Non-alcoholic fatty liver disease (NAFLD) affects one in every three subjects in the occidental world. The vast majority will not progress, but a relevant minority will develop liver cirrhosis and its complications. The classical gold standard for diagnosing and staging NAFLD and assessing fibrosis is liver biopsy (LB). However, it has important sample error issues and subjectivity in the interpretation, apart from a small but real risk of complications. The decision to perform an LB is even harder in a condition so prevalent such as NAFLD, in which the probability of finding severe liver injury is low. In an attempt to overcome LB and to subcategorize patients with NAFLD in different prognoses allowing better management decisions, several non-invasive methods have been studied in the last decade. The literature is vast and confusing. This review will summarize which methods have been tested and how they perform, which tests are adequate for clinical practice and how they can change the management of these patients.

Keywords: Non-alcoholic fatty liver disease; Non-alcoholic steatohepatitis; Fibrosis; Non-invasive diagnosis.

Received 31 August 2012; received in revised form 4 November 2012; accepted 7 November 2012

* Corresponding author. Address: Serviço de Gastroenterologia, Hospital de Santa Maria, Av. Prof. Egas Moniz, 1649-035 Lisboa, Portugal. Tel.: +351 217985187; fax: +351 217985142.
E-mail address: hlcortezpinto@netcabo.pt (H. Cortez-Pinto).

Abbreviations: NAFLD, non-alcoholic fatty liver disease; NASH, non-alcoholic steatohepatitis; T2DM, type 2 diabetes mellitus; LB, liver biopsy; Se, sensitivity; Sp, specificity; ROC, receiver operating characteristics; AUROC, area under ROC; NPV, negative predictive value; PPV, positive predictive value; FLI, fatty liver index; US, ultrasonography; BMI, body mass index; GGT, -glutamyltranspeptidase; LAP, lipid accumulation product; MRS, magnetic resonance spectroscopy; AST, aspartate aminotransferase; ALT, alanine aminotransferase; US-FLI, ultrasound fatty liver indicator; CT, computed tomography; CAP, controlled attenuation parameter; BMI, body mass index; CK18, cytokeratin 18; CBT, C-caffeine breath test; HA, hyaluronic acid; TIMP-1, tissue inhibitor of metalloproteinases-1; ARFI, acoustic radiation force impulse.
Review

high Sp. The most common approach to evaluate a test has been the analysis of the Receiving Operating Characteristic (ROC) curves and the area under ROC (AUROC), which evaluates the probability of a test discriminating a true positive (Se) against the probability of finding a false positive result (1-Sp). When the AUROC is higher than 0.8, it suggests good accuracy. It is a valuable tool but must be analysed carefully, particularly when comparing different tests. Although Se and Sp are invariant for a diagnostic test, they may depend on the characteristics of the population, such as age, gender and severity of disease. Thus, it may not be accurate to compare AUROC of different tests in different studies, with different populations and no statistical work done comparing them.

Identification and quantification of hepatic steatosis

The first challenge is when to suspect NAFLD. Suspicion will not be driven by clinical manifestations, since most patients are asymptomatic. Symptomatic patients present unspecific complaints such as fatigue, abdominal discomfort and, only seldom, manifestations of advanced liver disease. There are, however, high-risk populations in whom the prevalence is so high that per se is enough to raise the hypothesis of NAFLD. In fact, up to two thirds of patients with obesity and T2DM, present with hepatic steatosis [8]. Also, liver tests, namely aminotransferases, are usually normal, and when increased, typically present mild elevation with a fluctuant pattern [9]. Isolated increase in alkaline phosphatase is not frequent, but it has been reported in 10% of patients referred to tertiary care [10].

