 2020. The panel decided to have an approach based on the etiology of liver disease, since this allows comparing homogeneous groups of patients. The complexity of several cases of multifactorial disease was discussed, but the panel felt that evidence in this field is not strong enough to drive recommendations on NITs use in this scenario; the recommendations pertinent to the main etiology responsible for liver disease should be then applied, considering additional caution in the interpretation of the results. The main topics that the panel decided to address include the following, for which novel data are available:

- a) identification of cases of advanced liver fibrosis in the general population, which requires special considerations given the low prevalence in this setting;
- b) assessment of liver disease severity and prognosis in patients with excessive use of alcohol, since this is an increasing burden worldwide [15];
- c) assessment of liver disease severity and prognosis in patients with chronic hepatitis C after achieving sustained virological response, since guidance on this topic is an unmet need in hepatology;
- d) assessment of liver disease severity and prognosis in patients with NAFLD/NASH, as well as monitoring liver lesions under treatment, since NAFLD is massively increasing worldwide and novel therapies for NASH are being tested and will require the identification of the correct group of patients [16];
- e) assessment of liver disease severity and prognosis in patients with cholestatic and autoimmune liver disease (PBC, PSC, AIH), since these are emerging causes of liver disease [16] and were only partly addressed in the previous guidelines;
- f) assessment of compensated advanced chronic liver disease (cACLD) and portal hypertension, since the identification of this stage of the disease is key to improve patients' outcome [17].

The Panel decided to develop PICO questions with a homogeneous format for each section. PICO questions were sent to the Delphi panel composed by 19 international experts in hepatology, pathology, radiology and primary care from Europe, Asia and America, and by one patient, and were commented and voted using an online platform. Consensus of over 75% of voting members of the Delphi Panel was needed to consider a question approved.

Based on the PICO questions, a literature search was performed using PubMed, and expanding to Embase, Google Scholar, Scopus when needed. References from papers were searched and identified further. The initial key words were: “Non-invasive test” OR “elastography” OR “imaging” OR “serum markers” OR “magnetic resonance” OR “computerized tomography” AND “Liver cirrhosis” OR “Chronic liver Disease” OR “steatosis” OR “fibrosis”. Further, more specific key words were also utilized as: “NAFLD”, “NASH”, “SVR”, “PSC”, “PBC”, “Autoimmune hepatitis”, “decompensation”, “portal hypertension”, “cACLD”, “CSPH”, “Varices” for each specific topic of the guideline. The selection of references was based on appropriateness of study design, number of patients, and publication in peer review journals. Whenever available, meta-analyses were used; else, original data were privileged. The resulting literature database was made available to all members of the panel.

The Level of evidence based on the Oxford Centre for Evidence-based Medicine and the QUADAS-2 tool for accuracy of diagnostic studies were used to judge the quality of the evidence [18].

Each expert took responsibility and made proposals for statements for a specific section of the guideline and shared tables of evidence and text with the full panel.

The Panel met on two occasions, one during an international meeting, and one ad hoc at the EASL premises in Geneva, and had six ad hoc teleconferences for discussion and voting.

All recommendations were discussed and approved by all participants. The strength of the recommendations in these guidelines has been graded according to the Oxford Centre for Evidence-Based Medicine (OCEBM) [19]. The level of evidence (LoE) classifications and recommendations are therefore based in two categories: strong or weak. The CPGs were sent to the Delphi Panel for their review and vote. The vote result was taken into account as follows: less than 50% approval: re-write recommendation and resubmit to the Delphi panel; 50%-75% approval: re-write/improve the recommendation, but no resubmission to the Delphi panel; 75-90% approval: no need to re-write the recommendation but the document will take into account the comments $\geq 90\%$ approval: assumed as consensus, no change needed but small corrections possible.

The suggested changes were taken into account in a revised version, which was finally sent to the attention of the EASL Governing Board together with a response letter

regarding each of the points raised by the Delphi panel members. **The Delphi Panel agreement on each of the statements and recommendations is shown in the Appendix.**

The recommendations were subsequently approved by the EASL Governing Board.

This document is intended to be valid until April 2025 unless the EASL Governing Board indicates the need of an earlier update.

General Population

What is the accuracy of non-invasive scores in patients at risk of liver disease from low-prevalence populations as compared to liver biopsy?

Recommendations

- **Non-invasive fibrosis tests should be used for ruling out rather than diagnosing advanced fibrosis in low prevalence populations (LoE 1, Strong recommendation).**
- **Non-invasive fibrosis tests should be preferentially used in patients at risk of advanced liver fibrosis (such as patients with metabolic risk factors and/or harmful use of alcohol) and not in unselected general populations (LoE 2, Strong recommendation).**
- **ALT, AST and platelet count should be part of the routine investigations in primary care in patients with suspected liver disease, so that simple non-invasive scores can be readily calculated (LoE 2, Strong recommendation).**
- **The automatic calculation and systematic reporting of simple non-invasive fibrosis tests such as FIB-4, in populations at risk of liver fibrosis (individuals with metabolic risk factors and/or harmful use of alcohol) in primary care, is recommended in order to improve risk stratification and linkage to care (LoE 2, Strong recommendation).**

The development, validation and widespread use of non-invasive fibrosis tests has changed clinical practice in hepatology and has reduced the need for liver biopsies. Moreover, these tests are becoming increasingly available, while at the same time the epidemiology of chronic liver disease is changing, with NAFLD and alcohol-related liver disease (ALD) becoming the main cause of liver related morbidity and mortality. As a

consequence of the above, the context of use for non-invasive fibrosis tests is changing; they are increasingly used in populations at risk of liver disease to test for the presence of advanced fibrosis. The prevalence of advanced fibrosis in such settings is considerably lower compared to the prevalence seen in secondary/tertiary care, where these tests have been developed and validated. In a large meta-analysis of the diagnostic accuracy of non-invasive tests, which almost exclusively included studies performed in secondary care, the prevalence of advanced fibrosis was 37%, 29%, 19% and 51% in patients with chronic hepatitis B, chronic hepatitis C, NAFLD and ALD, respectively [20]. Particularly for NAFLD and alcohol-related liver disease, the prevalence of advanced fibrosis in unselected populations at risk is <5% [21] and <10% [22], respectively. The different context of use therefore raises the question of the diagnostic performance of these tests in populations with low prevalence of advanced fibrosis.

It is very likely that non-invasive fibrosis tests will have lower sensitivity and higher specificity when applied in populations with lower disease prevalence due to the well described spectrum effect [9], as shown in a study in patients with ALD from a primary and secondary care setting [22]. On the opposite, in secondary/tertiary care settings who are referred patients with more advanced disease, the positive predictive value of NITs is expected to be higher (higher a priori probability of observing true positive cases) (see Table 3). Therefore, in populations of low prevalence, non-invasive tests are far better for ruling out rather than diagnosing the presence of advanced fibrosis. This implies the use of at least two tiers of non-invasive fibrosis tests for selecting patients from low prevalence populations for further investigations and follow up in order to reduce false positive results. It also offers the possibility for the use of a simple non-invasive fibrosis test (such as FIB-4) in populations at risk for liver disease (such as patients with type 2 diabetes; potentially: people living with HIV), to rule out those with low probability of having advanced fibrosis and prompt further testing for those with indeterminate and positive results. Automatic calculation of such tests when liver blood tests are requested can potentially improve the risk stratification in patients at risk of advanced fibrosis. FIB-4 is simpler to calculate and performs better than other simple NITs in head-to-head comparisons, particularly in NAFLD. All simple NIT panels include AST, therefore AST, together with ALT and platelet count, should be routinely performed in primary care as part of the liver blood test panel.

In spite of their potential to act as 'gate-keeping test' in primary care liver fibrosis screening pathways, the simple fibrosis scores only include indirect markers of liver damage (AST,

ALT), risk factors (age, BMI, diabetes) or liver function and portal hypertension (platelet count, cholesterol) and are not direct markers of liver fibrosis. Consequently, physicians should not blindly use FIB-4 or similar indirect NITs as singular decision tools [23]; due to their easy testing, repeated measurement can be performed, and this strategy is currently being evaluated [24]. If a suspicion of liver disease remains even after a normal NIT value, the patients should be referred for more accurate testing.

In order to minimize the spectrum effect, it is essential that non-invasive tests are applied to populations with risk factors for liver disease rather than in unselected populations. This is because unselected populations have an increased range of potential differential diagnoses for positive results, which would normally be identified with closer patient evaluation and selection [23]. Moreover, it is essential that patients with abnormal liver blood tests are comprehensively investigated for the etiology of the abnormality before or in parallel with non-invasive fibrosis assessment [25].

Can non-invasive scores, serum markers, liver stiffness, and imaging methods identify advanced fibrosis in patients at risk of liver disease from low-prevalence populations as compared to clinical acumen?

Statement

- **Non-invasive scores, serum markers, liver stiffness and imaging methods can identify advanced fibrosis in patients at risk from low-prevalence populations significantly better than clinical acumen alone (LoE 1).**

Recommendations

- **Individuals at risk for advanced fibrosis due to metabolic risk factors and/or harmful use of alcohol should be entered in appropriate risk stratification pathways using non-invasive fibrosis tests (LoE 1, Strong recommendation).**
- **The selection of non-invasive tests and the design of diagnostic pathways for testing low prevalence populations for advanced fibrosis should be performed in consultation with a liver specialist (LoE 3, Strong recommendation)**

There have been several studies of non-invasive fibrosis tests in populations with variable risk factors for liver disease, from unselected to patients with several predefined risk factors. In a systematic review that included 19 studies, in which 11 non-invasive tests were evaluated, the prevalence of advanced fibrosis depended on the risk factors of the included cohorts [26]. Two studies performed in the general population identified advanced fibrosis in 0.9% of participants using FibroTest™ (cut-off 0.59) [27] and 2% using FibroScan® (cut-off 9.6 KPa) [28]. In studies targeting people at risk of NAFLD, the prevalence of advanced fibrosis ranged from 3.7% to 30% [29]. Significant fibrosis was present in 11-18% of people at risk of alcohol-related liver disease [30]. A study performed in 4,021 young adults (mean age 24 years) using Fibroscan®, revealed that 20% had suspected steatosis (CAP values ≥ 248 dB/m) and 2.7% suspicion of fibrosis (liver stiffness values ≥ 7.9 KPa)[31]. The above estimates are based on non-invasive tests, therefore the true prevalence of advanced fibrosis in such populations is at least 50% lower, taking into account the low prevalence of the target condition which results in suboptimal positive predictive values of the non-invasive tests (Table 1).

Liver biopsy was performed in selected patients who had a positive non-invasive test in some studies [26, 32-38]. In contrast, no patients who tested negative were biopsied, making it impossible to calculate the specificity of non-invasive tests for advanced fibrosis in the context of low prevalence populations. Conversely, not all patients with a positive test were biopsied and this could be due to a selection bias, leading to an overestimation of the sensitivity of non-invasive fibrosis tests in low prevalence populations.

In a study that tested 128 patients from primary centers of municipal alcohol rehabilitation, liver biopsy was performed in all subjects and the prevalence of advanced fibrosis was 6% [22]. The specificity of ELF™ (cut-off 10.5), FibroTest™ (cut-off 0.58), Fibroscan® (cut-off 15 KPa) and 2D-SWE (cut-off 16.4 KPa) for advanced fibrosis was 97%, 93%, 97% and 97% respectively with sensitivities of 75%, 63%, 86% and 88%, respectively.

The implementation of pathways to test populations at risk of advanced fibrosis results in a significant increase in the detection of cases with advanced fibrosis/cirrhosis, compared to the standard of care. In a study using Fibroscan® in patients with hazardous alcohol intake or type 2 diabetes in 4 general practices, the number of patients with cirrhosis doubled compared to the period before the study commencement [38]. In a community, pathways for patients with NAFLD using two-tier non-invasive testing with FIB-4 followed by ELF™ in patients with indeterminate FIB-4 results, improved the detection of advanced fibrosis 4-

fold and reduced unnecessary referrals by 88% [37]. Modeling suggests that only concordant non-invasive tests can produce diagnostic accuracy comparable to a liver biopsy and that currently single non-invasive tests do not have sufficient diagnostic accuracy, particularly for the diagnosis of cirrhosis [39]. Several cost-effectiveness analyses have shown that testing populations at risk for liver disease but with low prevalence of advanced fibrosis is cost-effective [21, 37, 40-43];[44].

The selection of non-invasive test in particular patients should be in accordance with the known indications and limitations of such tests (for instance avoid FibroTest™ in patients with Gilbert's or transient elastography in patients with heart failure). It is therefore advisable that hepatologists are involved and consulted when non-invasive tests and pathways are designed and implemented in populations at risk outside secondary care. Figure 1 summarizes an algorithm that could be used for such a selection.

Alcohol-related liver disease

What is the accuracy of non-invasive scores, serum markers, liver stiffness, and imaging markers to diagnose alcohol-related liver disease (liver fibrosis, alcoholic hepatitis and steatosis) compared to liver biopsy in patients with a chronic harmful use of alcohol?

Recommendations

- **In patients with alcohol-related liver disease LSM by TE <8 kPa is recommended to rule-out advanced fibrosis in clinical practice, with the following NITs as alternatives, if TE is not available (LoE 3; strong recommendation)**
 - **Patented tests: ELF™ <9.8 or FibroMeter™ <0.45 or FibroTest® < 0.48**
 - **Non-patented tests: FIB-4 <1.3**
- **Upon referral of patients at risk of alcohol-related liver disease, LSM by TE≥12-15 kPa is recommended to rule in advanced fibrosis, after considering causes of false positives (LoE 2; strong recommendation).**
- **In patients with elevated liver stiffness and biochemical evidence of hepatic inflammation (AST or GGT >2xULN), LSM by TE should be repeated after at**

least one week of alcohol abstinence or reduced drinking (LoE 3; strong recommendation).

Liver fibrosis

Alcohol-related liver disease is the dominant cause of liver-related mortality and morbidity worldwide [45]. Furthermore, patients with alcohol-related cirrhosis are diagnosed at later stages of disease, die at a younger age and are more likely to experience liver-related complications than patients with liver disease of any other etiology [46, 47]. Therefore, non-invasive tests for alcohol associated liver damage is appealing, as early disease detection could lead to reduced drinking, thereby interrupting disease progression [48]. The most robust evidence involves transient elastography for the diagnosis of advanced fibrosis in patients recruited from secondary and tertiary care centers. Since 2015, six single-etiology studies, [49-52] one Cochrane meta-analysis, [53, 54] and two individual patient data meta-analysis has assessed LSM by TE in ALD [55]. TE has excellent diagnostic accuracy for advanced fibrosis, with AUROCs above 0.90. For significant fibrosis, the diagnostic accuracy is good, with AUROCs around 0.85. Unfortunately, early diagnostic studies mostly assessed fibrosis using scoring systems developed for chronic viral hepatitis (METAVIR). These scores likely underestimate early stages of alcohol-related fibrosis, while for bridging fibrosis and cirrhosis, diagnostic estimates are probably reliable across histological fibrosis scoring systems. Therefore, our recommendations focus on advanced fibrosis.

The available evidence is mostly of moderate level, except one diagnostic test study at high level of evidence[22]. This study is also the only study to recruit patients from a primary care setting, which is why the main body of evidence concerns a population with a high prevalence of advanced fibrosis. A further concern of most early studies, is that most did not clearly exclude patients with obvious cirrhosis, thereby potentially overestimating diagnostic accuracies.

The Cochrane meta-analysis reported a 92% summary sensitivity for advanced fibrosis at a TE cut-off around 9.5 kPa (range 8.0 to 10.5 kPa)[53] while the most recent meta-analysis (n=5648 patients; ALD n=946) found a sensitivity of 94% at a 8 kPa cut-off [55]. Therefore, we recommend ruling out advanced fibrosis in patients with TE below 8-10 kPa. The most recent individual patient data meta-analysis reported a specificity of 92% for

advanced fibrosis at a cut-off of 15 kPa, and 89% for a 12 kPa cut-off, in line with the high-quality, single-center study [22] which found a specificity of 95% for advanced fibrosis at 15 kPa, and a corresponding PPV of 84% (23% prevalence of advanced fibrosis).

Consequently, advanced fibrosis may be suspected in ALD patients with TE \geq 12-15 kPa, but only after excluding causes of false positives.

Other technologies for liver stiffness measurement, pSWE and 2D-SWE, may perform similar or almost similar to TE, but only one center has done head-to-head comparison between TE and 2D-SWE (Supersonic Aixplorer) [56] and only two recent studies have assessed pSWE using the Virtual Touch technique (Siemens Acuson 2000), with no comparator [57, 58].

It has been debated whether active use of alcohol may cause false positive liver stiffness measurements. While abstinence reduces liver stiffness in detoxification studies, this reduction is paralleled by a reduction in biochemical markers of liver inflammation such as AST and GGT [59-62]. Consequently, it is the alcohol-related steatohepatitis rather than alcohol *per se*, which increases liver stiffness. In one study, a week of detoxification reduced TE by 22%, and TE correlated with AST and GGT both at baseline and after detoxification [62]. A second study reported a 16% decrease in TE after five days of hospitalisation for detoxification, in parallel with a 48% decrease in AST, from 77 to 40 U/L [60]. In outpatients, four weeks of detoxification led to a reduction in TE of 25%, together with a 29% reduction in AST, from 42 to 30 U/L, and a 58% reduction in GGT, from 153 to 64 U/L. In studies of diagnostic accuracy, the optimal cut-off values are 2-3 kPa higher in those with AST elevation 1-2xULN, and even more in patients with elevations $>$ 2xULN [49, 52]. In contrast, in an outpatient setting of ALD patients with little biochemical evidence of hepatic inflammation, active drinking was not a predictor of false positive TE measurements. Consequently, AST of more than twice ULN should raise caution for false positive liver stiffness measurements. In patients with elevated liver stiffness *and* biochemical evidence of liver inflammation, we therefore suggest repeating the measurement after at least one week of abstinence or reduced drinking, in parallel with biochemical retesting.

Several serum markers have also been evaluated for diagnosing alcohol-related liver fibrosis, both patented such as FibroTest®, Hepascore, FibroMeter™ and ELF™ test; and non-commercial algorithms of routine biochemistry such as FIB-4 and Forns'. FIB-4 and Forns' have good diagnostic accuracies for advanced fibrosis. Their low cost and wide

accessibility make them particularly suited to rule out advanced fibrosis in low-prevalence populations. This is supported by a NPV of 95% for FIB-4 <3.25, and a NPV of 97% for Forns' index <6.8, in a study of 128 primary care patients with a 6% prevalence of advanced fibrosis [22]. The value for ruling out advanced fibrosis in primary care is however only evaluated by one study, and so needs independent validation. Due to risk of misclassifications, the non-patented fibrosis scores cannot be recommended to rule in advanced fibrosis.

The patented markers have higher diagnostic accuracies than the non-patented, with AUROCs similar or close to LSM by TE, but cut-offs vary substantially from study to study and would therefore need to be aligned and validated. There is a similar lack of studies investigating combination markers, either in parallel [51], or sequential [22]. In cases of discrepancy between TE and patented serum markers, TE seems more reliable [22, 51].

Cost-benefit of using non-invasive tests for alcohol-related fibrosis

Of note, recent evidence suggests that TE and the ELF™ test are cost-beneficial in patients with an excess use of alcohol [41, 63]. The two studies both used 40-year males as exemplar, from Scandinavia and Spain, respectively. Both studies found cost-benefit of a sequential strategy using ELF™ followed by TE, if ELF™ is positive. The incremental cost-effectiveness ratios (ICERs) were €13,400 per QALY [63], and \$5,387-\$8,430 per QALY [41]. However, single use of TE was the most cost-beneficial strategy in secondary care in one study [41], while annual ELF™ alone was the optimal strategy for patients with ALD in another study [63].

Alcoholic hepatitis

The existing evidence since 2015 on non-invasive markers for diagnosing alcoholic hepatitis consists of only two studies of moderate evidence on cytokeratin-18 (CK-18) based markers of cell-death [60, 64], and one study on the AshTest [65]. The markers show moderate diagnostic accuracies (AUROCs of 0.84 and below), and the three studies are heterogeneous in their histological definition of alcoholic hepatitis, and in their patient cohorts. It is therefore not possible to recommend any NIT for use in patients suspected of AH.

In one study, both total (M65) and caspase-cleaved CK-18 (M30) correlated with histological ballooning, but they had inadequate diagnostic accuracy (AUROCs<0.80) for detecting patients with steatohepatitis, defined as a NAFLD activity score ≥ 5 [60]. Since

NAS includes steatosis, in addition to ballooning and lobular inflammation, and since NAS has been designed for NAFLD which although similar it is not identical with ALD histologically, the score is probably not suited as an outcome for alcohol-related hepatic inflammatory activity.

For diagnosing biopsy-verified alcoholic hepatitis, using the AHHS score [66], another study also tested the cytokeratin-18-markers M65 and M30 [64]. The cut-offs of M65 and M30 for ruling in alcoholic hepatitis were far higher than the cut-offs reported for diagnosing steatohepatitis, indicating a more severely ill patient population.

Steatosis

While steatosis remains a key feature of acute alcohol-related liver injury, it is not possible to recommend any non-invasive tests for diagnosing alcohol-related steatosis, as only one study exists [22]. They evaluated CAP using the FibroScan equipment. While CAP had superior diagnostic accuracy compared to bright liver echo pattern assessed by ultrasonography, the diagnostic accuracies were modest.

What is the accuracy of non-invasive scores, serum markers, stiffness of liver tissue, and imaging methods to predict liver-related outcomes compared to liver biopsy, HVPG, Child Pugh score or MELD, in patients with a chronic harmful use of alcohol?

Evidence from mixed-etiology studies suggests that NITs are prognostic in patients with compensated cirrhosis/compensated advanced chronic liver disease (cACLD; more details on this definition are provided elsewhere in this guideline). This is very likely the case for alcohol-etiology as well, although there is just one prognostic single-etiology study [67]. However, this study only reported FibroTest, FibroMeter and Hepascore, assessed liver-related death as the only outcome, and included almost one-third with cirrhosis at baseline, not clearly excluding those with evidence of decompensated disease. The prognostic values (AUROC for 8-year survival or non-liver disease-related death) were 0.79 for FibroTest, 0.80 for Fibrometer, and 0.78 for Hepascore.

Since 2015, 12 studies have explored prognostic markers in cohorts of patients with alcohol-related cirrhosis, either decompensated or a combination of decompensated and compensated cirrhosis [60, 68-78]. All studies are explorative, and most found that

MELD score performed better or equal to the marker under investigation. MELD remains the recommended prognostic tool for prediction of short-term mortality and morbidity in decompensated cirrhosis.

Due to scarce evidence, we cannot at this point make any etiology-specific recommendations regarding prognostic markers in alcohol-related, compensated liver disease

HCV Post-SVR/post-antiviral therapy

What is the accuracy of non-invasive scores, serum markers, liver stiffness, and imaging methods compared to liver biopsy to stage liver fibrosis in patients with HCV compensated advanced chronic liver disease (cACLD) who achieved sustained virological response?

Statement

- **Non-invasive scores and LSM by TE and other elastography methods are not accurate in detecting fibrosis regression after SVR in HCV patients diagnosed with cACLD prior to antiviral therapy (LoE 3)**

Recommendations

- **The routine use of non-invasive scores and LSM by TE and other elastography methods is currently not recommended to detect fibrosis regression after SVR in HCV patients (LoE 3; strong recommendation).**
- **Cut-offs of LSM by TE used in patients with untreated HCV should not be used to stage liver fibrosis after SVR (LoE 4; strong recommendation).**

Regression of fibrosis in HCV patients with cACLD has been described after SVR in patients treated with IFN-based therapies. A study in 38 HCV patients with cirrhosis with paired pre- and post-treatment liver biopsies (median interval 79 months) showed cirrhosis regression (decrease ≥ 1 METAVIR stage) in 61% of patients [79]. With the advent of DAA leading to SVR in most HCV patients with cirrhosis [80], fibrosis regression will likely

become even more common. However, post-SVR liver biopsies are not the standard of care. It is therefore a critical issue whether non-invasive methods are able to capture fibrosis regression and stage fibrosis after SVR in HCV patients with cACLD who still have residual risk of liver-related complications.

A recent meta-analysis including 24 studies (n=2934 HCV patients, SVR 75%, DAAs only n=6) with paired LSM by TE, reported a median relative LSM decline from baseline of 28% (IQR, 21.8–34.8), 6–12 months after the end of therapy in SVR patients whereas no change was observed in non SVR patients [81]. In the subgroup of 261 SVR patients classified as advanced fibrosis or cirrhosis (LSM >9.5 kPa), 47% had post-treatment LSM <9.5 kPa [81]. However, most of the included studies in this meta-analysis were retrospective, IFN-based, with small sample size, and short follow-up after SVR. In addition, LSM confounders such as NAFLD, diabetes and alcohol were not taken into account. Most importantly, only 187 patients had biopsy-proven cirrhosis and none had paired liver biopsies. It is consequently impossible to conclude whether the observed LSM decrease is related to resolution of hepatic inflammation or to regression of liver fibrosis. As for DAA, in a large real-life Italian cohort of 749 HCV patients with cACLD treated with DAA (SVR 97%), a significant LSM decrease was observed between baseline and SVR12 (mean LSM 19.3 (\pm 11.2) vs 14.2 (\pm 11.7) kPa, respectively) but with a short follow-up and no post-SVR liver biopsies [82]. Interestingly, in a study [83] with paired pre- and post-treatment LSM and liver biopsies (median interval 61 months) in 33 HCV patients with cirrhosis, the diagnostic accuracy of TE for diagnosing cirrhosis after SVR (cut-off 12 kPa) was suboptimal (95% specificity, but 61% sensitivity, meaning low value for ruling out cirrhosis). Another study in 112 patients with recurrent HCV infection post-OLT and with paired liver biopsies 12 months after SVR (84 out of them also with paired LSM by TE; 34 out of them with cirrhosis), showed that LSM decrease was significantly higher in patients with fibrosis regression compared to those without (47% vs 30%, p=0.02) [84]. However, the percentage of LSM decrease did not accurately predict fibrosis regression (AUROC=0.65) [84]. The same study also demonstrated that LSM by TE 1 year after SVR can accurately predict the presence of advanced fibrosis with an AUROC of 0.90. The best LSM cut-offs to rule-out and rule-in advanced fibrosis were, respectively, 10.6 and 14 kPa. [84]. Another issue is the significant variations in LSM using TE over time reported in untreated patients with chronic liver disease [85].

Similarly, post-SVR LSM decrease has been reported using other devices such as pSWE (Virtual Touch) [86-88] and MRE [89]. Along this line, post-SVR decreases has

been reported with non-invasive serum biomarkers like APRI, FIB-4 or ELF™ [84, 86, 88, 90]. These studies, even if with some contrasting results [90], also showed a good diagnostic accuracy of LSM by pSWE (AUROC from 0.88 and 0.91) [87, 88] and of APRI, FIB-4 and ELF™ [84, 86] for the diagnosis of advanced fibrosis after SVR and having liver biopsy as reference. It should be kept in mind, however, that thresholds for LSM and NITs used in untreated viral hepatitis have proven inaccurate after SVR [83, 84, 86, 87], and it is necessary to validate the newer (lower) cut-offs in larger studies. Based on the high specificity, and awaiting for further data, it seems reasonable to consider that patients with LSM >12 kPa after SVR have a high likelihood of persistent cACLD.

In summary, altogether these results question the accuracy of non-invasive tests to predict fibrosis regression and presence of cACLD after SVR. Studies with larger sample sizes and longer follow-up are necessary to establish the role of non-invasive methods in the follow-up of HCV patients with cACLD after viral clearance.

What is the accuracy of non-invasive scores, serum markers, liver stiffness, and imaging methods to predict clinical outcomes (decompensation; HCC) compared to liver biopsy, HVPG, Child Pugh score or MELD in patients with HCV compensated advanced chronic liver disease (cACLD) who achieved sustained virological response?

Statement

- **In patients with cACLD previous to antiviral therapy for HCV, LSM post-SVR could be helpful to refine the stratification of residual liver-related risk, and awaiting further confirmatory data yearly repetition of LSM can be done (LoE 3).**

Recommendations

- **Patients with cACLD previous to antiviral therapy for HCV should continue to be monitored for HCC and portal hypertension irrespective of the results of NITs post SVR (LoE 3; strong recommendation).**

In patients with HCV-related cACLD, SVR reduces the risk of liver-related complications such as hepatic decompensation, hepatocellular carcinoma (HCC) as well as all-cause mortality [91, 92]. However a residual risk of liver-related complications still persists after SVR, particularly HCC occurrence, and the role of non-invasive tools in stratifying this risk remains debated [91, 92].

The presence of clinically significant portal hypertension (CSPH, HVPG ≥ 10 mmHg), has been shown to be the strongest prognostic determinant in patients with cACLD [93]. It has been established that an HVPG reduction of 10% or more after therapy is associated with a decreased risk of first variceal hemorrhage [17, 94].

In a multicenter prospective study of 226 HCV patients with cirrhosis and CSPH, SVR after DAA therapy significantly reduced HVPG ($>10\%$ in 62% of patients) measured 24 weeks after therapy, compared to baseline [95]. However, CSPH persisted in most patients (78%) despite SVR, indicating persistent risk of decompensation. In another recent retrospective single center study in 90 HCV patients with portal hypertension (HVPG ≥ 6 mmHg), and SVR after DAA, a HVPG reduction $\geq 10\%$ was reported in 67 patients with pre-treatment CSPH and translated into a clinical benefit. In particular, patients who were compensated on inclusion and showed a HVPG decrease $\geq 10\%$, were completely protected from hepatic decompensation in the follow-up [96]. In the same cohort published earlier in 57 HCV patients with paired HVPG and TE, before and after SVR, the relative change in LSM was an independent predictor of a HVPG decrease $>10\%$ in the subgroup of patients ($n=40$) with baseline CSPH [97]. However, the performance of TE for diagnosing an HVPG reduction $\geq 10\%$ was inadequate (AUROC < 0.8) [96]. Similarly, in the Lens study [95], LSM decreased markedly at SVR24 and SVR96 [98], but changes in LSM did not correlate with HVPG changes, nor with the risk of clinical decompensation. CSPH persisted in up to 53-65% of patients at SVR96 [98]. Despite these negative data, LSM after SVR had a good accuracy for the diagnosis of post-treatment CSPH, with an AUROC ranging from 0.80 [95] to 0.93 [96]. Values of LSM by TE over 21-23 kPa were invariably associated with persistence of CSPH, while low LSM values did not rule-out CSPH (30% of patients with liver stiffness measurement < 13.6 kPa still had CSPH) [95, 96]. On the other hand, in the context of post-OLT, one year post-SVR LSM had a high diagnostic accuracy to rule-out CSPH (AUROC=0.88) [84]. Consistent with these data, a cohort study on 230 HCV cirrhotic patients on SVR by DAA and 151 of them with follow-up LSM and upper endoscopy suggested that LSM after SVR could predict varices progression after 36 months [99].

In a large retrospective single-center cohort of 505 HCV patients with cirrhosis treated with DAA and followed for a median time of 25 months, baseline LSM using TE independently predicted the occurrence of HCC at 3 years (20% vs. 5% in patients with LSM >30 kPa vs. LSM ≤30 kPa, respectively) [100]. When replacing LSM by FIB-4 in the model, FIB-4 ≥ 9 remained an independent predictor of HCC [100]. Another cohort study in 139 HCV patients with cirrhosis reported a lower LSM reduction, using TE, in patients developing HCC (median follow-up 15 months) with a difference in LSM from baseline to end-of-therapy lower than -30% being an independent predictor of HCC development [101]. Finally, a cohort study in 572 HCV patients with cACLD with SVR after DAA, showed that only few patients (5.6%) developed liver decompensation –all of them with baseline LSM >20 kPa, and that independent predictors of HCC occurrence were platelet count and LSM at one year of follow-up. Notably, the authors found that a follow-up LSM value <10 kPa, obtained in 40% of patients, identified a cohort at very low risk of HCC (<1/100 patient-years) [102].

Even if available evidence suggests that post-SVR LSM, using TE, can predict CSPH and HCC occurrence, given the significant LSM decrease observed after SVR, lower cut-offs should be defined and validated. Recent evidence suggests a decrease in liver-related events in patients with decrease of LSM after SVR [103]. Further studies are needed to investigate the ability of post-SVR non-invasive tests in predicting hepatic decompensation and death; awaiting further confirmatory data, we consider reasonable to perform yearly repetition of LSM.

NAFLD/NASH

What is the accuracy of non-invasive scores and imaging methods for the diagnosis of steatosis compared to liver biopsy in patients with metabolic risk factors and/or suspected NAFLD?

Statements

- **CAP is a promising point-of-care technique for rapid and standardized detection of steatosis. However, given its limited availability and lack of head-**

to-head studies compared to ultrasound, CAP cannot yet be recommended as first line technique (LoE 2).

- **Although there are no consensual cut-offs, values above 275 dB/m might be used to diagnose steatosis, since they showed over 90% sensitivity to detect steatosis (LoE 2).**

Recommendations

- **Non-invasive scores are not recommended for the diagnosis of steatosis in clinical practice (LoE 2; strong recommendation)**
- **Conventional ultrasound is recommended as first line tool for the diagnosis of steatosis in clinical practice, despite its well-known limitations (LoE 1; strong recommendation)**
- **MRI-PDFF is the most accurate non-invasive method for detecting and quantifying steatosis. However, it is not recommended as first line tool given its cost and limited availability. Therefore, it is more suited for clinical trials (LoE 2; strong recommendation)**

Several steatosis scores have been proposed for the detection of steatosis including the SteatoTest™, the fatty liver index (FLI), the hepatic steatosis index (HSI), the lipid accumulation product (LAP), the index of NASH (ION) and the NAFLD liver fat score (NAFLD-LFS) [104]. Although SteatoTest™, FLI, NAFLD-LFS, LAP and HSI have been independently validated [105-108], their diagnostic performances are difficult to compare. Indeed, they have been designed and validated against different standards: liver biopsy, ultrasonography, or magnetic resonance spectroscopy. Nevertheless, when FLI, NAFLD-LFS, and HSI were compared in a retrospective cohort of 324 patients with suspected NAFLD and liver biopsy, their diagnostic performances for detecting any steatosis (>5%) did not differ (AUROC 0.83, 0.80 and 0.81, respectively) [106]. Further studies are needed, but it should be acknowledged that these scores do not add much in clinical practice to the information provided by clinical, laboratory and imaging examinations done routinely in patients with suspected NAFLD.

Conventional ultrasonography (US) is the most commonly used imaging method for the diagnosis of steatosis, since it is widely available, innocuous, cheap and well

established [109]. In a large meta-analysis [110] (n=34 studies, 2815 patients with suspected or known liver diseases), pooled sensitivities and specificities of US to detect steatosis ($\geq 20\text{-}30\%$), taking liver biopsy as the reference, were 85% (80–89%) and 94% (87-97%), respectively. The main limitations of US are that it can only detect steatosis above 12.5-20% [111], is prone to inter operator variability and has reduced accuracy in patients with obesity [112].

Magnetic resonance proton density fat fraction (MRI-PDFF) is an accurate, reproducible, quantitative imaging-based technique that has the ability to quantify liver fat in its entire dynamic range [113]. Quantification of steatosis using MRI-PDFF highly correlates with magnetic resonance spectroscopy results [114]. In a recent meta-analysis (n=6 studies in 635 patients with biopsy-proven NAFLD) [115], the summary AUROC values of MRI-PDFF for detecting steatosis $\geq 5\%$, and $\geq 33\%$, $\geq 66\%$ were 0.98, 0.91, and 0.90, respectively. Pooled sensitivity and specificity were 93% and 94%, 74% and 90%, and 74% and 87%, respectively. Despite the high accuracy of MRI-PDFF for detecting and grading steatosis, cost and limited availability restrict its use in practice.

The ability to quantify steatosis by measuring ultrasonic attenuation of the echo wave, termed the controlled attenuation parameter (CAP), has been implemented on the FibroScan device [116]. In the first individual data meta-analysis [117] published (n=19 studies in 2735 patients (537 with NAFLD; 19.6%) with liver biopsies), AUROCs of CAP for detecting steatosis $\geq 5\text{-}10\%$, $\geq 33\%$ and $\geq 66\%$ were 0.82, 0.86, and 0.88, respectively. Pooled sensitivities were 69%, 77%, and 88% and specificities 82%, 81% and 78%, respectively. Optimal cutoffs of 248 dB/m, 268 dB/m and 280 dB/m were proposed but of note several covariates such as NAFLD, diabetes and BMI influenced CAP values. Nevertheless, the cut-off associated with significant steatosis ($>33\%$) was almost always >250 dB/m. In addition, most included studies were conducted in small sample size (<100 patients), heterogeneous populations (less than 20% with NAFLD) and were performed with the M probe. Two recent multi-center studies [118, 119] addressed the accuracy of CAP in large cohorts (n=393-450) of NAFLD patients, using M and XL probes as recommended by the device's automatic probe selection tool. Failure rate using the XL probe was much lower (3-4%) [118, 120] than that reported with the M probe (21%) [121]. Accuracy for detecting steatosis $\geq 5\%$ was good with AUROCs of 0.76-0.87. By contrast, accuracy was suboptimal for quantifying steatosis with AUROCs of 0.70-0.77 and 0.58-0.70 for steatosis $\geq 33\%$ and $\geq 66\%$, respectively. Cut-off values of 263 dB/m [119] and 274 dB/m [118] had high sensitivities and positive predictive values ($>90\%$) for detecting

steatosis ($\geq 5\%$). In a recent meta-analysis of individual data currently in press[122], CAP measured by the XL probe in 930 patients with NAFLD and histologically proven steatosis accuracy was good for identifying any grade of steatosis vs. absence of steatosis (AUROC 0.819; 95% CI 0.769-0.869), but suboptimal to differentiate mild steatosis from higher grades (S0-S1 vs. S2-S2; AUROC 0.754; 95% CI 0.720-0.787). According to this meta-analysis, the optimal cut-off (according to Youden's index) to detect any steatosis in NAFLD patients is 294 dB/m (Sens 0.790; Spec 0.740), but if a sensitivity of ≥ 0.90 was required, the cut-off dropped to 263 dB/m (95% CI 256-270)[122].

Quality criteria have been proposed (CAP IQR < 30 or 40 dB/m) [123, 124] but not externally validated [118]. When compared with MRI-PDFF for detecting and quantifying steatosis using liver biopsy as reference, CAP was outperformed by MRI-PDFF [125-127].

In summary, CAP is a promising point-of-care technique for rapid and standardized steatosis detection, with high applicability ($>95\%$) when using the XL probe. Although there are no consensual cut-offs, values above 275 dB/m have high sensitivities and PPV ($>90\%$) in NAFLD. However, CAP has suboptimal performance for quantifying steatosis and is outperformed by MRI-PDFF. CAP should be compared to US that, despite its limitations, remains the most widely used tool for first line steatosis detection.

What is the accuracy of non-invasive scores, serum markers, liver stiffness, and imaging methods to evaluate NAFLD severity (presence of NASH and staging of liver fibrosis) compared to liver biopsy?

Statement

In patients with NAFLD:

Liver biopsy remains the reference standard for the diagnosis of NASH, because none of the available non-invasive tests has acceptable accuracy (LoE 2)

Recommendations

In patients with NAFLD:

- **The following NITs are recommended to rule-out advanced fibrosis in clinical practice (LoE 1, strong):**
 - **LSM by TE <8 kPa**
 - **Patented tests: ELF™ <9.8 or FibroMeter™ <0.45 or FibroTest® < 0.48**
 - **Non-patented tests: FIB-4 <1.3 or NFS <-1.455**
- **Upon referral of a patient with FIB-4 over 1.3, the use of transient elastography and/or patented serum tests should be used to rule out/in advanced fibrosis (see Figure 1) (LoE 2, strong recommendation).**
- **MRE is the most accurate non-invasive method for staging liver fibrosis. However, it is only marginally better than other NITs for F3-F4 fibrosis and it is not recommended as first line NIT given its cost and limited availability (LoE 2; strong recommendation). Therefore, it is more suited for clinical trials.**

The diagnosis of nonalcoholic steatohepatitis (NASH) is clinically relevant because NASH is associated with faster liver fibrosis progression [109, 128]. Several serum markers or scores such as CK-18 fragments, combinations of clinical variables, combination of clinical variables with the PNPLA3 I148M variant, metabolomics or lipidomic-based scores, as well as imaging techniques have been proposed for the non-invasive diagnosis of NASH. However, contrasting results from literature, lack of validation studies, and lack of availability of some of the variables included in many scores limit the recommendation of the proposed tools in clinical practice [113, 129]. Thus, liver biopsy currently remains the reference standard for the diagnosis of NASH in NAFLD patients.

Liver fibrosis is the main prognostic driver in NAFLD patients, with advanced fibrosis being an independent risk factor for both hepatic and extrahepatic events and liver-related and global mortality [130] [131]. Thus, advanced liver fibrosis has been used as the main endpoint in studies on non-invasive tests in NAFLD patients. Proposed serum markers and scores for the assessment of fibrosis severity include NAFLD Fibrosis score (NFS), Fibrosis-4 (FIB-4), BARD score, AST to platelet ratio (APRI), AST to ALT ratio (AAR), eLIFT, HEPAMET score, pro-C3, FibroMeter™, FibroTest® and ELF™. The most validated are NFS and FIB-4, which are non-patented tests. NFS is based on the combination of six variables (age, BMI, AST/ALT ratio, platelet count, hyperglycemia and albumin) whereas FIB-4 is based only on the combination of age, AST, ALT and platelet

count. These scores use two cut-offs to rule-out or rule-in advanced fibrosis: one with high sensitivity (1.3 for FIB-4, and -1.455 for NFS) and another with high specificity (3.25 for FIB-4 and 0.676 for NFS). Advantages of NFS and FIB-4 are the following: 1) they are both based on simple variables widely available in clinical practice; 2) their results can be easily obtained at bedside on free online calculators; 3) their overall diagnostic accuracy for advanced fibrosis, as reported by a recent meta-analysis (n=36 studies in 9074 patients), is good with AUROCs of 0.80 for FIB-4 and 0.78 for NFS ([132]; 4) both can exclude the presence of advanced fibrosis with high negative predictive value (>90%) [132]. Disadvantages of NFS and FIB-4 are: 1) their positive predictive value for confirming advanced fibrosis is modest (<70%) with the risk of false positive results [132]; 2) about one third of patients fall in-between the upper and lower cut-off values giving an undetermined result [132]; 3) older age has been suggested to affect their diagnostic accuracy [133]. Therefore higher cut-offs have been proposed for ruling out advanced fibrosis in patients older than 65 years (2.0 for FIB-4, and 0.12 for NFS) but they need to be externally validated [133]; 4) preliminary evidence suggests lower performance of NFS in obese patients [134, 135] and in diabetic patients [136, 137], where FIB-4 could be preferred [136, 137]. The two most validated patented serum fibrosis biomarkers are FibroMeter™ and ELF™. ELF™ has been evaluated in an independent meta-analysis (n=11 studies in 4,452 patients) with an AUROC of 0.83 for detecting advanced fibrosis [138]. Overall, diagnostic accuracy of patented serum fibrosis tests for staging fibrosis is at least similar [139], if not higher [140], than that of FIB-4 and NFS, but their widespread application in clinical practice is limited by cost and availability.

Transient elastography is the most widely available device for LSM with the largest amount of data in the NAFLD setting. A large recent meta-analysis (M probe 17 studies; 2642 patients; XL probe 3 studies 318 patients) reported a good diagnostic accuracy for advanced fibrosis (AUC 0.87 with M probe and 0.86 with XL probe) and cirrhosis (AUC 0.92 with M probe and 0.94 with XL probe) [132]. The use of both M and XL probes reduces the failure rate to less than 5% of cases [118, 120]. A recent study suggests using the same LSM cut-offs for M probe in non-obese and XL probe in obese patients [141]. Transient elastography has a high NPV (above 90%) to rule-out advanced fibrosis but a modest PPV in NAFLD as compared to viral hepatitis; LSM has more often false positive results in NAFLD [118, 120]. Contrasting results exist about the impact of ALT levels, BMI, skin-to-capsule distance and steatosis/CAP on LSM accuracy and risk of false positive results [118, 134, 142-144]. There is no agreement in clinical practice on LSM cut-offs for

ruling-out advanced fibrosis, even though 8 kPa is the most validated one with NPV above 90% [113]. According to the results of a recent meta-analysis [55] values of LSM by TE > 12-15 kPa could be used to rule-in advanced fibrosis.

Regarding pSWE and 2D shear wave elastography (2D SWE), two recent meta-analyses [145, 146] suggest performance for detecting advanced fibrosis in keeping with those reported for FibroScan® [147]. However, they are less available in liver clinics and data in NAFLD patients remain limited.

Finally, magnetic resonance elastography (MRE) can be considered the most accurate non-invasive method for detecting advanced fibrosis. In a recent individual patient data meta-analysis, based on 3 studies in 230 patients, comparing MRE to TE [148], MRE outperformed TE for detecting advanced fibrosis (AUC 0.94 vs. 0.83, respectively, $p=0.001$) [148]. However, the amount of data in NAFLD remains limited. In addition, given its cost and limited availability, MRE cannot be recommended in clinical practice and is more suited for clinical trials.

Limitations of serum scores and TE together with the need to extend the search for NAFLD patients with fibrosis outside tertiary referral centers inspired clinical studies assessing whether combination strategies are better than the use of each method alone. A sequential combination of NFS or FIB-4 as first test -keeping in follow-up patients at low risk- followed by the use of TE in patients in the medium/high risk area was better than each test alone by obtaining a diagnostic accuracy ranging from 75% to 80% and by lowering the uncertainty area to <10% [134, 149]. Similar results have been reported when combining eLIFT score with FibroMeter™ [150] or FIB-4 with ELF™ score [151].

What is the accuracy of non-invasive scores, serum markers, liver stiffness, and imaging methods to predict liver-related outcomes in patients with NAFLD compared to liver biopsy, HVPG, Child Pugh score or MELD?

Recommendations

- **Serum scores (APRI, FIB-4, NFS, ELF™) and LSM by TE should be used to stratify the risk of liver-related outcomes in NAFLD (LoE 3; strong recommendation).**

- **Repeated measurements of NITs can be used to refine stratification of risk of liver-related events in NAFLD/NASH patients. Despite the lack of evidence regarding the optimal timeframe between subsequent LSM assessment, it seems reasonable to repeat NITs every 3 years in patients with early stage and every year in patients with advanced stage (LoE 3; weak recommendation)**

Available evidence suggests that non-invasive serum markers and elastography devices developed for predicting the presence of liver fibrosis can also have a role in predicting long-term prognosis of NAFLD patients.

A recent retrospective longitudinal study evaluated the ability of non-invasive scores to detect fibrosis progression in 292 NAFLD patients with paired liver biopsies (median time interval of 2.6 years) [152]. Changes over time of APRI, FIB-4 and NFS were significantly associated with fibrosis progression (defined as 1 fibrosis stage) (cross-validated C-statistic for detecting progression to advanced fibrosis of 0.82 for APRI, 0.81 for FIB-4 and 0.80 for NFS). FIB-4 and NFS had high negative predictive values (around 90%), but suboptimal positive predictive values for predicting progression to advanced fibrosis [152]. Furthermore, data from the simtuzumab trials showed that an ELF™ value ≥ 9.76 (sensitivity 77%, specificity 66%) can predict progression to cirrhosis in patients with F3 fibrosis [153].

In a retrospective cohort study of 320 patients with biopsy-proven NAFLD, NFS and FIB-4 accurately predicted the occurrence of liver events (AUROC 0.86 and 0.81, respectively), while having a lower accuracy for overall mortality (AUROC 0.70 and 0.67, respectively) [154]. The authors reported a progressive impairment in clinical outcomes from patients at low to those at intermediate and further to those at high risk for advanced fibrosis, but they did not compare the accuracy of non-invasive tests with histology. Similarly, an APRI value >1.5 significantly predicted the occurrence of HCC in an Asian cohort (N=6,508, median follow-up 5.6 years) of patients with ultrasonographic diagnosis of NAFLD [155]. Three other recent retrospective studies in patients with biopsy-proven NAFLD confirmed the good accuracy of both tests in predicting liver-related events and overall mortality [156-158]. One of these studies also showed that the severity of liver disease by histology was superior to NITs in predicting severe liver disease, but not in predicting overall mortality [157], while another reported similar diagnostic accuracy for predicting liver-related events and overall mortality considered together [156].

The good ability of FIB-4 in predicting not only liver-related events and overall mortality, but also liver-related mortality has been reported in a French study in 360 patients with biopsy-proven NAFLD and a median follow-up of 6.4 years [159]. A large US cohort study in 11,154 individuals (NHANES cohort) of whom 34% had NAFLD on ultrasound reported that those diagnosed as having advanced fibrosis using NFS had higher overall, liver-related and cardiovascular mortality [160]. Finally, in 250 compensated cirrhotic patients enrolled in the simtuzumab trials (median follow-up 30.9 months), ELF™, at a cut-off of 11.27, could predict (C-statistic 0.68, sensitivity 51%, specificity 72%) the onset of clinical events, similar to that observed for liver collagen content [153].

As for LSM using TE, the above mentioned study from France, LSM and FibroMeter™ had good accuracy for liver-related events as well as for liver-related and overall mortality [159]. Similar results regarding the accuracy of LSM and FIB-4 for liver-related mortality were reported in another French study; the authors also observed a similar accuracy for FibroTest® [161]. In a large population of 2251 NAFLD patients (diagnosed by ultrasound and with a short follow-up in median of 27 months) reported good performance of LSM in predicting overall mortality and liver complications (higher rate of events in patients with LSM >12 KPa) but not for the prediction of cardiovascular events and extra-hepatic cancers [162]. Consistently, baseline LSM independently predicted hepatic decompensation, HCC and liver-related death in 1,039 patients with NAFLD-related cACLD [163]. **PNPLA3 I148M variants are associated with higher risk of developing cirrhosis and HCC, but genetic testing is not used yet in clinical practice.**

Two recent retrospective studies investigated the impact of dynamic changes in FIB-4 and LSM on long-term outcomes. A population-based Sweden study on 40,729 individuals with availability of FIB-4 at 2 time points (baseline and within 5 years; mean time 2.4 years) showed that progression from a low- or intermediate- to a high-risk group was associated with an increased risk of severe liver disease (aHR 7.99 and 8.64, respectively) [24]. Similarly, a retrospective analysis of 533 patients with NAFLD-related cACLD and availability of LSM at baseline and within 1 year from the last follow-up (median time 37 months) showed that changes in LSM were independently associated with hepatic decompensation, HCC, overall mortality, and liver-related mortality (HR 1.96) [163]. Further prospective studies are needed to assess the impact of dynamic changes in non-invasive scores and LSM on long-term outcomes. Even if there is lack of evidence and the optimal timeframe remains to be found, it seems reasonable to repeat NITs every 3 years

in patients with early stage disease and every year in patients with advanced stage liver disease.