We should ask whether it is worth searching and diagnosing NAFLD in asymptomatic subjects with normal liver tests, since the majority will have non-progressive simple steatosis. Then again, hepatic steatosis is linked to an increase in cardiovascular risk and death. Particularly in diabetic patients, steatosis increases by more than 3-fold the risk for overall death [11] and cardiovascular disease [12]. Even in these high-risk patients, there is controversy regarding screening, among endocrinologists. Some do not recommend it, advocating that traditional scores should assess cardiovascular risk, while others consider diagnosis and evaluation of NAFLD as part of the management of DM being an indication for more intense monitoring and therapeutic intervention [13]. The guidelines on NAFLD from the American Association for the Study of Liver Diseases discourage screening for hepatic steatosis, even in high-risk patients such as diabetics. There is not enough evidence regarding diagnostic tools and treatment options, and there are no studies on the cost-effectiveness of a screening program [14]. There are however, situations in which an active search for NAFLD in asymptomatic subjects is warranted: liver donors for liver transplant, as steatosis is a risk for graft primary non-function, and in major hepatic resection [15], in which steatosis increases the risk for post-operative morbi-mortality [16].

The second question, if it is meaningful to quantify steatosis, is difficult to answer, since it has not been consistently demonstrated that the amount of fat influences prognosis. Also, it is yet to be elucidated if decreasing the amount of hepatic fat with therapeutic interventions will favourably affect the cardio-metabolic risk and the risk for progression to advanced liver disease.

Several diagnostic panels have been proposed to predict steatosis (Table 1). Steatotest incorporates 12 variables in an undisclosed formula, including α2-macroglobulin, haptoglobin, and apolipoprotein A1 [17]. In a French cohort of more than 700 patients, it showed reasonable accuracy, with 0.79 AUROC for moderate-severe steatosis, good negative predictive value (NPV), 93%, but small positive predictive value (PPV), 63%.

Another French group found similar results in 288 morbid obese subjects [18]. The Poynard’ group conducted a meta-analysis in morbid obese subjects, obtaining the same conclusions [19]. We have to acknowledge that AUROC was suboptimal, it has been validated only in French cohorts, it incorporates tests not used routinely and because the formula is not public, a fee is imposed for each test applied.

Bedogni et al. first proposed Fatty Liver Index (FLI) in 2006, as an algorithm derived from the population of the Dionysos Nutrition & Liver Study [20]. The gold standard was ultrasonography (US). The index varies between 0 and 100 and is calculated through a formula incorporating: body mass index (BMI), waist circumference, triglycerides and γ-glutamyltranspeptidase (GGT). It showed good accuracy in detecting NAFLD and it has been used in several population studies [21,22]. However, the gold standard used is far from ideal, and as such, the results should be interpreted carefully. Its main indication is for epidemiological studies, in an attempt to avoid US. Recently, a study on 2075 middle-aged Caucasians from the Regional Health Registry, followed for 15 years, showed that FLI independently associates with overall, cardiovascular and cancer-related mortality [23].

The same group also proposed Lipid Accumulation Product (LAP) that incorporates gender, waist circumference and triglycerides. After log-transformation, for each log unit increment, the risk for steatosis increased more than 4 folds [24]. Although this is a very simple test to apply, it needs validation by independent groups.

Recently, NAFLD Liver Fat Score [25] derived from a Finnish population. Gold standard was magnetic resonance spectroscopy (MRS). The score incorporates simple variables: presence of the metabolic syndrome and T2DM, fasting serum insulin, aspartate aminotransferase (AST) and AST/alanine aminotransferase (ALT) ratio. It yielded 95% Se and Sp. Information on PNPLA3 gene (rs738409) improved the accuracy for the prediction in less than 1% A Netherlands’ group confirmed these results [26]. It may be a test to take into account when assessing steatosis easily on the bench without recurring to radiology.

The best non-invasive tests for the diagnosis of steatosis are the imaging ones. US should be the first method to be used in a clinical setting. It is inexpensive, widely available and it has 60–94% Se and 66–97% Sp for hepatic steatosis [27–29]. However, US acuity decreases dramatically for mild steatosis. In a study on 100 living donors for liver transplant, US could not detect steatosis among other parameters. A score of at least 2 is highly indicative of NAFLD and it has been used in several populations and no statistical work done comparing them.