What is the accuracy of non-invasive scores, serum markers, liver stiffness, and imaging methods for patients' selection and evaluation of treatment response in NAFLD therapeutic trials, compared to liver biopsy?

Recommendations

- **Liver biopsy remains the reference for patient selection in phase 2b and phase 3 therapeutic trials and should be used for these purposes (LoE 1; strong recommendation).**
- **MRI-PDFF can be used to assess steatosis evolution under treatment (LoE 2; weak recommendation). However, the minimal decrease in MRI-PDFF that defines a clinically relevant change or treatment response needs to be better defined.**
- **Liver biopsy remains the reference to evaluate NASH resolution and liver fibrosis improvement and should be used for these purposes (LoE 2; strong recommendation).**

New drugs for NASH need to follow a highly standardized process before getting approval for use in clinical practice [164]. After phase 1, phase 2a trials demonstrate “on target effects” and provide pharmacokinetic and safety data. Then, phase 2b trials evaluate histological improvement in a significant subset of patients. Finally, phase 3 trials robustly confirm the histological improvement but also demonstrate the benefit regarding long-term clinical outcomes in large samples of patients. Study endpoints rely on non-invasive tests in phase 2a trials, whereas liver biopsy is used for phase 2b and 3 trials [165]. With the aim of selecting a subpopulation enriched in potential candidates, non-invasive tests are of interest to facilitate inclusions and reduce unnecessary screening liver biopsies in phase 2b and 3 therapeutic trials. Additionally, a non-invasive evaluation of treatment response instead of paired liver biopsies would increase the feasibility and the patient retention in clinical trials. Ultimately, beyond therapeutic trials, non-invasive tests validated for the identification of patients who need to be treated and for treatment response evaluation will

facilitate the practical management of patients once the new drugs will be available on the market.

Patient selection for therapeutic trials

According to international guidelines, pharmacological therapy should be reserved to NAFLD patients having an active disease and significant amount of liver fibrosis [30, 109]. It has been recently shown that NASH patients with a NAFLD activity score (NAS) ≥ 4 had a less-pronounced placebo response rate than those with a lower NAS [166]. Therefore, most of the phase 2b and phase 3 trials include patients with “fibrotic NASH” (NASH + NAS ≥ 4 + fibrosis stage F2-3). There is currently no validated test for the non-invasive diagnosis of NASH. The non-invasive tests able to accurately diagnose advanced F3/4 fibrosis are less accurate to identify earlier fibrosis stages and F2 patients [167]. Therefore, three tests have been recently developed specifically for the non-invasive diagnosis of fibrotic NASH: two blood tests, MACK-3 and NIS4, and the transient elastography-based FAST score [168-170]. The MACK-3 includes 4 serum markers (aspartate aminotransferase, glucose, insulin and CK18) as the NIS4 (miR-34a-5p, alpha2-macroglobulin, YKL-40, HbA1c), while FAST combines, according to a non-patented formula, aspartate aminotransferase with LSM and CAP values. The studies carried out by the developers showed good accuracy with AUROCs for detecting fibrotic NASH between 0.80 and 0.85 [168-170]. These tests require further external and independent validation in large cohorts.

Evaluation of treatment response in therapeutic trials

Weight loss is associated with decrease in liver steatosis, and new drugs have been developed for an antisteatotic effect. In these contexts, a precise evaluation of steatosis evolution is of interest to evaluate the effectiveness of the intervention.

Very recent studies have suggested that ProC3, a blood marker that directly reflects collagen formation during fibrogenesis, may be useful to identify responders to pharmacological treatment in NASH [171], but this requires confirmation in larger series. Cross sectional studies have demonstrated that MRI-PDFF provides a non-invasive, accurate, precise, sensitive, and reproducible quantification of liver steatosis [114]. The ability of MRI-PDFF to track change in liver steatosis has been evaluated as secondary endpoint in clinical trials with paired liver biopsies [172-175]. These preliminary studies have shown that changes in MRI-PDFF values well correlate with changes in steatosis on

liver biopsy. In addition, it has been suggested that MRI-PDFF could be more sensitive than liver biopsy to detect small changes in liver steatosis [176]. Therefore, MRI-PDFF appears as a promising tool to monitor steatosis evolution and is used as reference in phase 2a clinical trials evaluating drugs with antisteatotic mechanism of action. It should be acknowledged however that currently available evidence comes from small series of patients, which used the rough histological grades to monitor steatosis evolution. Larger studies, using as reference, precise and sensitive tools able to track subtle changes of steatosis on liver biopsy such as morphometry, are therefore required to definitely validate MRI-PDFF as the reference for the non-invasive evaluation of steatosis evolution under treatment. Additionally, the minimum MRI-PDFF decrease corresponding to a clinically relevant change or to treatment response needs to be better defined. New methods of ultrasonography and elastography are in development for the quantification of liver steatosis, but there is currently no data about their ability to monitor steatosis evolution under treatment.

The FDA (US Food and Drug Administration) and the EMA (European Medicines Agency) recognize two endpoints for the conditional approval of drugs in pre-cirrhotic patients: 1) resolution of NASH without worsening of liver fibrosis, and 2) at least one stage improvement in liver fibrosis without worsening of NASH [165]. There is currently no validated biomarker for liver inflammation and therefore no strong candidate for the non-invasive evaluation of NASH resolution. Ideally, the biomarker used to evaluate treatment response should be independent of the drug mechanism of action. In 200 adults with NASH, a ≥ 17 IU/L decrease of alanine aminotransferase at week 24 was the strongest predictor (odd ratio >10) of histological response as defined by a ≥ 2 -point improvement in NAS without worsening of fibrosis [177]. In a recent meta-analysis (n=7 studies; 346 patients) [178], MRI-PDFF responders (defined as relative decline in liver fat $\geq 30\%$) were more likely to have NASH resolution (41% vs 7%, $p < 0.001$; OR 5.45, 95% CI 1.53-19.46, $p = 0.009$) compared to MRI-PDFF non-responders. Such association between histological response and steatosis decrease was however not reproduced in another large study (n=121 patients) [179]. Thus, further studies are needed before any firm conclusions can be drawn. Moreover, as MRI-PDFF response has been evaluated in a short timeframe of months, it is unclear if the response is sustained in the long term and if it also translates to improvement in fibrosis.

Several non-invasive tests (serum markers and elastography) are accurate for the diagnosis of advanced fibrosis in NAFLD. Therapeutic trials represent a unique opportunity to evaluate their ability to monitor fibrosis evolution under treatment. Change over time of the blood test ELF™ was independently associated with an increased risk of disease progression in 217 NAFLD patients with advanced fibrosis from a phase 2b trial [153]. However, in another work including 54 F2-3 patients, the median relative change in liver stiffness by MRE was not significantly different between patients with fibrosis improvement (≥ 1 stage) and those without fibrosis improvement (-2.3% vs 3.0%) [174].

Studies about the non-invasive evaluation of treatment response remain scarce in the literature, and the few preliminary data available need confirmation in larger samples of patients. Consequently, liver biopsy currently remains the reference to evaluate NASH resolution and liver fibrosis improvement under therapy. The most relevant existing biomarkers and panels should be tested for this purpose, and extensive research should be conducted to find new candidate biomarkers. The many ongoing therapeutic trials in NASH include as secondary endpoint the evaluation of non-invasive tests, which will allow the accumulation of evidence about their ability to monitor treatment response.

Cholestatic and autoimmune liver disease (PBC, PSC, AIH)

What is the accuracy of non-invasive scores, serum markers, liver stiffness, and imaging methods to assess disease severity in comparison to liver biopsy in patients with primary biliary cholangitis and primary sclerosing cholangitis?

Recommendations

- **In patients with PBC, serum markers of fibrosis and non-invasive scores (combination of clinical and laboratory variables) are not recommended for fibrosis staging in clinical practice (LoE 3; strong recommendation).**
- **In patients with PBC, LSM by TE is the best surrogate marker for ruling-in severe fibrosis/cACLD and should be used for this purpose using a cut-off of 10 kPa (LoE 3; strong recommendation).**

- **In patients with PSC, LSM by TE above 9.5 kPa can be used to support the diagnosis of advanced fibrosis in compensated patients with normal bilirubin and without high-grade stenosis (LoE 3; weak recommendation).**

In general, studies on NITs in patients with PBC or PSC involve a small or very small number of patients.

In primary biliary cholangitis (PBC) similarly to other chronic liver diseases, advanced histological stages are associated with poor prognosis [180-185]; fibrosis stage was recently demonstrated to be an independent predictor of outcome even in patients with biochemical treatment response [186]. However, liver biopsy is no longer indicated in the diagnostic work up of PBC, unless in specific situations (absence of PBC specific antibodies, suspicion of coexistence of AIH or NASH or other co-morbidities) or in case of inadequate response to UDCA therapy in order to characterize histological lesions that underlie the resistance to treatment [187]. Moreover, the course of the disease may be progressive, despite UDCA treatment, thus non-invasive assessment of fibrosis is crucial both at diagnosis and during follow-up of these patients.

Serum biomarkers of liver fibrosis including serum levels of hyaluronic acid, procollagen III aminoterminal propeptide, collagen IV and FibroTest® do not have adequate accuracy to differentiate between early and advanced fibrosis in PBC [5]. Similarly, non-invasive scores, namely APRI, FIB-4, Aspartate transaminase to Alanine Transaminase Ratio (AAR), red blood cell distribution width to platelet ratio (RPR), red blood cell distribution width to lymphocyte ratio (RLR) and neutrophil to lymphocyte ratio, have a suboptimal diagnostic performance (AUROC < 0.80) in predicting histological stage in PBC [186, 188-196]. In one study, platelet count to spleen diameter (PC/SD) ratio showed a good diagnostic performance in predicting advanced fibrosis stage [197]. LSM by TE showed previously to correlate with liver fibrosis in PBC [184, 198, 199] and based on prospective data [198] a cut-off of 9 kPa was proposed to identify patients with vs. without significant fibrosis and 10.7 kPa for advanced fibrosis [5]. A study including 44 PBC patients confirmed a good accuracy of LSM by TE in predicting advanced fibrosis and cirrhosis (AUROCs 0.91 and 0.97, respectively), but reported higher optimal cut-offs for identification of advanced fibrosis and cirrhosis [189]. We suggest that an optimal cut off of 10 kPa should be used for rule in advanced fibrosis. One study including 41 PBC patients assessed the diagnostic performance of point shear wave elastography (pSWE) and reported promising results in prediction of both significant and advanced fibrosis in this

disease (AUROCs 0.81 and 0.91, respectively) [188]. Finally, preliminary data on the use of MRE in PBC were reported but they need to be further validated [200].

In primary sclerosing cholangitis (PSC), two studies published since the publication of the EASL-ALEH 2015 guidelines and including 62 and 39 PSC patients confirmed the good accuracy of LSM by TE in predicting advanced fibrosis (AUROC 0.95, sensitivity 90%, specificity 91%) and cirrhosis (AUROCs 0.98 and 0.90, sensitivity 69% and 78%, specificity 98% and 90%, respectively) with similar optimal cut-offs for predicting cirrhosis (14.4 kPa, and 13.7 kPa, respectively) [201, 202]. Moreover, in the Simtuzumab trial, the diagnostic performances of the optimal cut-offs for advanced fibrosis (≥ 9.6 kPa) and cirrhosis (≥ 14.4 kPa), reported by Corpechot *et al.*, were confirmed to have a good and excellent accuracy (AUROCs 0.80 and 0.95, sensitivity 74% and 100%, specificity 74% and 83%, respectively) [203]. Liver stiffness by MRE was assessed in 20 biopsy-proven PSC patients and the reported diagnostic accuracies for predicting fibrosis stage $\geq F1$, $\geq F2$, F4 were excellent (AUROCs 0.97, 0.97 and 0.99, respectively), however these data need to be confirmed in larger independent cohorts [204].

In patients with increased serum bilirubin due to the presence of a high-grade stenosis in the extrahepatic bile ducts, LS values need to be carefully interpreted due to the relevant risk of over-estimation of the fibrosis stage [204-206].

Preliminary data on spleen length measurement by ultrasound suggested a good diagnostic performance to identify cirrhosis (AUROC 0.85, sensitivity 73%, specificity 73%) when an optimal cut-off of 120 mm was applied [207].

What is the accuracy of non-invasive scores, serum markers, liver stiffness, and imaging methods to predict liver-related outcomes in patients with primary biliary cholangitis and primary sclerosing cholangitis compared to liver biopsy, HVPG, Child Pugh score or MELD?

Recommendations

Primary biliary cholangitis

- **In patients with PBC, non-invasive discrimination of early and advanced stage based on biochemical parameters (normal vs.**

abnormal albumin and bilirubin) and LSM by TE < or > 10 kPa is recommended at baseline (LoE 3, strong recommendation).

- **During treatment, risk stratification should be based on the assessment of response to therapy by using continuous (GLOBE and UK-PBC risk scores) and/or qualitative criteria (Paris II, Toronto, Rotterdam, Barcelona, Paris I) of response and LSM by TE (LoE 3, strong recommendation).**

Primary sclerosing cholangitis

- **In patients with PSC, both the ELF™ score and LSM by TE correlate with outcomes and they should be used for risk stratification both at baseline and during follow-up (LoE 3, strong recommendation).**

Patients with PBC treated with UDCA demonstrate different disease course depending on baseline (pre-treatment) features and biochemical response after 12 months of treatment and risk stratification is required. [187].

At baseline, the distinction of early from advanced disease stage is based on LSM by TE (LSM \leq 10 kPa or LSM $>$ 10 kPa), serum levels of bilirubin and albumin (both parameters normal vs. at least one parameter abnormal) and, when available, histology (absent or mild fibrosis vs. bridging fibrosis or cirrhosis) [187].

On-treatment, the evaluation of prognosis is based on the assessment of biochemical response to UDCA by using qualitative criteria (Paris-I, Paris-II, Rotterdam, Toronto, Rochester, Ehime criteria) or by the recently proposed quantitative criteria (UK-PBC score and GLOBE score). The GLOBE score (which includes age, total bilirubin, alkaline phosphatase, albumin and platelet count), derived and validated in a multi-center international cohort of PBC patients treated with UDCA was shown to accurately predicts LT-free survival at 5 and 10 years (c-statistics 0.81 and 0.82 in derivation and validation cohort) [208]. The UK-PBC score (including baseline albumin and platelet count, and bilirubin, AST or ALT and alkaline phosphatase 12 months after starting UDCA), derived and validated in a multicenter UK cohort of PBC patients treated with UDCA, accurately predict the risk of major outcomes (liver-related death, liver transplantation or bilirubin \geq 100 μ mol/L) at 5, 10, 15 years with reported AUROCs of 0.96, 0.95 and 0.94, respectively [209]. Both scores have been externally validated and were superior to qualitative criteria,

and to MELD and Child-Pugh scores [210-212] [213]. Further studies are needed to better define the applicability of the UK-PBC risk score in routine clinical practice.

Biochemical non-response as defined by the GLOBE score and APRI score above > 0.54 after 12 months of UDCA therapy, were recently shown to be independently associated with the risk of cirrhosis decompensation and their use in combination improve risk stratification in these patients [214]. Moreover, a recent study showed that serum level of GGT above 3.2 fold the upper limit of normal (ULN) at 12 months after treatment identifies patients at increased risk of liver transplantation or liver-related death independently of ALP values [215]. Thus, in addition to biochemical response, APRI score and GGT can be used to refine risk stratification in these patients. Finally, ALP normalization or serum bilirubin below $0.6 \times$ ULN after 12 months of treatment were recently associated with the lowest risk for LT or death in patients with PBC [216]. The ELF score has also been associated with clinical outcomes in PBC [217].

In addition, LSM by TE is indicated on treatment in the follow-up, since worsening of LSM predicts patients' outcome [5, 187, 198]. An increase of 2.1 kPa/year in LSM by TE was associated with a 8.4 times increased risk of adverse outcomes [198]. Despite the lack of evidence regarding the optimal timeframe between subsequent LS assessment, it seems reasonable to repeat LSM every 2 years in patients with early stage and every year in patients with advanced stage.

Primary sclerosing cholangitis is generally progressive and the natural history [218-220] is characterized by spontaneous fluctuation in bilirubin due to the occurrence of acute bacterial cholangitis, biliary stones or high-grade strictures. This explains the difficulty to accurately predict prognosis by applying classical prognostic models (Child-Pugh score and MELD score). Histological stage assessed by liver biopsy is strongly associated with clinical outcomes [221] and is still considered a robust surrogate endpoint for clinical trials in PSC [222].

The ELFTM score demonstrated a good accuracy in predicting liver transplant-free survival in several large independent cohorts of PSC patients with reported AUROCs ranging between 0.78 and 0.81 and optimal prognostic thresholds around 10 [223-227]. Recently, the prognostic values of the serological markers of extracellular matrix remodeling, PRO-C3 and PRO-C5, showed comparable accuracy to ELFTM in predicting LT-free survival (AUC 0.78, 0.74 vs. 0.81), moreover, PRO-C5 was able to predict LT-free survival independently from ELFTM score [226].

Four new composite scores including clinical, biochemical and radiological features were derived by using 3 large multi-center cohorts of PSC patients. The Amsterdam-Oxford model (AOM, including PSC subtype, age at PSC diagnosis, albumin, platelets, aspartate aminotransferase, alkaline phosphatase and bilirubin) showed moderate accuracy in prediction of LT and PSC-related death (c-statistic 0.68), however, it resulted well calibrated when applied both at diagnosis and during follow-up [228]. The AOM was then validated in an independent multi-center cohort showing increased accuracy that remains stable during follow-up (c-statistics at baseline, 1, 2, 3, 4 and at 5 years of follow-up: 0.67, 0.69, 0.72, 0.75, 0.75 and 0.75, respectively) [229]. The Primary Sclerosing Cholangitis Risk Estimate Tool (PREsTO, including bilirubin, albumin, serum alkaline phosphatase (SAP) times the upper limit of normal, platelet count, aspartate aminotransferase, hemoglobin, sodium, patient age, and number of years since PSC diagnosis), derived with a machine learning technique, demonstrated a good accuracy (c-statistic 0.90) to predict hepatic decompensation and excellent to predict LT and PSC-related death, exceeding that of MELD and Mayo risk score (c-statistics 0.96 vs. 0.73 and 0.84) [230]. Lastly, the Short-Term (RS_{ST}) and the Long-Term (RS_{LT}) UK-PSC risk score (including PSC type, age at diagnosis, haemoglobin at diagnosis, total bilirubin, albumin, platelet count, serum alkaline phosphatase at baseline and at year 2, and occurrence of variceal bleeding at year 2) showed good accuracy in predicting LT-free survival (c-statistics of both score \geq 0.80) and the RS_{ST} outperformed the Mayo risk score, APRI and MELD [231]. Further data are needed to understand the practical application of these scores in the clinical setting.

Baseline LSM by TE and the increase of LSM over time were associated with prognosis [232] and thus recommended in the previous EASL-ALEH 2015 guidelines for prognostic purpose in PSC. Subsequent studies confirmed the association of LSM values with liver-related outcomes in PSC patients [201, 203] and histological stage [203]. Optimal thresholds of LSM for the prediction of prognosis differed between studies, depending on outcomes considered. A large multicenter prospective study, by the International PSC Study group, to assess the prognostic value of LSM by TE (FICUS study) is now ongoing and an interim analysis confirmed the high predictive performance of LSM by TE (AUROC 0.88) with reported adjusted HRs of adverse outcomes of 4.2 for baseline LS values between 9.6 and 14.3 kPa and of 16.3 for baseline LSM values above 14.3 kPa, both compared to baseline LSM <9.6 kPa. Despite the lack of evidence regarding the optimal timeframe, it seems reasonable to repeat LSM by TE and/or ELF™ annually.

LSM by MRE was also associated with the risk of cirrhosis decompensation [204].

Spleen length at baseline and its changes during the follow-up were also associated with LT-free survival in PSC patients [207, 233] and the change of spleen volume seemed to predict liver-related outcomes better than Mayo risk score and MELD [234].

Finally, cholangiographic changes assessed by ERCP [235] and more recently, by MRI [236] were used for prognostic purpose. Two risk scores taking into account imaging features on magnetic resonance (without or with Gadolinium injection), called ANALI score, were shown to be independently associated with survival without adverse outcomes with reported c-statistics of 0.89 for the ANALI without gadolinium and 0.76 for the ANALI with gadolinium [237]. Moreover, the use in combination of ANALI score without gadolinium and LSM by TE is able to better stratify patients according to the risk of development of major outcomes [238]. A study comparing cholangiographic findings obtained by ERCP and MRI reported a weak correlation between cholangiographic findings and major outcomes [239]. Finally, the relative enhancement of liver parenchyma (RLE) after hepatospecific contrast agent (Primovist®) injection was correlated with markers of disease severity (ALP, INR), prognostic risk score and clinical outcomes [240]. In conclusion, in patients with PSC, recent evidences support the use of MRI, alone or in combination with TE, for risk stratification, similarly a number of prognostic scores were proposed and this data needs to be further confirmed.

What is the accuracy of non-invasive scores, serum markers, liver stiffness, and imaging methods to assess liver fibrosis, and to monitor disease course as compared to liver biopsy in patients with autoimmune hepatitis?

Recommendation

- **LSM by TE can be used in patients with treated AIH to monitor the disease course together with transaminases and IgG, and to stage liver fibrosis after at least 6 months of immunosuppressive therapy (LoE 3, weak recommendation).**

Several non-invasive methods used in viral and non-viral chronic liver disease to assess histological stage have been tested in autoimmune hepatitis including non-invasive scores (APRI, FIB-4, AST to ALT Ratio-AAR, NAFLD fibrosis score), LSM by TE, pSWE and 2D-SWE and imaging methods.

Non-invasive scores such as APRI, FIB-4 and AAR have a poor diagnostic accuracy in predicting liver fibrosis, especially in early fibrosis stage [241-245]. Indeed, the summary AUROCs of FIB-4, APRI and AAR for advanced fibrosis ($F \geq 3$) were 0.76, 0.74 and 0.73, respectively. Similarly, the summary AUROCs of FIB-4 and APRI for cirrhosis were 0.66 and 0.75, respectively [246]. One study including 53 patients with AIH, suggested that the NAFLD fibrosis score has an adequate accuracy to predict cirrhosis (AUROC 0.91, sensitivity 0.90 and specificity 0.89) [241]. However, this data needs to be further confirmed.

LSM by TE is positively correlated with histological fibrosis stage in AIH and is able to detect advanced fibrosis and cirrhosis with similar accuracy than in other chronic liver disease. However, hepatic inflammation is a known confounding factor that can lead to overestimation of liver stiffness, independently from fibrosis stage [5, 247].

Monitoring fibrosis progression during immunosuppressive therapy is crucial, especially in patients with insufficient response, intolerance or non-adherence. A study collectively including 94 patients with biopsy proven AIH showed that LSM assessed within the first 3 months from starting immunosuppressive treatment is more strongly correlated with histological disease activity and to a lesser degree with histological fibrosis stage. In particular, within the first 3 months ($n=34$ patients) the diagnostic performance of LSM for predicting advanced fibrosis showed an optimal cut-off of 10.4 kPa with reported AUROC, sensitivity and specificity of 0.80, 60% and 88%, respectively. Within 6-12 months from treatment initiation ($n=25$ patients) the same cut off predict advanced fibrosis with reported AUROC, sensitivity and specificity of 1.00, 100% and 100% and finally, after 4 years, the reported AUROC, sensitivity and specificity were 0.96, 95% and 94%, respectively [248]. In other three studies, collectively including 261 patients with AIH, LSM values for predicting advanced fibrosis vary within 8.2 and 12.1 kPa depending on the percentage of treatment-naïve patients included, with reported AUROCs, sensitivity and specificity of 0.74-0.90, 59%-80% and 83%-85%, respectively [241-243].

pSWE to detect histological fibrosis stage in 49 patients with AIH showed a moderate diagnostic accuracy to detect significant fibrosis, advanced fibrosis and cirrhosis (AUROCs 0.70, 0.76 and 0.75, respectively) [188]. In one study, 2D-SWE showed promising results

in predicting histological fibrosis stage in 103 patients affected by autoimmune liver diseases including 62 patients with AIH, 30 patients with PBC, 3 patients with PSC and 19 patients with PBC-AIH variant, but unfortunately data on diagnostic performance of pSWE for each single disease was not provided [249]. Finally, liver stiffness measured by MRE showed a good diagnostic performance in predicting advanced fibrosis and cirrhosis in 36 patients with AIH [250].

Platelet count to spleen diameter (PC/SD) ratio, assessed in 76 biopsy-proven AIH patients, showed a good diagnostic performance for predicting significant fibrosis, advanced fibrosis and cirrhosis (AUROC 0.84, 0.88, 0.97, respectively)[244].

To monitor disease course, complete biochemical remission, defined as normalization of transaminases and immunoglobulin G, was able to predict low histological activity and was the only independent predictor of histological fibrosis regression over time. Decrease of LSM during disease course was strongly linked to complete biochemical remission in one study [251].

Compensated advanced chronic liver disease (cACLD) and portal hypertension

What is the accuracy of non-invasive scores, serum markers, liver stiffness, and imaging methods to diagnose cACLD as compared to liver biopsy?

Recommendations

- **cACLD should be diagnosed by using second line tests (patented serum tests or elastography) in a specialized setting (LoE 2, strong recommendation)**
- **Fibrotest® or FibroMeter™ or ELF™ should be used to rule out cACLD if available (LoE 3, strong recommendation)**
- **LSM by TE should be used to rule-out and diagnose cACLD using the following cut-offs: < 8-10 kPa to rule-out; > 12-15 kPa to rule-in. Intermediate values require further testing (LoE 3 strong recommendation)**
- **pSWE and 2D-SWE should be used to rule-out and diagnose cACLD, with AUROCs >0.90 in the published meta-analyses (LoE 2, strong recommendation)**

- **Inter-system variability should be taken into account when interpreting the results of different elastography techniques, since values, ranges and cut-offs are not comparable (LoE 3, strong recommendation)**

The discrimination between severe fibrosis and compensated cirrhosis is often unclear since fibrosis can be inhomogeneously distributed within the liver, particularly in some etiologies (6), and since it is a dynamic process which can progress but also regress. Due to these considerations, and in order to better discriminate between patients at risk of developing portal hypertension and clinical decompensation, and patients in an earlier stage of chronic liver disease, it has been suggested to rename this clinical scenario including severe fibrosis and compensated cirrhosis as “compensated advanced chronic liver disease” (cACLD) (7).

Given its important prognostic implications, cACLD should be diagnosed by using second line tests (patented serum tests Fibrotest®, FibroMeter™ and ELF™ or elastography) in a specialized setting. The performance of serum markers and liver stiffness to diagnose significant fibrosis, severe fibrosis and cirrhosis in compensated patients has been extensively reviewed in the previous EASL guidelines [5].

Elastography updates are available in other recent guidelines from EFSUMB [247] and WFUMB [252]. Except for the novel data provided in the other specific sections of these guidelines, data on TE do not modify the previous recommendations and this method remains the best validated. Since 2015 there have been numerous publications and meta-analyses regarding the accuracy of pSWE and 2D SWE for liver fibrosis staging in comparison to liver biopsy. In addition to the data already available for HCV and HBV suggesting accuracies similar to TE, a meta-analysis on the performance of pSWE in 29 studies in patients with chronic liver disease due to non-viral etiologies [253] showed an AUROC of 0.94 for advanced fibrosis and cirrhosis.

As for 2D-SWE, on two meta-analysis, one including all etiologies [145] and one in NAFLD [132] it showed an accuracy similar to that of TE as for advanced fibrosis detection. As for the diagnosis of cirrhosis, in the meta-analysis by Hermann et al. [145] the AUROC of 2D-SWE was 0.92-0.95 (varying slightly among etiologies), and was 0.003-0.034 ($p=0.022$) larger than the AUROC of transient elastography. This difference was strongest in hepatitis B patients.

Intersystem variability should be taken into account, but as for cirrhosis, one study comparing 6 different systems showed a good to excellent agreement between measurements performed with different systems, with an interobserver agreement over 0.90 [254]. Nonetheless, knowledge of the specific cut-offs for each system has to be applied since they do not completely overlap.

In summary, LSM by TE remains the most validated tool to diagnose and rule-out advanced fibrosis and cirrhosis in all the major etiologies of chronic liver disease, holding a discriminating ability of over 0.90. Published cut-offs to diagnose cirrhosis vary from 11 to 27 kPa according to the etiology; however, cut-offs should be considered with caution due to the considerations regarding the prevalence of the fibrosis stage to be diagnosed in the target population. Rule-out and rule-in cut-offs can be used to minimize the risk of under- or overestimation. Furthermore, since it has been well demonstrated that the higher the liver stiffness, the higher the risk of advanced fibrosis and cirrhosis, approaches based on individualization of risk based on nomograms can be useful in this setting[50, 255].

The definition of cACLD provided by the Baveno VI recommendations encompasses advanced fibrosis and compensated cirrhosis and is based on LSM by TE alone (2 measurements on different days showing ≥ 10 kPa suggestive of cACLD; ≥ 15 kPa highly suggestive of cACLD) and is aimed at providing a simple non-invasive tool to help identifying asymptomatic patients at higher risk of developing clinical events in the absence of a confirmatory liver biopsy or in case of not contemporary liver biopsy, taking into account that fibrosis is a dynamic process that might regress from cirrhosis to lesser degree of fibrosis [8]. These criteria have been recently refined, in a validation study that included over 5,500 patients with chronic liver disease. The study showed that a cut-off of >12 KPa has $>90\%$ specificity for diagnosing cACLD, while a cut-off of <8 KPa (for NAFLD and ALD) or <7 KPa (for viral hepatitis) has $>90\%$ sensitivity for ruling-out cACLD [55]. In one study including patients with chronic liver disease of different etiologies, obesity and metabolic syndrome were associated with a high rate of false positive results of the ≥ 10 kPa criteria [256].

Magnetic resonance elastography using 2D gradient recalled echo (GRE) holds a high accuracy for fibrosis staging in all the main etiologies of liver disease [257] and is superior to TE in patients with NAFLD. However, its high cost and suboptimal availability limit its use in clinical practice.

As for conventional imaging methods, ultrasound, CT and MR are useful to identify signs of cirrhosis and portal hypertension (reviewed elsewhere)[258], but their accuracy to

identify cirrhosis in compensated patients does not exceed an AUROC of 0.75-0.80 in the reported studies. The nodularity of liver surface quantified by software analysis on CT scan images has been proposed and holds a high accuracy to detect cirrhosis (sensitivity 86%, specificity 92% using a cut-off of 2.75 in sections obtained in the portal venous phase)[259]. However, its use in asymptomatic patients cannot be routinely recommended due to the risk of radiation exposure. On the other hand, quantification of this parameter in patients undergoing CT for any other cause seems reasonable and could improve the detection of new cases of cACLD/cirrhosis. Several innovative methods, mostly based on MR techniques have been proposed [260] and include diffusion-weighted imaging (DWI), hepatocellular contrast-enhanced (HCE) MRI, T1 relaxometry, T1 ρ imaging, textural analysis, susceptibility-weighted imaging, and perfusion imaging. They are highly promising but need further evaluation and clinical validation and cannot yet be recommended for routine practice. Radiomics approaches are currently being developed to stage liver fibrosis based on US, CT and MR images, but are not ready for clinical implementation yet.

What is the accuracy of non-invasive scores, serum markers, liver stiffness, and imaging methods to diagnose clinically significant portal hypertension and to monitor portal hypertension in cACLD in comparison to HVPG measurement?

Recommendations

- **LSM by TE at a cut-off of >20-25 kPa should be used to diagnose CSPH in patients with cACLD (LoE 1, strong recommendation).**
- **Platelet count, spleen size and spleen stiffness should be used as additional NITs to further improve risk stratification for CSPH (LoE 3, strong recommendation).**
- **The presence of porto-systemic collaterals on ultrasound, CT or MR is a sign of CSPH in patients with cACLD and should be routinely reported (LoE 2, strong recommendation).**

- **For an exact assessment of the severity of portal hypertension in cACLD beyond presence and absence of CSPH and for assessment of the hemodynamic response to treatment, HVPG remains the only validated tool and should not be substituted by NITs (LoE 1, strong recommendation).**

Evidence regarding the use of serum markers of fibrosis to diagnose CSPH is scarce and data suggest an insufficient diagnostic accuracy, so their use is not recommended. In cACLD, a CTP score >5 points is associated with CSPH. Platelet count is inversely related to portal pressure, but its accuracy for CSPH does not exceed an AUROC of 0.75 in the published studies.

Von Willebrand Factor Antigen (vWF-Ag) has been shown to correlate with HVPG in two independent studies [261, 262], and predicted CSPH independent of Child Pugh score. A cut-off value of $\geq 241\%$, showed an AUROC of 0.85 to detect CSPH [262]. However, its use cannot be recommended yet due to lack of further validation.

Among imaging parameters, Liver Surface Nodularity Score (LSNS), a measurement of liver surface nodularity on routine CT, correlates with HVPG ($r=0.75$, $p<0.001$) and predicts CSPH with good accuracy (AUROC 0.88; cut-off 2.8: PPV 88%) [263]. In a pilot study including 30 patients, LSNS was measured on Gd-BOPTA-enhanced MRI and compared to CT, with similar results [264]. Several other parameters such as spleen size and portal vein diameter are associated with portal hypertension but show lower accuracy for the diagnosis of CSPH. On the other hand, the presence of porto-systemic collaterals on ultrasound, CT or MR is a highly specific sign of CSPH in patients with cACLD and is associated with the presence of gastroesophageal varices and with worse prognosis (see below). As such, porto-systemic collaterals should be routinely searched on routine imaging and reported.

Multiparametric magnetic resonance imaging showed promising results to predict CSPH in one small pilot study including 30 patients [265], but has not been validated yet.

LSM by TE (and more recently by pSWE and 2DSWE) is the most validated quantitative individual NIT for portal hypertension in compensated patients. Its linear correlation with HVPG is good but not excellent (AUROC 0.67-0.86). However, using a cut-off of 20-25 kPa LSM is able to identify CSPH with an AUROC of over 0.90; in the meta-analysis by You et al., the summary AUROC was 0.93 with a sensitivity of 87.5% (CI 75.8-93.9%) and

a specificity of 85.3% (95% CI 76.9-90.9%)[266]. As for etiology-specific cut-offs, in the recent meta-analysis including 9 studies and 679 patients [267], the summary sensitivity and specificity for CSPH in ALD patients at a cut-off of 21.8 kPa was 89% and 71%; while for severe portal hypertension at a cut-off of 29.1 kPa, sensitivity and specificity was 88% and 74%. However, 7 of 9 included studies had average HVPG above 12 mmHg. Together with the relatively high sensitivities and low specificities, this indicates spectrum bias, with probable inclusion of many decompensated cirrhosis patients, which limits the clinical value of the analysis.

The accuracy of LSM increases if this is combined with unrelated NITs, in particular platelet count and spleen size (LSPS[268]; PH risk score [269, 270] (see Table 4 for the most used formulas).

Due to the small number of studies with heterogeneous included population, and high variability of the cut-offs [271], pSWE cannot yet be recommended for the routine screening of CSPH in patients with cACLD. 2D-SWE has been tested in 9 studies against HVPG; on meta-analysis the AUROC was 0.88 (95% CI, 0.85-0.91), with a summary sensitivity of 85% and summary specificity of 85% [272]. In the published studies however, there is marked heterogeneity of cut-offs (16-38 kPa), and no recommendation can be given. A recent individual patient data meta-analysis suggested the use of 14 kPa as a cut-off of LSM by 2D-SWE to rule-out CSPH [56].

Spleen stiffness measured by TE, pSWE or 2DSWE has been tested in a limited number of studies vs. HVPG; while it's clear that this parameter correlates with portal pressure, it is unclear whether its performance is similar, inferior or superior to that of liver stiffness for the detection of CSPH. However, it seems reasonable to use spleen stiffness as a complementary NIT for CSPH, e.g. by applying both liver stiffness and spleen stiffness sequentially [273, 274]. The cut-off value of 40 kPa is highly sensitive (98%) to rule-out CSPH, while values above 46-52 kPa are over 90% specific to rule-it in in HCV treatment naïve cACLD [275].

LSM, serum markers and imaging parameters do not reflect changes of HVPG on medical therapy with non-selective beta-blockers (NSBB). Kim et al. [276] recently reported that changes in spleen stiffness measured by pSWE (Virtual Touch, Siemens, Germany) in 106 patients with cirrhosis and high risk esophageal varices before and on Carvedilol for primary prophylaxis, predicted the HVPG changes with good performance (0.80 in the training set

and 0.85 in the validation set). Marasco et al. suggested that SSM by TE could provide data on the hemodynamic response to NSBB as well [277]. Validation in independent cohorts is needed.

For an exact assessment of the severity of portal hypertension in cACLD beyond presence and absence of CSPH, HVPG remains the only validated tool and cannot be substituted by NITs.

What is the accuracy of non-invasive scores, serum markers, stiffness of liver tissue, and imaging methods to diagnose and exclude high-risk gastroesophageal varices in comparison to endoscopy?

Recommendations

- **In patients with cACLD due to untreated viral hepatitis, HIV-HCV coinfection, alcohol, NAFLD, PBC and PSC, the finding of LSM by TE < 20 kPa and platelet count >150 G/L (Baveno VI criteria) is a validated tool to rule-out high risk varices and avoid endoscopic screening. These criteria should be used whenever TE is available (LoE 1a; strong recommendation)**
- **Spleen stiffness can be used as an additional tool to refine the risk of high-risk varices in cACLD (LoE 2; weak recommendation).**
- **CT should not be used for primary screening for esophageal and gastric varices, but when doing a routine CT, varices should be looked for and reported (LoE 3, strong recommendation).**

Several NITs including laboratory tests (platelet count, individual component of the Child-Pugh score, MELD score); imaging signs (portal vein diameter and blood flow velocity, spleen size, nodularity of the liver surface, presence of porto-systemic collaterals), liver stiffness and spleen stiffness correlate with the presence and grade of gastroesophageal varices in patients with cACLD. None of them, taken individually, is sufficient to rule-in or rule-out varices and high-risk varices [278]. However, NITs used in combination achieve better results (e.g. platelet to spleen ratio [279]), and in particular the combination of liver stiffness and platelet count (and even more if spleen size is added, e.g. liver stiffness-spleen diameter to platelet ratio score – LSPS [268]) achieves a marked improvement of the diagnostic ability for varices and varices needing treatment as compared to any of the

individual NITs [269]. In a systematic review of the literature, LSM by TE < 20 kPa combined to a platelet count >150 G/L led invariably to miss less than 5% high-risk varices requiring treatment [17]. This led to an expert recommendation to use these non-invasive criteria (defined “Baveno VI” criteria) to spare endoscopy in patients with cACLD. Since the publication of the criteria, several studies and two meta-analysis [280, 281] confirmed the validity of this approach in all the major etiologies of liver disease including HIV-HCV coinfection [282] and patients who achieved SVR after treatment of HCV [283], showing rates of missed high-risk varices ranging from 0 to 2% and these criteria can be considered validated. Since the Baveno VI criteria are conservative and allow sparing no more than 10-25% of endoscopies, studies looking for expanded criteria have been published. On a recent meta-analysis, the standard criteria have been once more validated, but the number of missed varices appeared too high with the proposed “Expanded criteria”[284]. Etiology-specific cut-offs of liver stiffness [285] and the combination with spleen stiffness [286] might allow further reducing the number of unnecessary endoscopies without increasing the risk of missing high-risk varices above 5%, but further data are required. In addition, individualization of risk/benefit assessment using nomograms derived by well calibrated models is a promising approach for the future [255].

Spleen stiffness measured by TE or pSWE shows similar or even better accuracy (in some studies) vs. liver stiffness to identify patients at high risk or low-risk of high-risk varices. The most commonly reported cut-off using TE is 46 kPa. However, the failure rate using the standard probe of TE is high. A dedicated probe (100 Hz instead of 50 Hz frequency) has been recently commercialized. In the only paper published to date [287], 260 patients were prospectively included in two centers. Success rate for spleen stiffness measurement was significantly higher with the dedicated probe (92.5% vs 76.0% of standard probe, $p < 0.001$) and accuracy to detect high risk varices was superior and outperformed liver stiffness. The use of Baveno criteria alone, vs. combined to standard spleen stiffness vs. new probe spleen stiffness resulted in 8.1% vs. 26.5% vs. 38.9% spared endoscopies. The missed HRV rate was, respectively, 0% with Baveno criteria alone vs. 4.7% for combination with spleen stiffness (any of the probes).

In an open-label randomized controlled trial, a strategy based on liver and spleen stiffness measurement (LSSM) to prompt endoscopic screening was recently demonstrated similar to “endoscopy in all” as for the onset of the index variceal bleeding [288]. A meta-analysis of data from 45 studies including liver and spleen stiffness measurement by different

methods (TE, pSWE and 2DSWE) for high risk varices, showed an AUROC of spleen stiffness of 0.81 (slightly inferior to that of liver stiffness and LSPS), but with a high sensitivity (0.87 vs. 0.85 for liver stiffness) [289].

According to the available data, it seems reasonable to attempt measuring spleen stiffness in patients in whom liver stiffness cannot be measured, or in addition to liver stiffness to further refine risk stratification. Cut-offs should be chosen according to the technique used (pSWE, 2D-SWE or TE).

Liver and spleen stiffness measured by magnetic resonance elastography (MRE) have been tested in a limited number of studies to detect and exclude varices needing treatment. Studies include patients with compensated and decompensated cirrhosis, and despite they confirm that higher liver and spleen stiffness are observed in patients with high-risk varices, data is insufficient to warrant recommendations.

In patients with HCV-related cirrhosis within one year of starting antiviral therapy, the combination of platelet count (>120 G/L) and albumin (>2.6 g/dL) - the RESIST-HCV criteria [290] - might be sufficient to rule-out high-risk varices without measuring liver stiffness. These criteria yielded a negative predictive value of 97-99% in a large training and validation multi-centric cohort leading to spare about 25% of endoscopies and had similar results as the Baveno VI criteria. The combination of MELD score=6 and platelet count >150 G/L could reduce endoscopies by 54% without missing high-risk varices in one study [291]. However, these data have not been validated yet.

Large varices can be diagnosed on multidetector contrast-enhanced CT (MDCT) images with good accuracy. In a meta-analysis of 11 studies [292] the AUROC for the detection of esophageal and gastric varices was respectively 0.86 and 0.91. However, the studies included both compensated and decompensated patients, and no definite conclusion regarding the use in cACLD can be driven. Since CT is often performed in patients with cirrhosis, we consider reasonable to state that varices should be actively searched and reported on MDCT imaging.

What is the accuracy of non-invasive scores, serum markers, liver stiffness, and imaging methods to predict clinical decompensation, hepatocellular carcinoma and mortality in cACLD as compared to liver biopsy, HVPG, Child Pugh score or MELD?

Recommendations

- **In patients with cACLD, liver stiffness at diagnosis should be used in addition to liver function tests to stratify the risk of clinical decompensation and mortality (LoE 1, strong recommendation).**
- **Annual repeated measurements of liver stiffness can be used to refine risk stratification in patients with cACLD (LoE 5, weak recommendation)**
- **Liver stiffness can be used in addition to clinical variables and accepted risk scores to stratify the risk of hepatocellular carcinoma in patients with cACLD due to HBV (LoE 3, weak recommendation)**

Liver-related mortality is almost invariably preceded by clinical decompensation.

Therefore, clinical decompensation has to be considered the most relevant event to predict (together with the onset of hepatocellular carcinoma) in cACLD.

Several NITs hold prognostic value, but only few of them have been extensively validated in compensated patients.

Among serum markers and combination of blood tests, ELF™ and von Willebrand Factor [262] have been associated with the development of clinical decompensation and mortality in patients with cACLD. vWF had an accuracy similar to that of MELD score for mortality in one study (AUROC 0.71 for vWF-Ag vs 0.65 for MELD; $p = 0.2$). Its prognostic value was independent of HVPG values and associated with markers of bacterial translocation and inflammation in another study [293]. Data do not seem sufficient to recommend the use of vWF in clinical practice.

Liver stiffness by TE (and to lesser extent by pSWE and 2DSWE) is a strong and validated predictor of first clinical decompensation, risk of hepatocellular carcinoma and death in patients with compensated chronic liver disease [294, 295]. Its accuracy was similar to that of HVPG for predicting decompensation in two studies [296, 297]. Studies focusing on patients with cACLD confirmed the prognostic value of LSM by TE, which is maintained even above the threshold indicating clinically significant portal hypertension, indicating a higher risk with higher values [298, 299].

In untreated HCV-related cirrhosis, both liver and spleen stiffness predicted clinical decompensation, the latter showing a stronger predictive value (independent of MELD

score; cut-off for discrimination: 54 kPa) [297]. However, data regarding spleen stiffness are still insufficient to recommend its use for prognostic assessment in cACLD.

Similarly, few papers regarding the prognostic value of liver and spleen stiffness by pSWE and 2DSWE, and regarding changes of LSM in the follow-up are available. As expected, increased values correlate with worse prognosis.

Among imaging parameters, Liver Surface Nodularity Score on routine CT remained associated with clinical decompensation and mortality independent of MELD score in patients with cirrhosis [300]. In addition, signs of portal hypertension and in particular the presence and area of porto-systemic collaterals are strongly associated with the development of clinical decompensation in patients with compensated cirrhosis, independent of liver function [301, 302].

Figure 2 shows how simple and readily available NITs can be applied to support clinical decisions in patients with cACLD.

Detailed information regarding HCC screening is provided in the EASL CPGs on HCC and in each of the EASL CPGs referring to specific diseases. Data regarding the prediction of HCC mostly come from cross-sectional and longitudinal studies in Asian cohorts of patients with untreated HBV.

12 risk scores have been developed up to date based on clinical and laboratory characteristics (presence of cirrhosis, male gender, age, viral load) showing AUROCs 0.76-0.95 in the original cohorts.

The Page-B score and the modified PAGE-B scores have been externally validated and showed good results in Caucasian and Asian cohorts, with negative predictive values 0.95-0.99 at 5 years [303].

Liver stiffness is significantly associated with a higher risk of HCC, and changes in LS correlate with changes in the risk of HCC [304] in patients with cirrhosis due to HBV and HCV. A scoring system based on LSM, age, serum albumin and hepatitis B virus (HBV) DNA has been developed in patients with HBV [305] and ELF™ score can further refine risk stratification in patients with intermediate risk according to liver stiffness-HCC risk index. Interestingly, the magnitude of the reduction of LSM on antiviral therapy (HBV) [306, 307] or after achieving SVR (HCV) [101] is inversely associated with the risk of HCC in

patients with cirrhosis due to these etiologies. However, data is insufficient to indicate which patients can avoid HCC screening.