Identification and quantification of hepatic steatosis

The first challenge is when to suspect NAFLD. Suspicion will not be driven by clinical manifestations, since most patients are asymptomatic. Symptomatic patients present unspecific complaints such as fatigue, abdominal discomfort and, only seldom, manifestations of advanced liver disease. There are, however, high-risk populations in whom the prevalence is so high that per se is enough to raise the hypothesis of NAFLD. In fact, up to two thirds of patients with obesity and T2DM, present with hepatic steatosis [8]. Also, liver tests, namely aminotransferases, are usually normal, and when increased, typically present mild elevation with a fluctuant pattern [9]. Isolated increase in alkaline phosphatase is not frequent, but it has been reported in 10% of patients referred to tertiary care [10].

We should ask whether it is worth searching and diagnosing NAFLD in asymptomatic subjects with normal liver tests, since the majority will have non-progressive simple steatosis. Then again, hepatic steatosis is linked to an increase in cardiovascular risk and death. Particularly in diabetic patients, steatosis increases by more than 3-fold the risk for overall death [11] and cardiovascular disease [12]. Even in these high-risk patients, there is controversy regarding screening, among endocrinologists. Some do not recommend it, advocating that traditional scores should assess cardiovascular risk, while others consider diagnosis and evaluation of NAFLD as part of the management of DM being an indication for more intense monitoring and therapeutic intervention [13]. The guidelines on NAFLD from the American Association for the Study of Liver Diseases discourage screening for hepatic steatosis, even in high-risk patients such as diabetics. There is not enough evidence regarding diagnostic tools and treatment options, and there are no studies on the cost-effectiveness of a screening program [14]. There are however, situations in which an active search for NAFLD in asymptomatic subjects is warranted: liver donors for liver transplant, as steatosis is a risk for graft primary non-function, and in major hepatic resection [15], in which steatosis increases the risk for post-operative morbi-mortality [16].

The second question, if it is meaningful to quantify steatosis, is difficult to answer, since it has not been consistently demonstrated that the amount of fat influences prognosis. Also, it is yet to be elucidated if decreasing the amount of hepatic fat with therapeutic interventions will favourably affect the cardio-metabolic risk and the risk for progression to advanced liver disease.

Several diagnostic panels have been proposed to predict steatosis (Table 1). Steatotest incorporates 12 variables in an undisclosed formula, including α2-macroglobulin, haptoglobin, and apolipoprotein A1 [17]. In a French cohort of more than 700 patients, it showed reasonable accuracy, with 0.79 AUROC for moderate-severe steatosis, good negative predictive value (NPV), 93%, but small positive predictive value (PPV), 63%. Another French group found similar results in 288 morbid obese subjects [18]. The Poynard’ group conducted a meta-analysis in morbid obese subjects, obtaining the same conclusions [19]. We have to acknowledge that AUROC was suboptimal, it has been validated only in French cohorts, it incorporates tests not used routinely and because the formula is not public, a fee is imposed for each test applied.

Bedogni et al. first proposed Fatty Liver Index (FLI) in 2006, as an algorithm derived from the population of the Dionysos Nutrition & Liver Study [20]. The gold standard was ultrasonography (US). The index varies between 0 and 100 and is calculated through a formula incorporating: body mass index (BMI), waist circumference, triglycerides and γ-glutamyltranspeptidase (GGT). It showed good accuracy in detecting NAFLD and it has been used in several population studies [21,22]. However, the gold standard used is far from ideal, and as such, the results should be interpreted carefully. Its main indication is for epidemiological studies, in an attempt to avoid US. Recently, a study on 2075 middle-aged Caucasians from the Regional Health Registry, followed for 15 years, showed that FLI independently associates with overall, cardiovascular and cancer-related mortality [23].

The same group also proposed Lipid Accumulation Product (LAP) that incorporates gender, waist circumference and triglycerides. After log-transformation, for each log unit increment, the risk for steatosis increased more than 4 folds [24]. Although this is a very simple test to apply, it needs validation by independent groups.