References

- [1] Yano M, Kumada H, Kage M, Ikeda K, Shimamatsu K, Inoue O, et al. The long-term pathological evolution of chronic hepatitis C. *Hepatology* 1996;23:1334-1340.
- [2] Younossi ZM, Stepanova M, Rafiq N, Makhlouf H, Younoszai Z, Agrawal R, et al. Pathologic criteria for nonalcoholic steatohepatitis: interprotocol agreement and ability to predict liver-related mortality. *Hepatology* 2011;53:1874-1882.
- [3] Lackner C, Tiniakos D. Fibrosis and alcohol-related liver disease. *J Hepatol* 2019;70:294-304.
- [4] Garcia-Tsao G, Friedman S, Iredale J, Pinzani M. Now there are many (stages) where before there was one: In search of a pathophysiological classification of cirrhosis. *Hepatology* 2010;51:1445-1449.
- [5] European Association for Study of Liver, Asociacion Latinoamericana para el Estudio del H. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. *J Hepatol* 2015;63:237-264.
- [6] Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 2003;38:1449-1457.
- [7] Davison BA, Harrison SA, Cotter G, Alkhouiri N, Sanyal A, Edwards C, et al. Suboptimal reliability of liver biopsy evaluation has implications for randomized clinical trials. *J Hepatol* 2020.
- [8] Hytiroglou P, Theise ND. Regression of human cirrhosis: an update, 18 years after the pioneering article by Wanless et al. *Virchows Arch* 2018;473:15-22.
- [9] Usher-Smith JA, Sharp, Stephen J, Griffin SJ. The spectrum effect in tests for risk prediction, screening, and diagnosis. *Bmj* 2016;353.
- [10] Mehta SH, Lau B, Afdhal NH, Thomas DL. Exceeding the limits of liver histology markers. *J Hepatol* 2009;50:36-41.
- [11] Kjærgaard M, Thiele M, Jansen C, Stæhr Madsen B, Gørtzen J, Strassburg C, et al. High risk of misinterpreting liver and spleen stiffness using 2D shear-wave and transient elastography after a moderate or high calorie meal. *PLoS One* 2017;12:e0173992.
- [12] Ratchasettakul K, Rattanasiri S, Promson K, Sringam P, Sobhonslidsuk A. The inverse effect of meal intake on controlled attenuation parameter and liver stiffness as assessed by transient elastography. *BMC gastroenterology* 2017;17:50.
- [13] Vuppalanchi R, Weber R, Russell S, Gawrieh S, Samala N, Slaven JE, et al. Is Fasting Necessary for Individuals With Nonalcoholic Fatty Liver Disease to Undergo Vibration-Controlled Transient Elastography? *Am J Gastroenterol* 2019;114:995-997.
- [14] Cornberg M, Tacke F, Karlsen TH, European Association for the Study of the L. Clinical Practice Guidelines of the European Association for the study of the Liver - Advancing methodology but preserving practicability. *J Hepatol* 2019;70:5-7.
- [15] Pimpin L, Cortez-Pinto H, Negro F, Corbould E, Lazarus JV, Webber L, et al. Burden of liver disease in Europe: Epidemiology and analysis of risk factors to identify prevention policies. *J Hepatol* 2018;69:718-735.
- [16] Karlsen TH, Tacke F. <<The times they are a'changin'>> - Positioning the European Association for the Study of the Liver in the changing landscape of hepatology. *J Hepatol* 2018;68:873-875.
- [17] de Franchis R, Baveno VIF. Expanding consensus in portal hypertension: Report of the Baveno VI Consensus Workshop: Stratifying risk and individualizing care for portal hypertension. *J Hepatol* 2015;63:743-752.
- [18] Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of internal medicine* 2011;155:529-536.
- [19] Howick J, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, Liberati A, et al. Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence (Background Document). In: *Medicine*. OCEBM, editor.; 2011.
- [20] Crossan C, Tsochatzis EA, Longworth L, Gurusamy K, Davidson B, Rodriguez-Peralvarez M, et al. Cost-effectiveness of non-invasive methods for assessment and monitoring of liver fibrosis and cirrhosis in

- patients with chronic liver disease: systematic review and economic evaluation. *Health technology assessment (Winchester, England)* 2015;19:1-409.
- [21] Crossan C, Majumdar A, Srivastava A, Thorburn D, Rosenberg W, Pinzani M, et al. Referral pathways for patients with NAFLD based on non-invasive fibrosis tests: Diagnostic accuracy and cost analysis. *Liver Int* 2019;39:2052-2060.
- [22] Thiele M, Madsen BS, Hansen JF, Detlefsen S, Antonsen S, Krag A. Accuracy of the Enhanced Liver Fibrosis Test vs Fibrotest, Elastography and Indirect Markers in Detection of Advanced Fibrosis in Patients with Alcoholic Liver Disease. *Gastroenterology* 2018;154:1369-1379.
- [23] Hagström H, Talbäck M, Andreasson A, Walldius G, Hammar N. Ability of Noninvasive Scoring Systems to Identify Individuals in the Population at Risk for Severe Liver Disease. *Gastroenterology* 2020;158:200-214.
- [24] Hagström H, Talbäck M, Andreasson A, Walldius G, Hammar N. Repeated FIB-4 measurements can help identify individuals at risk of severe liver disease. *J Hepatol* 2020;73:1023-1029.
- [25] Newsome PN, Cramb R, Davison SM, Dillon JF, Foulerton M, Godfrey EM, et al. Guidelines on the management of abnormal liver blood tests. *Gut* 2018;67:6-19.
- [26] Harris R, Harman DJ, Card TR, Aithal GP, Guha IN. Prevalence of clinically significant liver disease within the general population, as defined by non-invasive markers of liver fibrosis: a systematic review. *Lancet Gastroenterol Hepatol* 2017;2:288-297.
- [27] Zelber-Sagi S, Lotan R, Shlomai A, Webb M, Harrari G, Buch A, et al. Predictors for incidence and remission of NAFLD in the general population during a seven-year prospective follow-up. *J Hepatol* 2012;56:1145-1151.
- [28] Wong VW, Chu WC, Wong GL, Chan RS, Chim AM, Ong A, et al. Prevalence of non-alcoholic fatty liver disease and advanced fibrosis in Hong Kong Chinese: a population study using proton-magnetic resonance spectroscopy and transient elastography. *Gut* 2012;61:409-415.
- [29] Caballeria L, Pera G, Arteaga I, Rodriguez L, Aluma A, Morillas RM, et al. High Prevalence of Liver Fibrosis Among European Adults With Unknown Liver Disease: A Population-Based Study. *Clin Gastroenterol Hepatol* 2018;16:1138-1145 e1135.
- [30] Chalasani N, Younossi Z, Lavine JE, Charlton M, Cusi K, Rinella M, et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. *Hepatology* 2018;67:328-357.
- [31] Abeysekera KWM, Fernandes GS, Hammerton G, Portal AJ, Gordon FH, Heron J, et al. Prevalence of steatosis and fibrosis in young adults in the UK: a population-based study. *Lancet Gastroenterol Hepatol* 2020;5:295-305.
- [32] Moessner BK, Andersen ES, Weis N, Laursen AL, Ingerslev J, Lethagen S, et al. Previously unrecognized advanced liver disease unveiled by transient elastography in patients with Haemophilia and chronic hepatitis C. *Haemophilia* 2011;17:938-943.
- [33] Roulot D, Costes JL, Buyck JF, Warzocha U, Gambier N, Czernichow S, et al. Transient elastography as a screening tool for liver fibrosis and cirrhosis in a community-based population aged over 45 years. *Gut* 2011;60:977-984.
- [34] Grattagliano I, Ubaldi E, Napoli L, Marulli CF, Nebiacolombo C, Cottone C, et al. Utility of noninvasive methods for the characterization of nonalcoholic liver steatosis in the family practice. The "VARES" Italian multicenter study. *Annals of hepatology* 2013;12:70-77.
- [35] Lemoine M, Shimakawa Y, Njie R, Njai HF, Nayagam S, Khalil M, et al. Food intake increases liver stiffness measurements and hampers reliable values in patients with chronic hepatitis B and healthy controls: the PROLIFICA experience in The Gambia. *Aliment Pharmacol Ther* 2014;39:188-196.
- [36] Harman DJ, Ryder SD, James MW, Jelpke M, Ottey DS, Wilkes EA, et al. Direct targeting of risk factors significantly increases the detection of liver cirrhosis in primary care: a cross-sectional diagnostic study utilising transient elastography. *BMJ open* 2015;5:e007516.
- [37] Srivastava A, Jong S, Gola A, Gailer R, Morgan S, Sennett K, et al. Cost-comparison analysis of FIB-4, ELF and fibroscan in community pathways for non-alcoholic fatty liver disease. *BMC gastroenterology* 2019;19:122.

- [38] Harman DJ, Ryder SD, James MW, Wilkes EA, Card TR, Aithal GP, et al. Obesity and type 2 diabetes are important risk factors underlying previously undiagnosed cirrhosis in general practice: a cross-sectional study using transient elastography. *Aliment Pharmacol Ther* 2018;47:504-515.
- [39] Majumdar A, Campos S, Gurusamy K, Pinzani M, Tsochatzis EA. Defining the Minimum Acceptable Diagnostic Accuracy of Noninvasive Fibrosis Testing in Cirrhosis: A Decision Analytic Modeling Study. *Hepatology* 2020;71:627-642.
- [40] Tanajewski L, Harris R, Harman DJ, Aithal GP, Card TR, Gkountouras G, et al. Economic evaluation of a community-based diagnostic pathway to stratify adults for non-alcoholic fatty liver disease: a Markov model informed by a feasibility study. *BMJ open* 2017;7:e015659.
- [41] Asphaug L, Thiele M, Krag A, Melberg HO. Cost-Effectiveness of Noninvasive Screening for Alcohol-Related Liver Fibrosis. *Hepatology* 2020;71:2093-2104.
- [42] Serra-Burriel M, Graupera I, Toran P, Thiele M, Roulot D, Wai-Sun Wong V, et al. Transient elastography for screening of liver fibrosis: Cost-effectiveness analysis from six prospective cohorts in Europe and Asia. *J Hepatol* 2019;71:1141-1151.
- [43] NifHaCE. N. Non-alcoholic fatty liver disease (NAFLD): assessment and management. . London; 2016.
- [44] Nouredin M, Jones C, Alkhouri N, Gomez EV, Dieterich DT, Rinella ME, et al. Screening for Nonalcoholic Fatty Liver Disease in Persons with Type 2 Diabetes in the United States Is Cost-effective: A Comprehensive Cost-Utility Analysis. *Gastroenterology* 2020;159:1985-1987 e1984.
- [45] WHO. Global status report on alcohol and health. Available from who.int/substance_abuse/publications/global_alcohol_report/en/; 2018.
- [46] Kim D, Adejumo AC, Yoo ER, Iqbal U, Li AA, Pham EA, et al. Trends in Mortality From Extrahepatic Complications in Patients With Chronic Liver Disease, From 2007 Through 2017. *Gastroenterology* 2019;157:1055-1066.e1011.
- [47] Ratib S, Fleming KM, Crooks CJ, Aithal GP, West J. 1 and 5 year survival estimates for people with cirrhosis of the liver in England, 1998–2009: A large population study. *J Hepatol* 2014;60:282-289.
- [48] European Association for the Study of the Liver. EASL Clinical Practice Guidelines: Management of alcohol-related liver disease. *J Hepatol* 2018;69:154-181.
- [49] Mueller S, Englert S, Seitz HK, Badea RI, Erhardt A, Bozaari B, et al. Inflammation-adapted liver stiffness values for improved fibrosis staging in patients with hepatitis C virus and alcoholic liver disease. *Liver Int* 2015;35:2514-2521.
- [50] Thiele M, Detlefsen S, Sevelsted Moller L, Madsen BS, Fuglsang Hansen J, Fialla AD, et al. Transient and 2-Dimensional Shear-Wave Elastography Provide Comparable Assessment of Alcoholic Liver Fibrosis and Cirrhosis. *Gastroenterology* 2016;150:123-133.
- [51] Voican CS, Louvet A, Trabut JB, Njike-Nakseu M, Dharancy S, Sanchez A, et al. Transient elastography alone and in combination with FibroTest((R)) for the diagnosis of hepatic fibrosis in alcoholic liver disease. *Liver Int* 2017;37:1697-1705.
- [52] Nguyen-Khac E, Thiele M, Voican C, Nahon P, Moreno C, Boursier J, et al. Non-invasive diagnosis of liver fibrosis in patients with alcohol-related liver disease by transient elastography: an individual patient data meta-analysis. *Lancet Gastroenterol Hepatol* 2018;3:614-625.
- [53] Pavlov CS, Casazza G, Nikolova D, Tsochatzis E, Burroughs AK, Ivashkin VT, et al. Transient elastography for diagnosis of stages of hepatic fibrosis and cirrhosis in people with alcoholic liver disease. *The Cochrane database of systematic reviews* 2015;1:CD010542.
- [54] Pavlov CS, Casazza G, Nikolova D, Tsochatzis E, Gluud C. Systematic review with meta-analysis: diagnostic accuracy of transient elastography for staging of fibrosis in people with alcoholic liver disease. *Aliment Pharmacol Ther* 2016;43:575-585.
- [55] Papatheodoridi M, Hiriart JB, Lupsor-Platon M, Bronte F, Boursier J, Elshaarawy O, et al. Refining the Baveno VI elastography criteria for the definition of compensated advanced chronic liver disease. *J Hepatol* 2020.
- [56] Thiele M, Hugger MB, Kim Y, Rautou PE, Elkrief L, Jansen C, et al. 2D shear wave liver elastography by Aixplorer to detect portal hypertension in cirrhosis: An individual patient data meta-analysis. *Liver Int* 2020;40:1435-1446.

- [57] Zhang D, Li P, Chen M, Liu L, Liu Y, Zhao Y, et al. Non-invasive assessment of liver fibrosis in patients with alcoholic liver disease using acoustic radiation force impulse elastography. *Abdominal imaging* 2015;40:723-729.
- [58] Kiani A, Brun V, Laine F, Turlin B, Morcet J, Michalak S, et al. Acoustic radiation force impulse imaging for assessing liver fibrosis in alcoholic liver disease. *World J Gastroenterol* 2016;22:4926-4935.
- [59] Gianni E, Forte P, Galli V, Razzolini G, Bardazzi G, Annese V. Prospective Evaluation of Liver Stiffness Using Transient Elastography in Alcoholic Patients Following Abstinence. *Alcohol Alcohol* 2017;52:42-47.
- [60] Mueller S, Nahon P, Rausch V, Peccerella T, Silva I, Yagmur E, et al. Caspase-cleaved K18 fragments increase during alcohol withdrawal and predict liver-related death in patients with ALD. *Hepatology* 2017;66:96-107.
- [61] Mueller S, Millionig G, Sarovska L, Friedrich S, Reimann FM, Pritsch M, et al. Increased liver stiffness in alcoholic liver disease: differentiating fibrosis from steatohepatitis. *World J Gastroenterol* 2010;16:966-972.
- [62] Gelsi E, Dainese R, Truchi R, Mariné-Barjoan E, Anty R, Autuori M, et al. Effect of Detoxification on Liver Stiffness Assessed by Fibroscan® in Alcoholic Patients. *Alcoholism: Clinical and Experimental Research* 2011;35:566-570.
- [63] Soto M, Sampietro-Colom L, Lasalvia L, Mira A, Jiménez W, Navasa M. Cost-effectiveness of enhanced liver fibrosis test to assess liver fibrosis in chronic hepatitis C virus and alcoholic liver disease patients. *World J Gastroenterol* 2017;23:3163-3173.
- [64] Bissonnette J, Altamirano J, Devue C, Roux O, Payance A, Lebrec D, et al. A prospective study of the utility of plasma biomarkers to diagnose alcoholic hepatitis. *Hepatology* 2017;66:555-563.
- [65] Rudler M, Mouri S, Charlotte F, Cluzel P, Ngo Y, Munteanu M, et al. Validation of AshTest as a Non-Invasive Alternative to Transjugular Liver Biopsy in Patients with Suspected Severe Acute Alcoholic Hepatitis. *PLoS one* 2015;10:e0134302.
- [66] Altamirano J, Miquel R, Katoonizadeh A, Abraldes JG, Duarte-Rojo A, Louvet A, et al. A Histologic Scoring System for Prognosis of Patients With Alcoholic Hepatitis. *Gastroenterology* 2014;146:1231-1239.
- [67] Naveau S, Gaude G, Asnacios A, Agostini H, Abella A, Barri-Ova N, et al. Diagnostic and prognostic values of noninvasive biomarkers of fibrosis in patients with alcoholic liver disease. *Hepatology* 2009;49:97-105.
- [68] Cho EJ, Kim MY, Lee JH, Lee IY, Lim YL, Choi DH, et al. Diagnostic and Prognostic Values of Noninvasive Predictors of Portal Hypertension in Patients with Alcoholic Cirrhosis. *PLoS One* 2015;10:e0133935.
- [69] Kasztelan-Szczerbinska B, Surdacka A, Celinski K, Rolinski J, Zwolak A, Miacz S, et al. Prognostic Significance of the Systemic Inflammatory and Immune Balance in Alcoholic Liver Disease with a Focus on Gender-Related Differences. *PLoS one* 2015;10:e0128347.
- [70] Leeming DJ, Veidal SS, Karsdal MA, Nielsen MJ, Trebicka J, Busk T, et al. Pro-C5, a marker of true type V collagen formation and fibrillation, correlates with portal hypertension in patients with alcoholic cirrhosis. *Scand J Gastroenterol* 2015;50:584-592.
- [71] Waidmann O, Brunner F, Herrmann E, Zeuzem S, Piiper A, Kronenberger B. Cytokeratin 18-based cell death markers indicate severity of liver disease and prognosis of cirrhotic patients. *Liver Int* 2016;36:1464-1472.
- [72] Bruns T, Nuraldeen R, Mai M, Stengel S, Zimmermann HW, Yagmur E, et al. Low serum transferrin correlates with acute-on-chronic organ failure and indicates short-term mortality in decompensated cirrhosis. *Liver Int* 2017;37:232-241.
- [73] da Silva TE, Costa-Silva M, Correa CG, Denardin G, Alencar MLA, Coelho M, et al. Clinical Significance of Serum Adiponectin and Resistin Levels in Liver Cirrhosis. *Annals of hepatology* 2018;17:286-299.
- [74] Jansen C, Möller P, Meyer C, Kolbe CC, Bogs C, Pohlmann A, et al. Increase in liver stiffness after transjugular intrahepatic portosystemic shunt is associated with inflammation and predicts mortality. *Hepatology* 2018;67:1472-1484.

- [75] Kim JH, Lee M, Park SW, Kang M, Kim M, Lee SH, et al. Validation of modified fibrosis-4 index for predicting hepatocellular carcinoma in patients with compensated alcoholic liver cirrhosis. *Medicine* 2018;97:e13438.
- [76] Lindner F, Mühlberg R, Wiegand J, Tröltzsch M, Hoffmeister A, Keim V, et al. Predictive value of liver and spleen stiffness in advanced alcoholic cirrhosis with refractory ascites. *Z Gastroenterol* 2018;56:561-568.
- [77] Macdonald S, Andreola F, Bachtiger P, Amoros A, Pavesi M, Mookerjee R, et al. Cell death markers in patients with cirrhosis and acute decompensation. *Hepatology* 2018;67:989-1002.
- [78] Seo YS, Park SY, Kim MY, Kim SG, Park JY, Yim HJ, et al. Serum cystatin C level: An excellent predictor of mortality in patients with cirrhotic ascites. *J Gastroenterol Hepatol* 2018;33:910-917.
- [79] D'Ambrosio R, Aghemo A, Rumi MG, Ronchi G, Donato MF, Paradis V, et al. A morphometric and immunohistochemical study to assess the benefit of a sustained virological response in hepatitis C virus patients with cirrhosis. *Hepatology* 2012;56:532-543.
- [80] European Association for the Study of the Liver. Electronic address eee, Clinical Practice Guidelines Panel C, representative EGB, Panel m. EASL recommendations on treatment of hepatitis C: Final update of the series(). *J Hepatol* 2020;73:1170-1218.
- [81] Singh S, Facciorusso A, Loomba R, Falck-Ytter YT. Magnitude and Kinetics of Decrease in Liver Stiffness After Antiviral Therapy in Patients With Chronic Hepatitis C: A Systematic Review and Meta-analysis. *Clin Gastroenterol Hepatol* 2018;16:27-38 e24.
- [82] Persico M, Rosato V, Aglitti A, Precone D, Corrado M, De Luna A, et al. Sustained virological response by direct antiviral agents in HCV leads to an early and significant improvement of liver fibrosis. *Antivir Ther* 2018;23:129-138.
- [83] D'Ambrosio R, Aghemo A, Fraquelli M, Rumi MG, Donato MF, Paradis V, et al. The diagnostic accuracy of Fibroscan for cirrhosis is influenced by liver morphometry in HCV patients with a sustained virological response. *J Hepatol* 2013;59:251-256.
- [84] Mauro E, Crespo G, Montironi C, Londoño MC, Hernández-Gea V, Ruiz P, et al. Portal pressure and liver stiffness measurements in the prediction of fibrosis regression after sustained virological response in recurrent hepatitis C. *Hepatology* 2018;67:1683-1694.
- [85] Nascimbeni F, Lebray P, Fedchuk L, Oliveira CP, Alvares-da-Silva MR, Varault A, et al. Significant variations in elastometry measurements made within short-term in patients with chronic liver diseases. *Clin Gastroenterol Hepatol* 2015;13:763-771 e761-766.
- [86] Tachi Y, Hirai T, Kojima Y, Miyata A, Ohara K, Ishizu Y, et al. Liver stiffness measurement using acoustic radiation force impulse elastography in hepatitis C virus-infected patients with a sustained virological response. *Aliment Pharmacol Ther* 2016;44:346-355.
- [87] Chen SH, Lai HC, Chiang IP, Su WP, Lin CH, Kao JT, et al. Performance of Acoustic Radiation Force Impulse Elastography for Staging Liver Fibrosis in Patients With Chronic Hepatitis C After Viral Eradication. *Clin Infect Dis* 2020;70:114-122.
- [88] Hsu WF, Lai HC, Su WP, Lin CH, Chuang PH, Chen SH, et al. Rapid decline of noninvasive fibrosis index values in patients with hepatitis C receiving treatment with direct-acting antiviral agents. *BMC gastroenterology* 2019;19:63.
- [89] Tada T, Kumada T, Toyoda H, Sone Y, Takeshima K, Ogawa S, et al. Viral eradication reduces both liver stiffness and steatosis in patients with chronic hepatitis C virus infection who received direct-acting anti-viral therapy. *Aliment Pharmacol Ther* 2018;47:1012-1022.
- [90] D'Ambrosio R, Degasperis E, Aghemo A, Fraquelli M, Lampertico P, Rumi MG, et al. Serological Tests Do Not Predict Residual Fibrosis in Hepatitis C Cirrhotics with a Sustained Virological Response to Interferon. *PloS one* 2016;11:e0155967.
- [91] Carrat F, Fontaine H, Dorival C, Simony M, Diallo A, Hezode C, et al. Clinical outcomes in patients with chronic hepatitis C after direct-acting antiviral treatment: a prospective cohort study. *Lancet* 2019;393:1453-1464.
- [92] Calvaruso V, Cabibbo G, Cacciola I, Petta S, Madonia S, Bellia A, et al. Incidence of Hepatocellular Carcinoma in Patients With HCV-Associated Cirrhosis Treated With Direct-Acting Antiviral Agents. *Gastroenterology* 2018;155:411-421 e414.

- [93] Ripoll C, Groszmann R, Garcia-Tsao G, Grace N, Burroughs A, Planas R, et al. Hepatic venous pressure gradient predicts clinical decompensation in patients with compensated cirrhosis. *Gastroenterology* 2007;133:481-488.
- [94] D'Amico G, Garcia-Pagan JC, Luca A, Bosch J. Hepatic vein pressure gradient reduction and prevention of variceal bleeding in cirrhosis: a systematic review. *Gastroenterology* 2006;131:1611-1624.
- [95] Lens S, Alvarado-Tapias E, Marino Z, Londono MC, E LL, Martinez J, et al. Effects of All-Oral Anti-Viral Therapy on HVPG and Systemic Hemodynamics in Patients With Hepatitis C Virus-Associated Cirrhosis. *Gastroenterology* 2017;153:1273-1283 e1271.
- [96] Mandorfer M, Kozbial K, Schwabl P, Chromy D, Semmler G, Stattermayer AF, et al. Changes in Hepatic Venous Pressure Gradient Predict Hepatic Decompensation in Patients Who Achieved Sustained Virologic Response to Interferon-Free Therapy. *Hepatology* 2020;71:1023-1036.
- [97] Mandorfer M, Kozbial K, Schwabl P, Freissmuth C, Schwarzer R, Stern R, et al. Sustained virologic response to interferon-free therapies ameliorates HCV-induced portal hypertension. *J Hepatol* 2016;65:692-699.
- [98] Lens S, Baiges A, Alvarado-Tapias E, E LL, Martinez J, Fortea JI, et al. Clinical outcome and hemodynamic changes following HCV eradication with oral antiviral therapy in patients with clinically significant portal hypertension. *J Hepatol* 2020.
- [99] Puigvehi M, Londono MC, Torras X, Lorente S, Vergara M, Morillas RM, et al. Impact of sustained virological response with DAAs on gastroesophageal varices and Baveno criteria in HCV-cirrhotic patients. *J Gastroenterol* 2020;55:205-216.
- [100] Degasperi E, D'Ambrosio R, Iavarone M, Sangiovanni A, Aghemo A, Soffredini R, et al. Factors Associated With Increased Risk of De Novo or Recurrent Hepatocellular Carcinoma in Patients With Cirrhosis Treated With Direct-Acting Antivirals for HCV Infection. *Clin Gastroenterol Hepatol* 2019;17:1183-1191 e1187.
- [101] Ravaioli F, Conti F, Brillanti S, Andreone P, Mazzella G, Buonfiglioli F, et al. Hepatocellular carcinoma risk assessment by the measurement of liver stiffness variations in HCV cirrhotics treated with direct acting antivirals. *Dig Liver Dis* 2018;50:573-579.
- [102] Pons M, Rodriguez-Tajes S, Esteban JI, Marino Z, Vargas V, Lens S, et al. Non-invasive prediction of liver-related events in patients with HCV-associated compensated advanced chronic liver disease after oral antivirals. *J Hepatol* 2020;72:472-480.
- [103] Vutien P, Kim NJ, Moon AM, Pearson M, Su F, Berry K, et al. Fibroscan liver stiffness after anti-viral treatment for hepatitis C is independently associated with adverse outcomes. *Aliment Pharmacol Ther* 2020;52:1717-1727.
- [104] Stern C, Castera L. Non-invasive diagnosis of hepatic steatosis. *Hepatology international* 2017;11:70-78.
- [105] Poynard T, Lassailly G, Diaz E, Clement K, Caiazzo R, Tordjman J, et al. Performance of biomarkers FibroTest, ActiTest, SteatoTest, and NashTest in patients with severe obesity: meta analysis of individual patient data. *PloS one* 2012;7:e30325.
- [106] Fedchuk L, Nascimbeni F, Pais R, Charlotte F, Housset C, Ratzu V, et al. Performance and limitations of steatosis biomarkers in patients with nonalcoholic fatty liver disease. *Aliment Pharmacol Ther* 2014;40:1209-1222.
- [107] Cuthbertson DJ, Weickert MO, Lythgoe D, Sprung VS, Dobson R, Shoajee-Moradie F, et al. External validation of the fatty liver index and lipid accumulation product indices, using 1H-magnetic resonance spectroscopy, to identify hepatic steatosis in healthy controls and obese, insulin-resistant individuals. *Eur J Endocrinol* 2014;171:561-569.
- [108] Calori G, Lattuada G, Ragogna F, Garancini MP, Crosignani P, Villa M, et al. Fatty liver index and mortality: the Cremona study in the 15th year of follow-up. *Hepatology* 2011;54:145-152.
- [109] EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *J Hepatol* 2016;64:1388-1402.
- [110] Hernaez R, Lazo M, Bonekamp S, Kamel I, Brancati FL, Guallar E, et al. Diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver: a meta-analysis. *Hepatology* 2011;54:1082-1090.