Recently, NAFLD Liver Fat Score [25] derived from a Finnish population. Gold standard was magnetic resonance spectroscopy (MRS). The score incorporates simple variables: presence of the metabolic syndrome and T2DM, fasting serum insulin, aspartate aminotransferase (AST) and AST/alanine aminotransferase (ALT) ratio. It yielded 95% Se and Sp. Information on PNPLA3 gene (rs738409) improved the accuracy for the prediction in less than 1%. A Netherlands’ group confirmed these results [26]. It may be a test to take into account when assessing steatosis easily on the bench without recurring to radiology.

The best non-invasive tests for the diagnosis of steatosis are the imaging ones. US should be the first method to be used in a clinical setting. It is inexpensive, widely available and it has 60–94% Se and 66–97% Sp for hepatic steatosis [27–29]. However, US acuity decreases dramatically for mild steatosis. In a study on 100 living donors for liver transplant, US could not detect steatosis among other parameters. A score of at least 2 is highly indicative of NAFLD. US has several limitations: it is unreliable in the detection of mild steatosis, it has only up to 67% PPV [33], it is operator dependent with low inter and intra-observer agreement for steatosis, around 70% and 50% for steatosis presence and severity.
### Table 1. Complex scores for predicting steatosis.

<table>
<thead>
<tr>
<th>Author, Ref.</th>
<th>Score's name</th>
<th>N</th>
<th>Formula</th>
<th>Results</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poynard T et al., 2005, [17]</td>
<td>Steatotest</td>
<td>310 - training group 434 - 3 validation groups</td>
<td>Undisclosed formula incorporating: α2-MG, haptoglobin, apolipoprotein A1, total bilirubin, GGT, fasting glucose, triglycerides, cholesterol, ALT, age, gender and BMI</td>
<td>AUROC 0.79 (steatosis &gt;33%) 2 cut-offs: 0.3 and 0.72 - Se 90%, Sp 70%, PPV 63%, NPV 93%</td>
<td>In other French cohorts (total 494 patients) - similar results [18, 19]</td>
</tr>
</tbody>
</table>
| Bedogni G et al., 2006, [20] | Fatty Liver Index (FLI) | 216 patients 280 controls | \[
    \left( \frac{e^{0.953 \log_{10}(\text{triglycerides}) + 0.139 \times \text{BMI} + 0.718 \log_{10}(\text{GGT}) + 0.053 \times (\text{waist circumference}) - 15.745}}{1 + e^{0.953 \log_{10}(\text{triglycerides}) + 0.139 \times \text{BMI} + 0.718 \log_{10}(\text{GGT}) + 0.053 \times (\text{waist circumference}) - 15.745}} \right) \times 100
    \] | AUROC 0.84 (steatosis by US) 2 cut-offs, <30 for excluding and >60 for ruling - Se 87%, 86% Sp | Widely used in epidemiological studies |
| Kotronen A et al., 2009, [25] | NAFLD Liver Fat Score | 313 - training group 157 - validation group | -2.89 + 1.18 * (metabolic syndrome - yes = 1/no = 0) + 0.45 * (type 2 diabetes mellitus - yes = 1/no = 0) + 0.15 * (fasting serum insulin, mU/L) + 0.04 * (AST, IU/L) - 0.94 * (AST/ALT) | AUROC 0.87 2 cut-offs, -1.413 and +1.257 - Se and Sp 95% | A Netherlands’ group confirmed these results [26] |
| Bedogni G et al., 2010, [24] | Lipid Accumulation Product (LAP) | 588 subjects | (waist circumference - 65) * triglycerides if men and (waist circumference - 58) * triglycerides if women | For each log unit increase, OR for steatosis 4.28 |
| Ballestri S et al., 2012, [32] | Ultrasonographic Fatty Liver Indicator (US-FLI) | 53 patients | Brighter liver than kidney, whose intensity in contrast can be graded as mild/moderate (2 points) or severe (3 points). One extra point for each of the following: 1. Posterior attenuation of ultrasound beam, 2. Vessel blurring, 3. Difficult visualization of the gallbladder wall, 4. Difficult visualization of the diaphragm, 5. Areas of focal sparing | Score ≥2 - highly indicative of NAFLD Score ≥4 - predicts NASH: AUROC 0.796, NPV 94%, Se 48% |
| Zardi EM et al., 2011, [67] | Zardi’s US score | 94 patients (retrospective) | Two parameters: 1. Echo amplitude attenuation scored from 0 if absent, 1 mild and 2 severe, 2. Presence/absence of focal fat sparing | Score ≥1 could - Se 92%, Sp 75% |

α2-MG, alpha-2 macroglobulin; GGT, gamma-glutamyltranspeptidase; ALT, alanine aminotransferase; BMI, body mass index; AUROC, area under receiver operating characteristic; Se, sensitivity; Sp, specificity; PPV, positive predictive value; NPV, negative predictive value; NAFLD, non-alcoholic fatty liver disease; AST, aspartate aminotransferase; NASH, non-alcoholic steatohepatitis; HU, Hounsfield Units.
whereas LB quantifies the number of hepatocytes with fat. MRS detects the amount of triglycerides in the parenchyma. The discordances between both techniques can be high, are not perfect, with AUROC 0.95–0.97, 92–100% Se and 92–97% Sp [48]. The correlations with biopsy-assessed steatosis, although defined as normal up to 5.56% fat fraction, the 95th percentile study on 345 patients derived from the Dallas Heart Study cohort respectively [34]. Its accuracy significantly decreases in morbid obesity, with Se below 50% and 75% Sp in those patients [35]. In addition, graduation of steatosis can be biased by fibrosis, necrosis, oedema, and extra-hepatic adipose tissue. Recently, the hepatic/renal ratio and hepatic echo/intensity attenuation rate have been used to quantify steatosis. Xia et al. [36], standardized those measurements by introducing a tissue-mimicking phantom for a computer-assisted US program, allowing reproducibility among different US equipments. They found a good correlation between the quantification of hepatic fat by MRS and by using the following formula: liver fat content (%)=62.592 × US hepatic/renal ratio + US hepatic attenuation rate – 27.863. This is a promising technique needing further validation.

Hepatic Doppler has been evaluated for the diagnosis and grading of steatosis. Hepatic veins characteristically have a triphasic waveform, reflecting right atrial and inferior vena cava pressures [37]. In NAFLD patients, abnormal hepatic vein Doppler, biphasic or monophasic, is frequent, probably by compression of the hepatic vein by enlarged hepatocytes [37–39]. Portal venous flow is classically described as continuous, however, it is normal to detect feeble pulsatility in rhythm with the cardiac cycle. NAFLD patients frequently present a decrease in pulsatility index and mean velocity of the portal vein blood flow [37,40]. Furthermore, a decrease in hepatic artery resistance index has been consistently reported [39,41]. Those approaches need additional validation and standardization. Regarding steatosis grading, studies were not consensual.

Computed tomography (CT) has similar accuracy for NAFLD as US, being superior only for focal steatosis [42]. Steatosis is best detected in non-enhanced CT as a decrease in the attenuation of hepatic parenchyma [43]. There are three measures to determine hepatic steatosis: liver parenchymal attenuation, liver to spleen attenuation difference and ratio. They all achieved 100% Sp, but with low Se [44]. CT has several limitations, making it clinically unacceptable for screening steatosis: it exposes patients to radiation, other diffuse liver diseases may lead to misdiagnosis of steatosis, hepatic iron content increases attenuation, inducing false negatives for steatosis, and attenuation values are scanner dependent, not standardized between different manufacturers.