- [111] Bril F, Ortiz-Lopez C, Lomonaco R, Orsak B, Freckleton M, Chintapalli K, et al. Clinical value of liver ultrasound for the diagnosis of nonalcoholic fatty liver disease in overweight and obese patients. *Liver Internat* 2015;35:2139-2146.
- [112] de Moura Almeida A, Cotrim HP, Barbosa DBV, de Athayde LGM, Santos AS, Bitencourt AGV, et al. Fatty liver disease in severe obese patients: diagnostic value of abdominal ultrasound. *World J Gastroenterol* 2008;14:1415-1418.
- [113] Castera L, Friedrich-Rust M, Loomba R. Noninvasive Assessment of Liver Disease in Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology* 2019;156:1264-1281.e1264.
- [114] Caussy C, Reeder SB, Sirlin CB, Loomba R. Noninvasive, Quantitative Assessment of Liver Fat by MRI-PDFF as an Endpoint in NASH Trials. *Hepatology* 2018;68:763-772.
- [115] Gu J, Liu S, Du S, Zhang Q, Xiao J, Dong Q, et al. Diagnostic value of MRI-PDFF for hepatic steatosis in patients with non-alcoholic fatty liver disease: a meta-analysis. *European radiology* 2019;29:3564-3573.
- [116] Sasso M, Beaugrand M, de Ledinghen V, Douvin C, Marcellin P, Poupon R, et al. Controlled attenuation parameter (CAP): a novel VCTE guided ultrasonic attenuation measurement for the evaluation of hepatic steatosis: preliminary study and validation in a cohort of patients with chronic liver disease from various causes. *Ultrasound in medicine & biology* 2010;36:1825-1835.
- [117] Karlas T, Petroff D, Sasso M, Fan JG, Mi YQ, de Lédinthen V, et al. Individual patient data meta-analysis of controlled attenuation parameter (CAP) technology for assessing steatosis. *J Hepatol* 2017;66:1022-1030.
- [118] Eddowes PJ, Sasso M, Allison M, Tsochatzis E, Anstee QM, Sheridan D, et al. Accuracy of FibroScan Controlled Attenuation Parameter and Liver Stiffness Measurement in Assessing Steatosis and Fibrosis in Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology* 2019;156:1717-1730.
- [119] Siddiqui MS, Vuppalanchi R, Van Natta ML, Hallinan E, Kowdley KV, Abdelmalek M, et al. Vibration-Controlled Transient Elastography to Assess Fibrosis and Steatosis in Patients With Nonalcoholic Fatty Liver Disease. *Clin Gastroenterol Hepatol* 2019;17:156-163 e152.
- [120] Vuppalanchi R, Siddiqui MS, Van Natta ML, Hallinan E, Brandman D, Kowdley K, et al. Performance characteristics of vibration-controlled transient elastography for evaluation of nonalcoholic fatty liver disease. *Hepatology* 2018;67:134-144.
- [121] de Ledinghen V, Wong GL, Vergniol J, Chan HL, Hiriart JB, Chan AW, et al. Controlled attenuation parameter for the diagnosis of steatosis in non-alcoholic fatty liver disease. *J Gastroenterol Hepatol* 2016;31:848-855.
- [122] Petroff DB, V.; Newsome, P.N.; Shalimar; Voican, C.S.; Thiele, M.; de Ledinghen,V.; Baumeler,S.; Chan, W.K.; Perlemuter, G.; Cardoso, A.C.; Aggarwal,S.; Sasso, M.; Eddowes, P.J.; Allison, M.; Tsochatzis, E.; Anstee, Q.M.; Sheridan,D.; Cobbold, J.F.; Naveau, S.; Lupsor-Platon,M.; Mueller, S.; Krag, A.; Irlés-Depe, M.; Semela, D.; Wong, G.L.H.; Wong, V.W.S.; Vilela-Nogueira, C.A.; Grag, H.; Chazouilleres, O.; Wiegand, J.; Karlas, T. Assessment of hepatic steatosis by controlled attenuation parameter using M and XL probes: an individual patient meta-analysis. *Lancet Gastroenterol Hepatol* 2021.
- [123] Caussy C, Alqiraish MH, Nguyen P, Hernandez C, Cepin S, Fortney LE, et al. Optimal threshold of controlled attenuation parameter with MRI-PDFF as the gold standard for the detection of hepatic steatosis. *Hepatology* 2018;67:1348-1359.
- [124] Wong VW, Petta S, Hiriart JB, Camma C, Wong GL, Marra F, et al. Validity criteria for the diagnosis of fatty liver by M probe-based controlled attenuation parameter. *J Hepatol* 2017;67:577-584.
- [125] Imajo K, Kessoku T, Honda Y, Tomeno W, Ogawa Y, Mawatari H, et al. Magnetic Resonance Imaging More Accurately Classifies Steatosis and Fibrosis in Patients With Nonalcoholic Fatty Liver Disease Than Transient Elastography. *Gastroenterology* 2016;150:626-637 e627.
- [126] Park CC, Nguyen P, Hernandez C, Bettencourt R, Ramirez K, Fortney L, et al. Magnetic Resonance Elastography vs Transient Elastography in Detection of Fibrosis and Noninvasive Measurement of Steatosis in Patients With Biopsy-Proven Nonalcoholic Fatty Liver Disease. *Gastroenterology* 2017;152:598-607 e592.
- [127] Runge JH, Smits LP, Verheij J, Depla A, Kuiken SD, Baak BC, et al. MR Spectroscopy-derived Proton Density Fat Fraction Is Superior to Controlled Attenuation Parameter for Detecting and Grading Hepatic Steatosis. *Radiology* 2017:162931.

- [128] Singh S, Allen AM, Wang Z, Prokop LJ, Murad MH, Loomba R. Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies. *Clin Gastroenterol Hepatol* 2015;13:643-654 e641-649; quiz e639-640.
- [129] Verhaegh P, Bavalia R, Winkens B, Masclee A, Jonkers D, Koek G. Noninvasive Tests Do Not Accurately Differentiate Nonalcoholic Steatohepatitis From Simple Steatosis: A Systematic Review and Meta-analysis. *Clin Gastroenterol Hepatol* 2018;16:837-861.
- [130] Dulai PS, Singh S, Patel J, Soni M, Prokop LJ, Younossi Z, et al. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: Systematic review and meta-analysis. *Hepatology* 2017;65:1557-1565.
- [131] Taylor RS, Taylor RJ, Bayliss S, Hagstrom H, Nasr P, Schattenberg JM, et al. Association Between Fibrosis Stage and Outcomes of Patients With Nonalcoholic Fatty Liver Disease: A Systematic Review and Meta-Analysis. *Gastroenterology* 2020;158:1611-1625 e1612.
- [132] Xiao G, Zhu S, Xiao X, Yan L, Yang J, Wu G. Comparison of laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: A meta-analysis. *Hepatology* 2017;66:1486-1501.
- [133] McPherson S, Hardy T, Dufour JF, Petta S, Romero-Gomez M, Allison M, et al. Age as a Confounding Factor for the Accurate Non-Invasive Diagnosis of Advanced NAFLD Fibrosis. *Am J Gastroenterol* 2017;112:740-751.
- [134] Petta S, Wai-Sun Wong V, Bugianesi E, Fracanzani AL, Camma C, Hiriart JB, et al. Impact of Obesity and Alanine Aminotransferase Levels on the Diagnostic Accuracy for Advanced Liver Fibrosis of Noninvasive Tools in Patients With Nonalcoholic Fatty Liver Disease. *Am J Gastroenterol* 2019;114:916-928.
- [135] Joo SK, Kim W, Kim D, Kim JH, Oh S, Lee KL, et al. Steatosis severity affects the diagnostic performances of noninvasive fibrosis tests in nonalcoholic fatty liver disease. *Liver Int* 2018;38:331-341.
- [136] Bril F, McPhaul MJ, Caulfield MP, Clark VC, Soldevilla-Pico C, Firpi-Morell RJ, et al. Performance of Plasma Biomarkers and Diagnostic Panels for Nonalcoholic Steatohepatitis and Advanced Fibrosis in Patients With Type 2 Diabetes. *Diabetes Care* 2020;43:290-297.
- [137] Bertot LC, Jeffrey GP, de Boer B, MacQuillan G, Garas G, Chin J, et al. Diabetes impacts prediction of cirrhosis and prognosis by non-invasive fibrosis models in non-alcoholic fatty liver disease. *Liver Int* 2018;38:1793-1802.
- [138] Vali Y, Lee J, Boursier J, Spijker R, Loffler J, Verheij J, et al. Enhanced liver fibrosis test for the non-invasive diagnosis of fibrosis in patients with NAFLD: A systematic review and meta-analysis. *J Hepatol* 2020;73:252-262.
- [139] Anstee QM, Lawitz EJ, Alkhoury N, Wong VW, Romero-Gomez M, Okanoue T, et al. Noninvasive Tests Accurately Identify Advanced Fibrosis due to NASH: Baseline Data From the STELLAR Trials. *Hepatology* 2019;70:1521-1530.
- [140] Guillaume M, Moal V, Delabaudiere C, Zuberbuhler F, Robic MA, Lannes A, et al. Direct comparison of the specialised blood fibrosis tests FibroMeter(V2G) and Enhanced Liver Fibrosis score in patients with non-alcoholic fatty liver disease from tertiary care centres. *Aliment Pharmacol Ther* 2019;50:1214-1222.
- [141] Wong VW, Irlles M, Wong GL, Shili S, Chan AW, Merrouche W, et al. Unified interpretation of liver stiffness measurement by M and XL probes in non-alcoholic fatty liver disease. *Gut* 2019;68:2057-2064.
- [142] Petta S, Maida M, Macaluso FS, Di Marco V, Camma C, Cabibi D, et al. The severity of steatosis influences liver stiffness measurement in patients with nonalcoholic fatty liver disease. *Hepatology* 2015;62:1101-1110.
- [143] Petta S, Wong VW, Camma C, Hiriart JB, Wong GL, Marra F, et al. Improved noninvasive prediction of liver fibrosis by liver stiffness measurement in patients with nonalcoholic fatty liver disease accounting for controlled attenuation parameter values. *Hepatology* 2017;65:1145-1155.
- [144] Karlas T, Petroff D, Sasso M, Fan JG, Mi YQ, de Ledinghen V, et al. Impact of controlled attenuation parameter on detecting fibrosis using liver stiffness measurement. *Aliment Pharmacol Ther* 2018;47:989-1000.
- [145] Herrmann E, de Ledinghen V, Cassinotto C, Chu WC, Leung VY, Ferraioli G, et al. Assessment of biopsy-proven liver fibrosis by two-dimensional shear wave elastography: An individual patient data-based meta-analysis. *Hepatology* 2018;67:260-272.

- [146] Jiang W, Huang S, Teng H, Wang P, Wu M, Zhou X, et al. Diagnostic accuracy of point shear wave elastography and transient elastography for staging hepatic fibrosis in patients with non-alcoholic fatty liver disease: a meta-analysis. *BMJ open* 2018;8:e021787.
- [147] Cassinotto C, Boursier J, de Ledinghen V, Lebigot J, Lapuyade B, Cales P, et al. Liver stiffness in nonalcoholic fatty liver disease: A comparison of supersonic shear imaging, FibroScan, and ARFI with liver biopsy. *Hepatology* 2016;63:1817-1827.
- [148] Hsu C, Caussy C, Imajo K, Chen J, Singh S, Kaulback K, et al. Magnetic Resonance vs Transient Elastography Analysis of Patients With Nonalcoholic Fatty Liver Disease: A Systematic Review and Pooled Analysis of Individual Participants. *Clin Gastroenterol Hepatol* 2019;17:630-637 e638.
- [149] Petta S, Wong VW, Camma C, Hiriart JB, Wong GL, Vergniol J, et al. Serial combination of non-invasive tools improves the diagnostic accuracy of severe liver fibrosis in patients with NAFLD. *Aliment Pharmacol Ther* 2017;46:617-627.
- [150] Boursier J, de Ledinghen V, Leroy V, Anty R, Francque S, Salmon D, et al. A stepwise algorithm using an at-a-glance first-line test for the non-invasive diagnosis of advanced liver fibrosis and cirrhosis. *J Hepatol* 2017;66:1158-1165.
- [151] Srivastava A, Gailer R, Tanwar S, Trembling P, Parkes J, Rodger A, et al. Prospective evaluation of a primary care referral pathway for patients with non-alcoholic fatty liver disease. *J Hepatol* 2019;71:371-378.
- [152] Siddiqui MS, Yamada G, Vuppalanchi R, Van Natta M, Loomba R, Guy C, et al. Diagnostic Accuracy of Noninvasive Fibrosis Models to Detect Change in Fibrosis Stage. *Clin Gastroenterol Hepatol* 2019;17:1877-1885 e1875.
- [153] Sanyal AJ, Harrison SA, Ratziu V, Abdelmalek MF, Diehl AM, Caldwell S, et al. The Natural History of Advanced Fibrosis Due to Nonalcoholic Steatohepatitis: Data From the Simtuzumab Trials. *Hepatology* 2019;70:1913-1927.
- [154] Angulo P, Bugianesi E, Bjornsson ES, Charatcharoenwittaya P, Mills PR, Barrera F, et al. Simple noninvasive systems predict long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology* 2013;145:782-789 e784.
- [155] Kawamura Y, Arase Y, Ikeda K, Seko Y, Imai N, Hosaka T, et al. Large-scale long-term follow-up study of Japanese patients with non-alcoholic Fatty liver disease for the onset of hepatocellular carcinoma. *Am J Gastroenterol* 2012;107:253-261.
- [156] Sebastiani G, Alshaalan R, Wong P, Rubino M, Salman A, Metrakos P, et al. Prognostic Value of Non-Invasive Fibrosis and Steatosis Tools, Hepatic Venous Pressure Gradient (HVPG) and Histology in Nonalcoholic Steatohepatitis. *PloS one* 2015;10:e0128774.
- [157] Hagstrom H, Nasr P, Ekstedt M, Stal P, Hultcrantz R, Kechagias S. Accuracy of Noninvasive Scoring Systems in Assessing Risk of Death and Liver-Related Endpoints in Patients With Nonalcoholic Fatty Liver Disease. *Clin Gastroenterol Hepatol* 2019;17:1148-1156 e1144.
- [158] Onnerhag K, Hartman H, Nilsson PM, Lindgren S. Non-invasive fibrosis scoring systems can predict future metabolic complications and overall mortality in non-alcoholic fatty liver disease (NAFLD). *Scand J Gastroenterol* 2019;54:328-334.
- [159] Boursier J, Vergniol J, Guillet A, Hiriart JB, Lannes A, Le Bail B, et al. Diagnostic accuracy and prognostic significance of blood fibrosis tests and liver stiffness measurement by FibroScan in non-alcoholic fatty liver disease. *J Hepatol* 2016;65:570-578.
- [160] Kim D, Kim WR, Kim HJ, Therneau TM. Association between noninvasive fibrosis markers and mortality among adults with nonalcoholic fatty liver disease in the United States. *Hepatology* 2013;57:1357-1365.
- [161] Munteanu M, Pais R, Peta V, Deckmyn O, Moussalli J, Ngo Y, et al. Long-term prognostic value of the FibroTest in patients with non-alcoholic fatty liver disease, compared to chronic hepatitis C, B, and alcoholic liver disease. *Aliment Pharmacol Ther* 2018;48:1117-1127.
- [162] Shili-Masmoudi S, Wong GL, Hiriart JB, Liu K, Chermak F, Shu SS, et al. Liver stiffness measurement predicts long-term survival and complications in non-alcoholic fatty liver disease. *Liver Int* 2020;40:581-589.

- [163] Petta S, Sebastiani G, Viganò M, Ampuero J, Wai-Sun Wong V, Boursier J, et al. Monitoring Occurrence of Liver-Related Events and Survival by Transient Elastography in Patients With Nonalcoholic Fatty Liver Disease and Compensated Advanced Chronic Liver disease. *Clin Gastroenterol Hepatol* 2020.
- [164] Ratziu V. A critical review of endpoints for non-cirrhotic NASH therapeutic trials. *J Hepatol* 2018;68:353-361.
- [165] Rinella ME, Tacke F, Sanyal AJ, Anstee QM, participants of the AEW. Report on the AASLD/EASL joint workshop on clinical trial endpoints in NAFLD. *J Hepatol* 2019;71:823-833.
- [166] Ratziu V, Harrison SA, Francque S, Bedossa P, Leheret P, Serfaty L, et al. Elafibranor, an Agonist of the Peroxisome Proliferator-Activated Receptor- α and - δ , Induces Resolution of Nonalcoholic Steatohepatitis Without Fibrosis Worsening. *Gastroenterology* 2016;150:1147-1159 e1145.
- [167] Boursier J, Guillaume M, Leroy V, Irlès M, Roux M, Lannes A, et al. New sequential combinations of non-invasive fibrosis tests provide an accurate diagnosis of advanced fibrosis in NAFLD. *J Hepatol* 2019;71:389-396.
- [168] Boursier J, Anty R, Vonghia L, Moal V, Vanwolleghem T, Canivet CM, et al. Screening for therapeutic trials and treatment indication in clinical practice: MACK-3, a new blood test for the diagnosis of fibrotic NASH. *Aliment Pharmacol Ther* 2018;47:1387-1396.
- [169] Newsome PN, Sasso M, Deeks JJ, Paredes A, Boursier J, Chan WK, et al. FibroScan-AST (FAST) score for the non-invasive identification of patients with non-alcoholic steatohepatitis with significant activity and fibrosis: a prospective derivation and global validation study. *Lancet Gastroenterol Hepatol* 2020;5:362-373.
- [170] Harrison SA, Ratziu V, Boursier J, Francque S, Bedossa P, Majd Z, et al. A blood-based biomarker panel (NIS4) for non-invasive diagnosis of non-alcoholic steatohepatitis and liver fibrosis: a prospective derivation and global validation study. *Lancet Gastroenterol Hepatol* 2020;5:970-985.
- [171] Harrison SA, Rossi SJ, Paredes AH, Trotter JF, Bashir MR, Guy CD, et al. NGM282 Improves Liver Fibrosis and Histology in 12 Weeks in Patients With Nonalcoholic Steatohepatitis. *Hepatology* 2020;71:1198-1212.
- [172] Nouredin M, Lam J, Peterson MR, Middleton M, Hamilton G, Le TA, et al. Utility of magnetic resonance imaging versus histology for quantifying changes in liver fat in nonalcoholic fatty liver disease trials. *Hepatology* 2013;58:1930-1940.
- [173] Middleton MS, Heba ER, Hooker CA, Bashir MR, Fowler KJ, Sandrasegaran K, et al. Agreement Between Magnetic Resonance Imaging Proton Density Fat Fraction Measurements and Pathologist-Assigned Steatosis Grades of Liver Biopsies From Adults With Nonalcoholic Steatohepatitis. *Gastroenterology* 2017;153:753-761.
- [174] Jayakumar S, Middleton MS, Lawitz EJ, Mantry PS, Caldwell SH, Arnold H, et al. Longitudinal correlations between MRE, MRI-PDFF, and liver histology in patients with non-alcoholic steatohepatitis: Analysis of data from a phase II trial of selonsertib. *J Hepatol* 2019;70:133-141.
- [175] Loomba R. MRI-PDFF treatment response criteria in nonalcoholic steatohepatitis. *Hepatology* 2020.
- [176] Le TA, Chen J, Changchien C, Peterson MR, Kono Y, Patton H, et al. Effect of colestevlam on liver fat quantified by magnetic resonance in nonalcoholic steatohepatitis: a randomized controlled trial. *Hepatology* 2012;56:922-932.
- [177] Loomba R, Sanyal AJ, Kowdley KV, Terrault N, Chalasani NP, Abdelmalek MF, et al. Factors Associated With Histologic Response in Adult Patients With Nonalcoholic Steatohepatitis. *Gastroenterology* 2019;156:88-95 e85.
- [178] Stine JG, Munaganuru N, Barnard A, Wang JL, Kaulback K, Argo CK, et al. Change in MRI-PDFF and Histologic Response in Patients with Nonalcoholic Steatohepatitis: A Systematic Review and Meta-Analysis. *Clin Gastroenterol Hepatol* 2020.
- [179] Bril F, Barb D, Lomonaco R, Lai J, Cusi K. Change in hepatic fat content measured by MRI does not predict treatment-induced histological improvement of steatohepatitis. *J Hepatol* 2020;72:401-410.
- [180] Roll J, Boyer JL, Barry D, Klatskin G. The prognostic importance of clinical and histologic features in asymptomatic and symptomatic primary biliary cirrhosis. *N Engl J Med* 1983;308:1-7.
- [181] Poupon RE, Bonnard AM, Chretien Y, Poupon R. Ten-year survival in ursodeoxycholic acid-treated patients with primary biliary cirrhosis. The UDCA-PBC Study Group. *Hepatology* 1999;29:1668-1671.

- [182] Poupon R, Chazouilleres O, Balkau B, Poupon RE. Clinical and biochemical expression of the histopathological lesions of primary biliary cirrhosis. UDCA-PBC Group. *J Hepatol* 1999;30:408-412.
- [183] Corpechot C, Abenavoli L, Rabahi N, Chretien Y, Andreani T, Johanet C, et al. Biochemical response to ursodeoxycholic acid and long-term prognosis in primary biliary cirrhosis. *Hepatology* 2008;48:871-877.
- [184] Floreani A, Caroli D, Variola A, Rizzotto ER, Antoniazzi S, Chiaramonte M, et al. A 35-year follow-up of a large cohort of patients with primary biliary cirrhosis seen at a single centre. *Liver Int* 2011;31:361-368.
- [185] Carbone M, D'Amato D, Hirschfield GM, Jones DEJ, Mells GF. Letter: histology is relevant for risk stratification in primary biliary cholangitis. *Aliment Pharmacol Ther* 2020;51:192-193.
- [186] Murillo Perez CF, Hirschfield GM, Corpechot C, Floreani A, Mayo MJ, van der Meer A, et al. Fibrosis stage is an independent predictor of outcome in primary biliary cholangitis despite biochemical treatment response. *Aliment Pharmacol Ther* 2019;50:1127-1136.
- [187] European Association for the Study of the Liver. Electronic address eee, European Association for the Study of the L. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. *J Hepatol* 2017;67:145-172.
- [188] Park DW, Lee YJ, Chang W, Park JH, Lee KH, Kim YH, et al. Diagnostic performance of a point shear wave elastography (pSWE) for hepatic fibrosis in patients with autoimmune liver disease. *PloS one* 2019;14:e0212771.
- [189] Koizumi Y, Hirooka M, Abe M, Tokumoto Y, Yoshida O, Watanabe T, et al. Comparison between real-time tissue elastography and vibration-controlled transient elastography for the assessment of liver fibrosis and disease progression in patients with primary biliary cholangitis. *Hepatology research* 2017;47:1252-1259.
- [190] Meng J, Xu H, Liu X, Wu R, Niu J. Increased red cell width distribution to lymphocyte ratio is a predictor of histologic severity in primary biliary cholangitis. *Medicine* 2018;97:e13431.
- [191] Jiang X, Wang Y, Su Z, Yang F, Lv H, Lin L, et al. Red blood cell distribution width to platelet ratio levels in assessment of histologic severity in patients with primary biliary cholangitis. *Scand J Clin Lab Invest* 2018;78:258-263.
- [192] Olmez S, Sayar S, Avcioglu U, Tenlik I, Ozaslan E, Koseoglu HT, et al. The relationship between liver histology and noninvasive markers in primary biliary cirrhosis. *Eur J Gastroenterol Hepatol* 2016;28:773-776.
- [193] Wang H, Xu H, Wang X, Wu R, Gao X, Jin Q, et al. Red Blood Cell Distribution Width to Platelet Ratio is Related to Histologic Severity of Primary Biliary Cirrhosis. *Medicine* 2016;95:e3114.
- [194] Stasi C, Leoncini L, Biagini MR, Arena U, Madaia S, Laffi G, et al. Assessment of liver fibrosis in primary biliary cholangitis: Comparison between indirect serum markers and fibrosis morphometry. *Dig Liver Dis* 2016;48:298-301.
- [195] Tahtaci M, Yurekli OT, Bolat AD, Balci S, Akin FE, Buyukasik NS, et al. Increased mean platelet volume is related to histologic severity of primary biliary cirrhosis. *Eur J Gastroenterol Hepatol* 2015;27:1382-1385.
- [196] Milas GP, Karageorgiou V, Cholongitas E. Red cell distribution width to platelet ratio for liver fibrosis: a systematic review and meta-analysis of diagnostic accuracy. *Expert review of gastroenterology & hepatology* 2019;13:877-891.
- [197] Wang Z, Liu X, Xu H, Qu L, Zhang D, Gao P. Platelet count to spleen thickness ratio is related to histologic severity of primary biliary cholangitis. *Medicine* 2018;97:e9843.
- [198] Corpechot C, Carrat F, Poujol-Robert A, Gaouar F, Wendum D, Chazouilleres O, et al. Noninvasive elastography-based assessment of liver fibrosis progression and prognosis in primary biliary cirrhosis. *Hepatology* 2012;56:198-208.
- [199] Gomez-Dominguez E, Mendoza J, Garcia-Buey L, Trapero M, Gisbert JP, Jones EA, et al. Transient elastography to assess hepatic fibrosis in primary biliary cirrhosis. *Aliment Pharmacol Ther* 2008;27:441-447.
- [200] Osman KT, Maselli DB, Idilman IS, Rowan DJ, Viehman JK, Harmsen WS, et al. Liver Stiffness Measured by Either Magnetic Resonance or Transient Elastography Is Associated With Liver Fibrosis and Is an Independent Predictor of Outcomes Among Patients With Primary Biliary Cholangitis. *J Clin Gastroenterol* 2020.

- [201] Ehlken H, Wroblewski R, Corpechot C, Arrive L, Rieger T, Hartl J, et al. Validation of Transient Elastography and Comparison with Spleen Length Measurement for Staging of Fibrosis and Clinical Prognosis in Primary Sclerosing Cholangitis. *PLoS one* 2016;11:e0164224.
- [202] Krawczyk M, Ligocka J, Ligocki M, Raszeja-Wyszomirska J, Milkiewicz M, Szparecki G, et al. Does transient elastography correlate with liver fibrosis in patients with PSC? Laennec score-based analysis of explanted livers. *Scand J Gastroenterol* 2017;52:1407-1412.
- [203] Muir AJ, Levy C, Janssen HLA, Montano-Loza AJ, Shiffman ML, Caldwell S, et al. Simtuzumab for Primary Sclerosing Cholangitis: Phase 2 Study Results With Insights on the Natural History of the Disease. *Hepatology* 2019;69:684-698.
- [204] Eaton JE, Dzyubak B, Venkatesh SK, Smyrk TC, Gores GJ, Ehman RL, et al. Performance of magnetic resonance elastography in primary sclerosing cholangitis. *J Gastroenterol Hepatol* 2016;31:1184-1190.
- [205] Millionig G, Reimann FM, Friedrich S, Fonouni H, Mehrabi A, Buchler MW, et al. Extrahepatic cholestasis increases liver stiffness (FibroScan) irrespective of fibrosis. *Hepatology* 2008;48:1718-1723.
- [206] Ehlken H, Lohse AW, Schramm C. Transient elastography in primary sclerosing cholangitis-the value as a prognostic factor and limitations. *Gastroenterology* 2014;147:542-543.
- [207] Ehlken H, Wroblewski R, Corpechot C, Arrive L, Lezius S, Hartl J, et al. Spleen size for the prediction of clinical outcome in patients with primary sclerosing cholangitis. *Gut* 2016;65:1230-1232.
- [208] Lammers WJ, Hirschfield GM, Corpechot C, Nevens F, Lindor KD, Janssen HL, et al. Development and Validation of a Scoring System to Predict Outcomes of Patients With Primary Biliary Cirrhosis Receiving Ursodeoxycholic Acid Therapy. *Gastroenterology* 2015;149:1804-1812 e1804.
- [209] Carbone M, Sharp SJ, Flack S, Paximadas D, Spiess K, Adgey C, et al. The UK-PBC risk scores: Derivation and validation of a scoring system for long-term prediction of end-stage liver disease in primary biliary cholangitis. *Hepatology* 2016;63:930-950.
- [210] Yang F, Yang Y, Wang Q, Wang Z, Miao Q, Xiao X, et al. The risk predictive values of UK-PBC and GLOBE scoring system in Chinese patients with primary biliary cholangitis: the additional effect of anti-gp210. *Aliment Pharmacol Ther* 2017;45:733-743.
- [211] Cheung KS, Seto WK, Fung J, Lai CL, Yuen MF. Prognostic Factors for Transplant-Free Survival and Validation of Prognostic Models in Chinese Patients with Primary Biliary Cholangitis Receiving Ursodeoxycholic Acid. *Clinical and translational gastroenterology* 2017;8:e100.
- [212] Efe C, Tascilar K, Henriksson I, Lytvyak E, Alalkim F, Trivedi H, et al. Validation of Risk Scoring Systems in Ursodeoxycholic Acid-Treated Patients With Primary Biliary Cholangitis. *Am J Gastroenterol* 2019;114:1101-1108.
- [213] Cheung AC, Gulamhusein AF, Juran BD, Schlicht EM, McCauley BM, de Andrade M, et al. External validation of the United Kingdom-primary biliary cholangitis risk scores of patients with primary biliary cholangitis treated with ursodeoxycholic acid. *Hepatology* 2018;67:676-682.
- [214] Harms MH, Lammers WJ, Thorburn D, Corpechot C, Invernizzi P, Janssen HLA, et al. Major Hepatic Complications in Ursodeoxycholic Acid-Treated Patients With Primary Biliary Cholangitis: Risk Factors and Time Trends in Incidence and Outcome. *Am J Gastroenterol* 2018;113:254-264.
- [215] Gerussi A, Bernasconi DP, O'Donnell SE, Lammers WJ, Van Buuren H, Hirschfield G, et al. Measurement of Gamma Glutamyl Transferase to Determine Risk of Liver Transplantation or Death in Patients With Primary Biliary Cholangitis. *Clin Gastroenterol Hepatol* 2020.
- [216] Murillo Perez CF, Harms MH, Lindor KD, van Buuren HR, Hirschfield GM, Corpechot C, et al. Goals of Treatment for Improved Survival in Primary Biliary Cholangitis: Treatment Target Should Be Bilirubin Within the Normal Range and Normalization of Alkaline Phosphatase. *Am J Gastroenterol* 2020;115:1066-1074.
- [217] Mayo MJ, Parkes J, Adams-Huet B, Combes B, Mills AS, Markin RS, et al. Prediction of clinical outcomes in primary biliary cirrhosis by serum enhanced liver fibrosis assay. *Hepatology* 2008;48:1549-1557.
- [218] Weismuller TJ, Trivedi PJ, Bergquist A, Imam M, Lenzen H, Ponsioen CY, et al. Patient Age, Sex, and Inflammatory Bowel Disease Phenotype Associate With Course of Primary Sclerosing Cholangitis. *Gastroenterology* 2017;152:1975-1984 e1978.

- [219] Berntsen NL, Klingenberg O, Juran BD, Benito de Valle M, Lindkvist B, Lazaridis KN, et al. Association Between HLA Haplotypes and Increased Serum Levels of IgG4 in Patients With Primary Sclerosing Cholangitis. *Gastroenterology* 2015;148:924-927 e922.
- [220] Mendes FD, Jorgensen R, Keach J, Katzmann JA, Smyrk T, Donlinger J, et al. Elevated serum IgG4 concentration in patients with primary sclerosing cholangitis. *Am J Gastroenterol* 2006;101:2070-2075.
- [221] de Vries EM, de Krijger M, Farkkila M, Arola J, Schirmacher P, Gotthardt D, et al. Validation of the prognostic value of histologic scoring systems in primary sclerosing cholangitis: An international cohort study. *Hepatology* 2017;65:907-919.
- [222] Ponsioen CY, Chapman RW, Chazouilleres O, Hirschfield GM, Karlsen TH, Lohse AW, et al. Surrogate endpoints for clinical trials in primary sclerosing cholangitis: Review and results from an International PSC Study Group consensus process. *Hepatology* 2016;63:1357-1367.
- [223] Vesterhus M, Hov JR, Holm A, Schruppf E, Nygard S, Godang K, et al. Enhanced liver fibrosis score predicts transplant-free survival in primary sclerosing cholangitis. *Hepatology* 2015;62:188-197.
- [224] de Vries EM, Farkkila M, Milkiewicz P, Hov JR, Eksteen B, Thorburn D, et al. Enhanced liver fibrosis test predicts transplant-free survival in primary sclerosing cholangitis, a multi-centre study. *Liver Int* 2017;37:1554-1561.
- [225] Vesterhus M, Holm A, Hov JR, Nygard S, Schruppf E, Melum E, et al. Novel serum and bile protein markers predict primary sclerosing cholangitis disease severity and prognosis. *J Hepatol* 2017;66:1214-1222.
- [226] Nielsen MJ, Thorburn D, Leeming DJ, Hov JR, Nygard S, Moum B, et al. Serological markers of extracellular matrix remodeling predict transplant-free survival in primary sclerosing cholangitis. *Aliment Pharmacol Ther* 2018;48:179-189.
- [227] Dhillon AK, Kremer AE, Kummen M, Boberg KM, Elferink RPO, Karlsen TH, et al. Autotaxin activity predicts transplant-free survival in primary sclerosing cholangitis. *Sci Rep* 2019;9:8450.
- [228] de Vries EM, Wang J, Williamson KD, Leeflang MM, Boonstra K, Weersma RK, et al. A novel prognostic model for transplant-free survival in primary sclerosing cholangitis. *Gut* 2018;67:1864-1869.
- [229] Goet JC, Floreani A, Verhelst X, Cazzagon N, Perini L, Lammers WJ, et al. Validation, clinical utility and limitations of the Amsterdam-Oxford model for primary sclerosing cholangitis. *J Hepatol* 2019;71:992-999.
- [230] Eaton JE, Vesterhus M, McCauley BM, Atkinson EJ, Schlicht EM, Juran BD, et al. Primary Sclerosing Cholangitis Risk Estimate Tool (PREsTo) Predicts Outcomes of the Disease: A Derivation and Validation Study Using Machine Learning. *Hepatology* 2020;71:214-224.
- [231] Goode EC, Clark AB, Mells GF, Srivastava B, Spiess K, Gelson WTH, et al. Factors Associated With Outcomes of Patients With Primary Sclerosing Cholangitis and Development and Validation of a Risk Scoring System. *Hepatology* 2019;69:2120-2135.
- [232] Corpechot C, Gaouar F, El Naggar A, Kemgang A, Wendum D, Poupon R, et al. Baseline values and changes in liver stiffness measured by transient elastography are associated with severity of fibrosis and outcomes of patients with primary sclerosing cholangitis. *Gastroenterology* 2014;146:970-979; quiz e915-976.
- [233] Jung F, Cazzagon N, Vettorazzi E, Corpechot C, Chazouilleres O, Arrive L, et al. Rate of Spleen Length Progression Is a Marker of Outcome in Patients With Primary Sclerosing Cholangitis. *Clin Gastroenterol Hepatol* 2019;17:2613-2615.
- [234] Khoshpouri P, Ameli S, Ghasabeh MA, Pandey A, Zarghampour M, Varzaneh FN, et al. Correlation between quantitative liver and spleen volumes and disease severity in primary sclerosing cholangitis as determined by Mayo risk score. *Eur J Radiol* 2018;108:254-260.
- [235] Ponsioen CY, Reitsma JB, Boberg KM, Aabakken L, Rauws EA, Schruppf E. Validation of a cholangiographic prognostic model in primary sclerosing cholangitis. *Endoscopy* 2010;42:742-747.
- [236] Ruiz A, Lemoine S, Carrat F, Corpechot C, Chazouilleres O, Arrive L. Radiologic course of primary sclerosing cholangitis: assessment by three-dimensional magnetic resonance cholangiography and predictive features of progression. *Hepatology* 2014;59:242-250.

- [237] Lemoine S, Cazzagon N, El Mouhadi S, Trivedi PJ, Dohan A, Kemgang A, et al. Simple Magnetic Resonance Scores Associate With Outcomes of Patients With Primary Sclerosing Cholangitis. *Clin Gastroenterol Hepatol* 2019;17:2785-2792 e2783.
- [238] Cazzagon N, Lemoine S, El Mouhadi S, Trivedi PJ, Gaouar F, Kemgang A, et al. The Complementary Value of Magnetic Resonance Imaging and Vibration-Controlled Transient Elastography for Risk Stratification in Primary Sclerosing Cholangitis. *Am J Gastroenterol* 2019;114:1878-1885.
- [239] Tenca A, Mustonen H, Lind K, Lantto E, Kolho KL, Boyd S, et al. The role of magnetic resonance imaging and endoscopic retrograde cholangiography in the evaluation of disease activity and severity in primary sclerosing cholangitis. *Liver Int* 2018;38:2329-2339.
- [240] Schulze J, Lenzen H, Hinrichs JB, Ringe B, Manns MP, Wacker F, et al. An Imaging Biomarker for Assessing Hepatic Function in Patients With Primary Sclerosing Cholangitis. *Clin Gastroenterol Hepatol* 2019;17:192-199 e193.
- [241] O EA, Buchter M, H AB, Korth J, Canbay A, Gerken G, et al. Performance and Utility of Transient Elastography and Non-Invasive Markers of Liver Fibrosis in Patients with Autoimmune Hepatitis: A Single Centre Experience. *Hepat Mon* 2016;16:e40737.
- [242] Xu Q, Sheng L, Bao H, Chen X, Guo C, Li H, et al. Evaluation of transient elastography in assessing liver fibrosis in patients with autoimmune hepatitis. *J Gastroenterol Hepatol* 2017;32:639-644.
- [243] Guo L, Zheng L, Hu L, Zhou H, Yu L, Liang W. Transient Elastography (FibroScan) Performs Better Than Non-Invasive Markers in Assessing Liver Fibrosis and Cirrhosis in Autoimmune Hepatitis Patients. *Med Sci Monit* 2017;23:5106-5112.
- [244] Sheptulina A, Shirokova E, Nekrasova T, Blum H, Ivashkin V. Platelet count to spleen diameter ratio non-invasively identifies severe fibrosis and cirrhosis in patients with autoimmune hepatitis. *J Gastroenterol Hepatol* 2016;31:1956-1962.
- [245] Nishikawa H, Enomoto H, Iwata Y, Hasegawa K, Nakano C, Takata R, et al. Clinical significance of serum *Wisteria floribunda* agglutinin positive Mac-2-binding protein level and high-sensitivity C-reactive protein concentration in autoimmune hepatitis. *Hepatology research* 2016;46:613-621.
- [246] Wu S, Yang Z, Zhou J, Zeng N, He Z, Zhan S, et al. Systematic review: diagnostic accuracy of non-invasive tests for staging liver fibrosis in autoimmune hepatitis. *Hepatology international* 2019;13:91-101.
- [247] Dietrich CF, Bamber J, Berzigotti A, Bota S, Cantisani V, Castera L, et al. EFSUMB Guidelines and Recommendations on the Clinical Use of Liver Ultrasound Elastography, Update 2017 (Long Version). *Ultraschall Med* 2017;38:e16-e47.
- [248] Hartl J, Denzer U, Ehlken H, Zenouzi R, Peiseler M, Sebode M, et al. Transient elastography in autoimmune hepatitis: Timing determines the impact of inflammation and fibrosis. *J Hepatol* 2016;65:769-775.
- [249] Zeng J, Huang ZP, Zheng J, Wu T, Zheng RQ. Non-invasive assessment of liver fibrosis using two-dimensional shear wave elastography in patients with autoimmune liver diseases. *World J Gastroenterol* 2017;23:4839-4846.
- [250] Wang J, Malik N, Yin M, Smyrk TC, Czaja AJ, Ehman RL, et al. Magnetic resonance elastography is accurate in detecting advanced fibrosis in autoimmune hepatitis. *World J Gastroenterol* 2017;23:859-868.
- [251] Hartl J, Ehlken H, Sebode M, Peiseler M, Krech T, Zenouzi R, et al. Usefulness of biochemical remission and transient elastography in monitoring disease course in autoimmune hepatitis. *J Hepatol* 2018;68:754-763.
- [252] Ferraioli G, Wong VW, Castera L, Berzigotti A, Sporea I, Dietrich CF, et al. Liver Ultrasound Elastography: An Update to the World Federation for Ultrasound in Medicine and Biology Guidelines and Recommendations. *Ultrasound Med Biol* 2018;44:2419-2440.
- [253] Lin Y, Li H, Jin C, Wang H, Jiang B. The diagnostic accuracy of liver fibrosis in non-viral liver diseases using acoustic radiation force impulse elastography: A systematic review and meta-analysis. *PLoS one* 2020;15:e0227358.
- [254] Ferraioli G, De Silvestri A, Lissandrin R, Maiocchi L, Tinelli C, Filice C, et al. Evaluation of Inter-System Variability in Liver Stiffness Measurements. *Ultraschall in der Medizin* 2019;40:64-75.

- [255] Abraldes JG, Bureau C, Stefanescu H, Augustin S, Ney M, Blasco H, et al. Noninvasive tools and risk of clinically significant portal hypertension and varices in compensated cirrhosis: The "Anticipate" study. *Hepatology* 2016;64:2173-2184.
- [256] Piccinni R, Rodrigues SG, Montani M, Murgia G, Delgado MG, Casu S, et al. Controlled attenuation parameter reflects steatosis in compensated advanced chronic liver disease. *Liver Int* 2020;40:1151-1158.
- [257] Singh S, Venkatesh SK, Wang Z, Miller FH, Motosugi U, Low RN, et al. Diagnostic performance of magnetic resonance elastography in staging liver fibrosis: a systematic review and meta-analysis of individual participant data. *Clin Gastroenterol Hepatol* 2015;13:440-451 e446.
- [258] Kennedy P, Bane O, Hectors SJ, Fischman A, Schiano T, Lewis S, et al. Noninvasive imaging assessment of portal hypertension. *Abdominal radiology* 2020;45:3473-3495.
- [259] Smith AD, Branch CR, Zand K, Subramony C, Zhang H, Thaggard K, et al. Liver Surface Nodularity Quantification from Routine CT Images as a Biomarker for Detection and Evaluation of Cirrhosis. *Radiology* 2016;280:771-781.
- [260] Mathew RP, Venkatesh SK. Imaging of Hepatic Fibrosis. *Curr Gastroenterol Rep* 2018;20:45.
- [261] La Mura V, Reverter JC, Flores-Arroyo A, Raffa S, Reverter E, Seijo S, et al. Von Willebrand factor levels predict clinical outcome in patients with cirrhosis and portal hypertension. *Gut* 2011;60:1133-1138.
- [262] Ferlitsch M, Reiberger T, Hoke M, Salzl P, Schwengerer B, Ulbrich G, et al. von Willebrand factor as new noninvasive predictor of portal hypertension, decompensation and mortality in patients with liver cirrhosis. *Hepatology* 2012;56:1439-1447.
- [263] Sartoris R, Rautou PE, Elkrief L, Pollorsi G, Durand F, Valla D, et al. Quantification of Liver Surface Nodularity at CT: Utility for Detection of Portal Hypertension. *Radiology* 2018;289:698-707.
- [264] De Vos N, Sartoris R, Cauchy F, Rautou PE, Vilgrain V, Ronot M. Performance of liver surface nodularity quantification for the diagnosis of portal hypertension in patients with cirrhosis: comparison between MRI with hepatobiliary phase sequences and CT. *Abdominal radiology* 2020;45:365-372.
- [265] Palaniyappan N, Cox E, Bradley C, Scott R, Austin A, O'Neill R, et al. Non-invasive assessment of portal hypertension using quantitative magnetic resonance imaging. *J Hepatol* 2016;65:1131-1139.
- [266] You MW, Kim KW, Pyo J, Huh J, Kim HJ, Lee SJ, et al. A Meta-analysis for the Diagnostic Performance of Transient Elastography for Clinically Significant Portal Hypertension. *Ultrasound Med Biol* 2017;43:59-68.
- [267] Song J, Ma Z, Huang J, Luo Y, Zykus R, Kumar A, et al. Reliability of Transient Elastography-Based Liver Stiffness for Diagnosing Portal Hypertension in Patients with Alcoholic Liver Disease: A Diagnostic Meta-Analysis with Specific Cut-Off Values. *Ultraschall in der Medizin* 2020;41:60-68.
- [268] Kim BK, Han KH, Park JY, Ahn SH, Kim JK, Paik YH, et al. A liver stiffness measurement-based, noninvasive prediction model for high-risk esophageal varices in B-viral liver cirrhosis. *Am J Gastroenterol* 2010;105:1382-1390.
- [269] Berzigotti A, Seijo S, Arena U, Abraldes JG, Vizzutti F, Garcia-Pagan JC, et al. Elastography, spleen size, and platelet count identify portal hypertension in patients with compensated cirrhosis. *Gastroenterology* 2013;144:102-111 e101.
- [270] Pons M, Augustin S, Scheiner B, Guillaume M, Rosselli M, Rodrigues SG, et al. Noninvasive Diagnosis of Portal Hypertension in Patients With Compensated Advanced Chronic Liver Disease. *Am J Gastroenterol* 2020.
- [271] Berzigotti A. Non-invasive evaluation of portal hypertension using ultrasound elastography. *J Hepatol* 2017;67:399-411.
- [272] Suh CH, Kim KW, Park SH, Lee SS, Kim HS, Tirumani SH, et al. Shear Wave Elastography as a Quantitative Biomarker of Clinically Significant Portal Hypertension: A Systematic Review and Meta-Analysis. *AJR Am J Roentgenol* 2018;210:W185-W195.
- [273] Jansen C, Bogs C, Verlinden W, Thiele M, Möller P, Görtzen J, et al. Algorithm to rule out clinically significant portal hypertension combining Shear-wave elastography of liver and spleen: a prospective multicentre study. *Gut* 2016;65:1057-1058.
- [274] Jansen C, Bogs C, Verlinden W, Thiele M, Möller P, Görtzen J, et al. Shear-wave elastography of the liver and spleen identifies clinically significant portal hypertension: A prospective multicentre study. *Liver Int* 2017;37:396-405.

- [275] Cescon M, Colecchia A, Cucchetti A, Peri E, Montrone L, Ercolani G, et al. Value of transient elastography measured with FibroScan in predicting the outcome of hepatic resection for hepatocellular carcinoma. *Ann Surg* 2012;256:706-712; discussion 712-703.
- [276] Kim HY, So YH, Kim W, Ahn DW, Jung YJ, Woo H, et al. Non-invasive response prediction in prophylactic carvedilol therapy for cirrhotic patients with esophageal varices. *J Hepatol* 2019;70:412-422.
- [277] Marasco G, Dajti E, Ravaioli F, Alemanni LV, Capuano F, Gjini K, et al. Spleen stiffness measurement for assessing the response to beta-blockers therapy for high-risk esophageal varices patients. *Hepatology international* 2020;14:850-857.
- [278] Qi X, Berzigotti A, Cardenas A, Sarin SK. Emerging non-invasive approaches for diagnosis and monitoring of portal hypertension. *Lancet Gastroenterol Hepatol* 2018;3:708-719.
- [279] Giannini E, Botta F, Borro P, Risso D, Romagnoli P, Fasoli A, et al. Platelet count/spleen diameter ratio: proposal and validation of a non-invasive parameter to predict the presence of oesophageal varices in patients with liver cirrhosis. *Gut* 2003;52:1200-1205.
- [280] Marot A, Trépo 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.
- [281] Szakacs Z, Eross B, Soos A, Matrai P, Szabo I, Petervari E, et al. Baveno Criteria Safely Identify Patients With Compensated Advanced Chronic Liver Disease Who Can Avoid Variceal Screening Endoscopy: A Diagnostic Test Accuracy Meta-Analysis. *Front Physiol* 2019;10:1028.
- [282] Merchante N, Saroli Palumbo C, Mazzola G, Pineda JA, Tellez F, Rivero-Juarez A, et al. Prediction of Esophageal Varices by Liver Stiffness and Platelets in Persons with HIV infection and Compensated Advanced Chronic Liver Disease. *Clin Infect Dis* 2019.
- [283] 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.
- [284] 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.
- [285] Petta S, Sebastiani G, Bugianesi E, Viganò M, Wong VW, Berzigotti A, et al. Non-invasive prediction of esophageal varices by stiffness and platelet in non-alcoholic fatty liver disease cirrhosis. *J Hepatol* 2018;69:878-885.
- [286] Colecchia A, Ravaioli F, Marasco G, Colli A, Dajti E, Di Biase AR, et al. A combined model based on spleen stiffness measurement and Baveno VI criteria to rule out high-risk varices in advanced chronic liver disease. *J Hepatol* 2018;69:308-317.
- [287] Stefanescu H, Marasco G, Cales P, Fraquelli M, Rosselli M, Ganne-Carrie N, et al. A novel spleen-dedicated stiffness measurement by FibroScan(R) improves the screening of high-risk oesophageal varices. *Liver Int* 2020;40:175-185.
- [288] Wong GL, Liang LY, Kwok R, Hui AJ, Tse YK, Chan HL, et al. Low Risk of Variceal Bleeding in Patients With Cirrhosis After Variceal Screening Stratified by Liver/Spleen Stiffness. *Hepatology* 2019;70:971-981.
- [289] Manatsathit W, Samant H, Kapur S, Ingviya T, Esmadi M, Wijarnpreecha K, et al. Accuracy of liver stiffness, spleen stiffness, and LS-spleen diameter to platelet ratio score in detection of esophageal varices: Systemic review and meta-analysis. *J Gastroenterol Hepatol* 2018;33:1696-1706.
- [290] Calvaruso V, Cacciola I, Licata A, Madonia S, Benigno R, Petta S, et al. Is Transient Elastography Needed for Noninvasive Assessment of High-Risk Varices? The REAL Experience. *Am J Gastroenterol* 2019;114:1275-1282.
- [291] Jangouk P, Turco L, De Oliveira A, Schepis F, Villa E, Garcia-Tsao G. Validating, deconstructing and refining Baveno criteria for ruling out high-risk varices in patients with compensated cirrhosis. *Liver Int* 2017;37:1177-1183.
- [292] Tseng YJ, Zeng XQ, Chen J, Li N, Xu PJ, Chen SY. Computed tomography in evaluating gastroesophageal varices in patients with portal hypertension: A meta-analysis. *Dig Liver Dis* 2016;48:695-702.

- [293] Mandorfer M, Schwabl P, Paternostro R, Pomej K, Bauer D, Thaler J, et al. Von Willebrand factor indicates bacterial translocation, inflammation, and procoagulant imbalance and predicts complications independently of portal hypertension severity. *Aliment Pharmacol Ther* 2018;47:980-988.
- [294] Singh S, Fujii LL, Murad MH, Wang Z, Asrani SK, Ehman RL, et al. Liver stiffness is associated with risk of decompensation, liver cancer, and death in patients with chronic liver diseases: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2013;11:1573-1584 e1571-1572; quiz e1588-1579.
- [295] Wang J, Li J, Zhou Q, Zhang D, Bi Q, Wu Y, et al. Liver stiffness measurement predicted liver-related events and all-cause mortality: A systematic review and nonlinear dose-response meta-analysis. *Hepatol Commun* 2018;2:467-476.
- [296] Robic MA, Procopet B, Métivier S, Péron JM, Selves J, Vinel JP, et al. Liver stiffness accurately predicts portal hypertension related complications in patients with chronic liver disease: a prospective study. *J Hepatol* 2011;55:1017-1024.
- [297] Colecchia A, Colli A, Casazza G, Mandolesi D, Schiumerini R, Reggiani LB, et al. Spleen stiffness measurement can predict clinical complications in compensated HCV-related cirrhosis: a prospective study. *J Hepatol* 2014;60:1158-1164.
- [298] Margini C, Murgia G, Stirnimann G, De Gottardi A, Semmo N, Casu S, et al. Prognostic Significance of Controlled Attenuation Parameter in Patients With Compensated Advanced Chronic Liver Disease. *Hepatol Commun* 2018;2:929-940.
- [299] Mendoza Y, Cocciolillo S, Murgia G, Chen T, Margini C, Sebastiani G, et al. Noninvasive Markers of Portal Hypertension Detect Decompensation in Overweight or Obese Patients With Compensated Advanced Chronic Liver Disease. *Clin Gastroenterol Hepatol* 2020.
- [300] Smith AD, Zand KA, Florez E, Sirous R, Shlapak D, Souza F, et al. Liver Surface Nodularity Score Allows Prediction of Cirrhosis Decompensation and Death. *Radiology* 2017;283:711-722.
- [301] Simon-Talero M, Roccarina D, Martinez J, Lampichler K, Baiges A, Low G, et al. Association Between Portosystemic Shunts and Increased Complications and Mortality in Patients With Cirrhosis. *Gastroenterology* 2018;154:1694-1705 e1694.
- [302] Praktiknjo M, Simon-Talero M, Romer J, Roccarina D, Martinez J, Lampichler K, et al. Total area of spontaneous portosystemic shunts independently predicts hepatic encephalopathy and mortality in liver cirrhosis. *J Hepatol* 2020;72:1140-1150.
- [303] Voulgaris T, Papatheodoridi M, Lampertico P, Papatheodoridis GV. Clinical utility of hepatocellular carcinoma risk scores in chronic hepatitis B. *Liver Int* 2020;40:484-495.
- [304] Jung KS, Kim SU, Ahn SH, Park YN, Kim DY, Park JY, et al. Risk assessment of hepatitis B virus-related hepatocellular carcinoma development using liver stiffness measurement (FibroScan). *Hepatology* 2011;53:885-894.
- [305] Wong GL, Chan HL, Wong CK, Leung C, Chan CY, Ho PP, et al. Liver stiffness-based optimization of hepatocellular carcinoma risk score in patients with chronic hepatitis B. *J Hepatol* 2014;60:339-345.
- [306] Kim BS, Seo YS, Kim YS, Lee CH, Lee HA, Um SH, et al. Reduced risk of hepatocellular carcinoma by achieving a subcirrhotic liver stiffness through antiviral agents in hepatitis B virus-related advanced fibrosis or cirrhosis. *J Gastroenterol Hepatol* 2018;33:503-510.
- [307] Kim BK, Oh HJ, Park JY, Kim DY, Ahn SH, Han KH, et al. Early on-treatment change in liver stiffness predicts development of liver-related events in chronic hepatitis B patients receiving antiviral therapy. *Liver Int* 2013;33:180-189.

Table 1. Common measures for evaluating the diagnostic accuracy of non-invasive fibrosis tests.

Sensitivity	Probability that a patient with the condition (e.g. advanced fibrosis) tests positive
Specificity	Probability that a patient without the condition tests negative
Positive predictive value	Probability that a patient who tests positive has the condition
Negative predictive value	Probability that a patient who tests negative does not have the condition
Area under the receiver operating curve (AUROC)	The diagnostic ability of a binary classifier at a specific cut-off, i.e. the probability that this classifier will correctly rank a randomly chosen person with the disease higher than a randomly chosen person without the disease
Positive likelihood ratio	How many times more likely positive index test results are in the diseased group compared to the non-diseased group. Estimated as $\text{sens}/(1-\text{spec})$
Negative likelihood ratio	How many times less likely negative index test results are in the diseased group compared to the non-diseased group. Estimated as $(1-\text{sens})/\text{spec}$

Table 2. Advantages and disadvantages of the main non-invasive tests used to diagnose and stage liver fibrosis.

	Serum markers		Transient elastography	pSWE	2DSWE	MRE
Advantages	Non-Patented • Good reproducibility • High applicability (95%) • No cost and wide availability • Well validated • Can be performed in the outpatient clinic • Prognostic value of some has been validated for some etiologies of chronic liver disease on population level	Patented Good reproducibility • High applicability (95%) • Well validated • Can be performed in the outpatient clinic • Prognostic value of some has been validated for some etiologies of chronic liver disease	• Most widely used and validated technique • Point-of-care (bedside; rapid, easy to learn) • Quality criteria well defined • Good reproducibility • High performance for cirrhosis (AUROC >0.9) • Prognostic value in compensated cirrhosis well validated	• Can be performed in combination with regular US if the device is provided with adequate software • ROI smaller than TE and location chosen by the operator • Higher applicability than TE (ascites and obesity) • Performance equivalent to that of TE for advanced fibrosis and cirrhosis • Prognostic value in cirrhosis • High applicability for spleen stiffness measurement	• Can be performed in combination with regular US if the device is provided with adequate software • Large ROI that can be adjusted in size and location chosen by the operator • Measures liver stiffness in real-time • Good applicability • High performance for the diagnosis of significant fibrosis and cirrhosis • Prognostic value in compensated cirrhosis	• Can be implemented on a regular MRI machine • Examination of the whole liver • Higher applicability than TE (ascites and obesity) • High performance for the earlier fibrosis stage and for diagnosis of cirrhosis

<p>Disadvantages</p>	<ul style="list-style-type: none"> • Non liver-specific • Performance not as good as TE and patented serum markers • False positive results with FIB-4 and NFS in case of age > 65 yrs 	<ul style="list-style-type: none"> • Cost • Non liver-specific • Performance not as good as TE for cirrhosis • False positive results in case of extrahepatic inflammatory conditions, profibrotic, extrahepatic disease and other (e.g. hemolysis, Gilbert syndrome) 	<ul style="list-style-type: none"> • Requires a dedicated device • ROI cannot be chosen • Applicability (>95%) lower than serum biomarker: (obesity, ascites, operator experience) • False positive in case of acute hepatitis, extrahepatic cholestasis, liver congestion, food intake and excessive alcohol intake 	<ul style="list-style-type: none"> • False positive in case of acute hepatitis, extrahepatic cholestasis, liver congestion, food intake and excessive alcohol intake 	<ul style="list-style-type: none"> • False positive in case of acute hepatitis, extrahepatic cholestasis, liver congestion, food intake and excessive alcohol intake 	<ul style="list-style-type: none"> • Not applicable in case of iron overload • Requires a MRI facility • Time-consuming • Costly • No clear data on prognostic value
-----------------------------	--	---	---	---	---	---

Table 3. This table shows that for the same value of specificity and sensitivity, the negative predictive value decreases and the positive predictive value increases with increasing prevalence of advanced fibrosis.

Prevalence of advanced fibrosis	Sensitivity	Specificity	Positive predictive value	Negative predictive value
10%	80%	80%	31%	97%
20%	80%	80%	50%	94%
30%	80%	80%	63%	90%
40%	80%	80%	73%	86%
50%	80%	80%	80%	80%

Table 4. Combination of tests used to assess the risk of CSPH and varices in cirrhosis.

Test	Formula	Suggested cut-off	Sensitivity and Specificity in cACLD
LSPS [268] [269]	LS by TE × (spleen size in mm/platelet count in G/L)	1.08 to exclude CSPH 2.06 to diagnose CSPH 3.21 to rule-out/rule-in varices (any size)	Se 90%, Spec 91% Se 92%, Spec 90% Se 81%, Spec 86%
PH Risk score [269]	$5.953 + 0.188 \times \text{LS} + 1.583 \times \text{sex}$ (1: male; 0: female) + $26.705 \times \text{spleen diameter in mm/platelet count in G/L}$	0.06 to exclude CSPH 0.82 to diagnose CSPH	Se 90%, Spec 91% Se 93%, Spec 90%
Platelet to spleen ratio [279]	(platelet count in G/L) / (maximum spleen bipolar diameter in mm by ultrasound)	909 to rule-out/rule-in varices (any size)	Se 100%, Spec 71%

Appendix. Delphi Round Agreement on the Statements and Recommendations of the present CPGs.

Statement / Recommendation	Delphi Panel agreement
Non-invasive fibrosis tests should be used for ruling out rather than diagnosing advanced fibrosis in low prevalence populations (LoE 1, Strong recommendation).	100%
Non-invasive fibrosis tests should be preferentially used in patients at risk of advanced liver fibrosis (such as patients with metabolic risk factors and/or harmful use of alcohol) and not in unselected general populations (LoE 2, Strong recommendation).	100%
ALT, AST and platelet count should be part of the routine investigations in primary care in patients with suspected liver disease, so that simple non-invasive scores can be readily calculated (LoE 2, Strong recommendation).	95%
The automatic calculation and systematic reporting of simple non-invasive fibrosis tests such as FIB-4, in populations at risk of liver fibrosis (individuals with metabolic risk factors and/or harmful use of alcohol) in primary care, is recommended in order to improve risk stratification and linkage to care (LoE 2, Strong recommendation).	100%
Non-invasive scores, serum markers, liver stiffness and imaging methods can identify advanced fibrosis in patients at risk from low-prevalence populations significantly better than clinical acumen alone (LoE 1).	89%
Individuals at risk for advanced fibrosis due to metabolic risk factors and/or harmful use of alcohol should be entered in appropriate risk stratification pathways using non-invasive fibrosis tests (LoE 1, Strong recommendation).	100%
The selection of non-invasive tests and the design of diagnostic pathways for testing low prevalence populations for advanced fibrosis should be performed in consultation with a liver specialist (LoE 3, Strong recommendation)	89%
<p>In patients with alcohol-related liver disease LSM by TE <8 kPa is recommended to rule-out advanced fibrosis in clinical practice, with the following NITs as alternatives, if TE is not available (LoE 3; strong recommendation)</p> <ul style="list-style-type: none"> - Patented tests: ELF™ <9.8 or FibroMeter™ <0.45 or FibroTest® < 0.48 - Non-patented tests: FIB-4 <1.3 	89%
Upon referral of patients at risk of alcohol-related liver disease, LSM by TE ≥12-15 kPa is recommended to rule in advanced fibrosis, after considering causes of false positives (LoE 2; strong recommendation).	89%

In patients with elevated liver stiffness and biochemical evidence of hepatic inflammation (AST or GGT >2xULN), LSM by TE should be repeated after at least one week of alcohol abstinence or reduced drinking (LoE 3; strong recommendation).	95%
Non-invasive scores and LSM by TE and other elastography methods are not accurate in detecting fibrosis regression after SVR in HCV patients diagnosed with cACLD prior to antiviral therapy (LoE 3)	89%
The routine use of non-invasive scores and LSM by TE and other elastography methods is currently not recommended to detect fibrosis regression after SVR in HCV patients (LoE 3; strong recommendation).	89%
Cut-offs of LSM by TE used in patients with untreated HCV should not be used to stage liver fibrosis after SVR (LoE 4; strong recommendation).	95%
In patients with cACLD previous to antiviral therapy for HCV, LSM post-SVR could be helpful to refine the stratification of residual liver-related risk, and awaiting further confirmatory data yearly repetition of LSM can be done (LoE 3).	100%
Patients with cACLD previous to antiviral therapy for HCV should continue to be monitored for HCC and portal hypertension irrespective of the results of NITs post SVR (LoE 3; strong recommendation).	79%
Non-invasive scores are not recommended for the diagnosis of steatosis in clinical practice (LoE 2; strong recommendation)	84%
Conventional ultrasound is recommended as first line tool for the diagnosis of steatosis in clinical practice, despite its well-known limitations (LoE 1; strong recommendation)	100%
MRI-PDFF is the most accurate non-invasive method for detecting and quantifying steatosis. However, it is not recommended as first line tool given its cost and limited availability. Therefore, it is more suited for clinical trials (LoE 2; strong recommendation)	100%
CAP is a promising point-of-care technique for rapid and standardized detection of steatosis. However, given its limited availability and lack of head-to-head studies compared to ultrasound, CAP cannot yet be recommended as first line technique (LoE 2).	89%
Although there are no consensual cut-offs, values above 275 dB/m might be used to diagnose steatosis, since they showed over 90% sensitivity to detect steatosis (LoE 2).	79%
In patients with NAFLD: Liver biopsy remains the reference standard for the diagnosis of NASH, because none of the available non-invasive tests has acceptable accuracy (LoE 2)	89%
In patients with NAFLD: The following NITs are recommended to rule-out advanced fibrosis in clinical practice (LoE 1, strong):	94%

<ul style="list-style-type: none"> - - LSM by TE <8 kPa - - Patented tests: ELFTM <9.8 or FibroMeterTM <0.45 or FibroTest® < 0.48 - - Non-patented tests: FIB-4 <1.3 or NFS <-1.455 	
Upon referral of a patient with FIB-4 over 1.3, the use of transient elastography and/or patented serum tests should be used to rule out/in advanced fibrosis (see Figure 1) (LoE 2, strong recommendation).	94%
MRE is the most accurate non-invasive method for staging liver fibrosis. However, it is only marginally better than other NITs for F3-F4 fibrosis and it is not recommended as first line NIT given its cost and limited availability (LoE 2; strong recommendation). Therefore, it is more suited for clinical trials	89%
Serum scores (APRI, FIB-4, NFS, ELF TM) and LSM by TE should be used to stratify the risk of liver-related outcomes in NAFLD (LoE 3; strong recommendation).	89%
Repeated measurements of NITs can be used to refine stratification of risk of liver-related events in NAFLD/NASH patients. Despite the lack of evidence regarding the optimal timeframe between subsequent LSM assessment, it seems reasonable to repeat NITs every 3 years in patients with early stage and every year in patients with advanced stage (LoE 3; weak recommendation)	94%
Liver biopsy remains the reference for patient selection in phase 2b and phase 3 therapeutic trials and should be used for these purposes (LoE 1; strong recommendation).	95%
MRI-PDFF can be used to assess steatosis evolution under treatment (LoE 2; weak recommendation). However, the minimal decrease in MRI-PDFF that defines a clinically relevant change or treatment response needs to be better defined.	95%
Liver biopsy remains the reference to evaluate NASH resolution and liver fibrosis improvement and should be used for these purposes (LoE 2; strong recommendation).	84%
In patients with PBC, serum markers of fibrosis and non-invasive scores (combination of clinical and laboratory variables) are not recommended for fibrosis staging in clinical practice (LoE 3; strong recommendation).	79%
In patients with PBC, LSM by TE is the best surrogate marker for ruling-in severe fibrosis/cACLD and should be used for this purpose using a cut-off of 10 kPa (LoE 3; strong recommendation).	89%
In patients with PSC, LSM by TE above 9.5 kPa can be used to support the diagnosis of advanced fibrosis in compensated patients with normal bilirubin and without high-grade stenosis (LoE 3; weak recommendation).	89%
In patients with PBC, non-invasive discrimination of early and advanced stage based on biochemical parameters (normal vs. abnormal albumin and bilirubin) and LSM by TE < or > 10 kPa is recommended at baseline (LoE 3, strong recommendation).	94%
During treatment, risk stratification should be based on the assessment of response to therapy by using continuous (GLOBE and	89%

UK-PBC risk scores) and/or qualitative criteria (Paris II, Toronto, Rotterdam, Barcelona, Paris I) of response and LSM by TE (LoE 3, strong recommendation).	
In patients with PSC, both the ELF™ score and LSM by TE correlate with outcomes and they should be used for risk stratification both at baseline and during follow-up (LoE 3, strong recommendation).	89%
LSM by TE can be used in patients with treated AIH to monitor the disease course together with transaminases and IgG, and to stage liver fibrosis after at least 6 months of immunosuppressive therapy (LoE 3, weak recommendation).	84%
cACLD should be diagnosed by using second line tests (patented serum tests or elastography) in a specialized setting (LoE 2, strong recommendation)	100%
Fibrotest® or FibroMeter™ or ELF™ should be used to rule out cACLD if available (LoE 3, strong recommendation)	83%
LSM by TE should be used to rule-out and diagnose cACLD using the following cut-offs: < 8-10 kPa to rule-out; > 12-15 kPa to rule-in. Intermediate values require further testing (LoE 3 strong recommendation)	89%
pSWE and 2D-SWE should be used to rule-out and diagnose cACLD, with AUROCs >0.90 in the published meta-analyses (LoE 2, strong recommendation)	100%
Inter-system variability should be taken into account when interpreting the results of different elastography techniques, since values, ranges and cut-offs are not comparable (LoE 3, strong recommendation)	89%
LSM by TE at a cut-off of >20-25 kPa should be used to diagnose CSPH in patients with cACLD (LoE 1, strong recommendation).	89%
Platelet count, spleen size and spleen stiffness should be used as additional NITs to further improve risk stratification for CSPH (LoE 3, strong recommendation).	89%
The presence of porto-systemic collaterals on ultrasound, CT or MR is a sign of CSPH in patients with cACLD and should be routinely reported (LoE 2, strong recommendation).	100%
For an exact assessment of the severity of portal hypertension in cACLD beyond presence and absence of CSPH and for assessment of the hemodynamic response to treatment, HVPg remains the only validated tool and should not be substituted by NITs (LoE 1, strong recommendation).	84%
In patients with cACLD due to untreated viral hepatitis, HIV-HCV coinfection, alcohol, NAFLD, PBC and PSC, the finding of LSM by TE < 20 kPa and platelet count >150 G/L (Baveno VI criteria) is a validated tool to rule-out high risk varices and avoid endoscopic screening. These criteria should be used whenever TE is available (LoE 1a; strong recommendation)	100%
Spleen stiffness can be used as an additional tool to refine the risk of high-risk varices in cACLD (LoE 2; weak recommendation).	89%

CT should not be used for primary screening for esophageal and gastric varices, but when doing a routine CT, varices should be looked for and reported (LoE 3, strong recommendation).	100%
In patients with cACLD, liver stiffness at diagnosis should be used in addition to liver function tests to stratify the risk of clinical decompensation and mortality (LoE 1, strong recommendation).	100%
Annual repeated measurements of liver stiffness can be used to refine risk stratification in patients with cACLD (LoE 5, weak recommendation)	89%
Liver stiffness can be used in addition to clinical variables and accepted risk scores to stratify the risk of hepatocellular carcinoma in patients with cACLD due to HBV (LoE 3, weak recommendation)	89%

Supplementary Material

Recommendations that remained unchanged as in the previous EASL CPGs

General statements

- Even though liver biopsy has been used as the reference method for the design, evaluation and validation of non-invasive tests, it is an imperfect gold standard. In order to optimize the value of liver biopsy for fibrosis evaluation, it is important to adhere to the following recommendations: (i) sample length >15 mm by a 16G needle; (ii) use of appropriate scoring systems according to liver disease etiology; and (iii) reading by an experienced (and if possible specialized) pathologist.
- Non-invasive tests reduce but do not abolish the need for liver biopsy; they should be used as an integrated system with liver biopsy according to the context.

Currently available non-invasive methods

- Non-invasive tests should always be interpreted by specialists in liver disease, according to the clinical context, considering the results of other tests (biochemical, radiological and endoscopic) and taking into account the recommended quality criteria for each test and its possible pitfalls (A1)
- Serum biomarkers can be used in clinical practice due to their high applicability (>95%) and good interlaboratory reproducibility. However, they should be preferably obtained in fasting patients (particularly those including hyaluronic acid) and following the manufacturer's recommendations for the patented tests (A1)
- TE is a fast, simple, safe and easy to learn procedure that is widely available. Its main limitation is the impossibility of obtaining results in case of ascites or morbid obesity and its limited applicability in case of obesity and limited operator experience (A1)
- TE should be performed by an experienced operator (>100 examinations) following a standardized protocol with the patient, fasting for at least 2 hours, in the supine position, right arm in full abduction, on the midaxillary line with the probe-tip placed in the 9th to 11th intercostal space with a minimum of 10 shots (A1)
- Correct interpretation of TE results in clinical practice must consider the following parameters:
 - IQR/ median value (<30%),
 - Serum aminotransferases levels (<5 x ULN),
 - BMI (use XL probe above 30 kg/m² or if skin-to-capsule distance is >25 mm),
 - Absence of extra-hepatic cholestasis
 - Absence of right heart failure, or other causes of congestive liver
 - Absence of ongoing excessive alcohol intake (A1)

- Although alternative techniques, such as pSWE/ARFI or 2D-SWE seem to overcome limitations of TE, their quality criteria for correct interpretation are not yet well defined (A1)
- At present correct interpretation of pSWE/ARFI results in clinical practice should systematically take into account the potentially confounding parameter:

fasting for at least 2 hours, transaminases levels (<5 x ULN), absence of extra-hepatic cholestasis and absence of right heart failure (A1)

Endpoints for staging liver fibrosis

- Detection of cACLD/cirrhosis represents the most relevant clinical endpoint in patients with chronic liver disease of any etiology (A1)
- Detection of cirrhosis indicates that patients should be monitored for complications related to PH and regularly screened for HCC (A1)

Performance of TE for staging liver fibrosis

- TE can be considered the non-invasive standard for the measurement of LS (A1)
- TE is well validated in untreated viral hepatitis with performance equivalent in hepatitis B and C and in HIV-HCV coinfection (A1)
- TE performs better for detection of cirrhosis than for detection of significant fibrosis (A1)
- TE is a reliable method for the diagnosis of cirrhosis in patients with chronic liver diseases, that generally performs better at ruling out than ruling in cirrhosis (with negative predictive value higher than 90%) (A1)

Performance of other ultrasound elastography techniques for staging liver fibrosis

- pSWE/ARFI performs better for detecting cirrhosis than significant fibrosis and is better validated in chronic hepatitis C than for hepatitis B, HIV-HCV coinfection, NAFLD and other liver diseases (A1)
- pSWE/ARFI shows equivalent performance to TE for detecting significant fibrosis and cirrhosis (A1)

Comparison of performance of TE and serum biomarkers for staging liver fibrosis

- TE and serum biomarkers have equivalent performance for detecting significant fibrosis in patients with untreated viral hepatitis (A1)
- TE is the most accurate non-invasive method for detecting cirrhosis in patients with untreated viral hepatitis (A1)

Algorithms combining different tests (LS and/or serum biomarkers)

- Among the different available strategies, algorithms combining TE and serum biomarkers appear to be the most attractive and validated one (A2)

- In patients with viral hepatitis C, when TE and serum biomarkers results are in accordance, the diagnostic accuracy is increased for detecting significant fibrosis but not for cirrhosis. In cases of unexplained discordance, a liver biopsy should be performed if the results would change the patient management. (A1)

Indications for non-invasive tests for staging liver disease in untreated viral hepatitis

HCV including HIV-HCV

- All HCV patients should be screened to exclude cirrhosis by TE if available. Serum biomarkers can be used in the absence of TE (A1)
- HCV patients who were diagnosed with cirrhosis based on non-invasive diagnosis should undergo screening for HCC and PH and do not need confirmatory liver biopsy (A1)

HBV

- TE has better prediction for advanced liver fibrosis and cirrhosis than serum biomarkers in chronic hepatitis B(B1)
- TE is best used to determine liver fibrosis in hepatitis B patients with active viraemia (HBV DNA >2000 IU/ml)but normal ALT (A1)
- TE can be used to exclude severe fibrosis and cirrhosis in inactive carriers (HBeAg-negative, low viral load (HBV DNA <2000 IU/ml) and normal ALT). Liver biopsy should only be considered in doubtful cases after TE(A1)
- LS measurement should be interpreted with caution among patients with elevated ALT, and should not be used in patients with very high ALT levels (>10 x ULN)(A1)

Use of non-invasive methods when deciding for treatment in viral hepatitis B

- For the diagnosis of significant fibrosis a combination of tests with concordance may provide the highest diagnostic accuracy (A2)
- Non-invasive tests should be utilized prior to therapy by treating non-specialists to make sure that patients with severe fibrosis/cirrhosis are referred for appropriate disease specific specialist evaluation (A1)
- Non-invasive assessment of liver fibrosis, using either serum biomarkers or TE, should be considered for patients with significant viraemia (HBV DNA >2000 IU/ml) when liver cirrhosis is suspected (A1)
- Among patients who have HBV DNA >2000 IU/ml, antiviral therapy should be considered for patients who have advanced fibrosis or cirrhosis as determined by non-invasive assessment of liver fibrosis, either by serum biomarkers or TE, regardless of the ALT levels(A1)