Magnetic resonance imaging (MRI) is superior to US in detecting and quantifying minor fat infiltration, being able to detect down to 3% of steatosis [45]. MRI exploits the difference in resonance frequencies between water and fat proton signals to quantify the signal fat fraction, i.e., liver signal attributable to fat and/or fat proton density fat fraction, i.e., fraction of mobile protons in the liver attributable to fat [33]. There are several MRI techniques that overcome the scope of this review [33,46]. More recently, MRS that directly measures proton signals from the acyl groups in hepatocyte triglyceride stores has shown incredible accuracy for diagnosing and quantifying steatosis. It obtains volume-selective MR spectra from liver volumes ranging from 1 to 27 cm³. A study on 345 patients derived from the Dallas Heart Study cohort defined as normal up to 5.56% fat fraction, the 95th percentile [47]. The correlations with biopsy-assessed steatosis, although high, are not perfect, with AUROC 0.95–0.97, 92–100% Se and 92–97% Sp [48]. The discordances between both techniques can be the result of limitations in the biopsy. In fact, MRS assesses a larger liver volume, being less susceptible to sampling errors. MRS detects the amount of triglycerides in the parenchyma whereas LB quantifies the number of hepatocytes with fat droplets, so both techniques do not quantify exactly the same thing. Finally, MRS is highly sensitive in detecting small amounts of triglycerides that may not be enough to form macrovesicles susceptible to histological visualization [42]. Taken together, could MRS be the new gold standard for hepatic steatosis diagnosis and quantification, overcoming LB? Arguments against it are: lack of information regarding necro-inflammation and fibrosis; moreover it is expensive and not broadly available. The accuracy of magnetic resonance-derived methods may decrease with significant fibrosis, raising the need for lower cut-offs in that situation [48].

Controlled attenuation parameter (CAP) is a new method that measures attenuation in the liver using signals acquired by a transient elastography probe (FibroScan®). It relies on the assumption that fat affects ultrasound propagation. Results range from 100 to 400 dB/m [49]. Sasso et al. [50] retrospectively evaluated CAP in 115 patients, showing AUROC 0.91 and 0.95 for steatosis higher than 10% and 33%, respectively, and good accuracy in grading steatosis. It was not affected by liver fibrosis. Subsequent prospective studies reported lower accuracy [51,52]. It has several advantages, it is non-ionising, easy to perform, results are operator independent not relying on subjective interpretation. It is also less susceptible to sample error as compared to LB since it evaluates 100 times more tissue. Of notice, failure to obtain measurements increases with increased BMI, and it is not yet available for the XL probes of FibroScan®. CAP performed better than Steatotest, FLI or Hepatic Steatosis Index [51]. In summary, in clinical practice, the best way to assess steatosis is US, although its accuracy decreases hugely for mild steatosis. In contrast, MRS is highly accurate for even minimal amounts of steatosis and it may even be more reliable than LB; however, its costs limit its use routinely, being a valuable tool for research purposes. Furthermore, FLI uses simple indices and may be very useful in large-scale epidemiological studies, since it avoids radiology. Also, the NAFLD Liver Fat Score can be easily used in clinical practice to discard steatosis in a specific group of patients such as with T2DM, in whom it can lead to changes in management.

Prediction of NASH

NAFLD has been subdivided into benign simple steatosis and NASH, the latter progressing to cirrhosis in up to 20% of cases [53]. Consequently, clinical differentiation has significant prognostic implications. However, since there is no effective treatment for NASH, and all NAFLD patients should be based on life-style intervention and metabolic disturbances correction, it is arguable whether the information of having NASH has an impact in the management. Nonetheless, those patients need closer follow-up, and are the ones who should be included in clinical trials.

Clinical signs and symptoms do not differentiate NASH from simple steatosis. Obesity, particularly central obesity, increases the risk for advanced disease. Also, dorsocervical lipohypertrophy is the anthropometric parameter most strongly associated with NASH and liver injury severity [54].

The metabolic syndrome has been associated with increased risk for NASH and fibrosis, among NAFLD patients [55,56], and can help select for LB [14].

Aminotransferases levels are not reliable in identifying NASH [9,57], presenting low AUROC, 0.6–0.7 [58–60]. Decreasing the
cut-off to 19 IU/L increases Se up to 70%, but at the expense of a dramatic decrease in Sp [61,62]. In summary, patients with NAFLD and increased aminotransferases levels are at higher risk of having NASH, but NASH cannot be excluded in patients with normal levels [63]. A simple model incorporating AST and the presence of DM showed great accuracy for NASH [64]. GGT can be of value. Although its association with NASH is not clear, it associates with increased mortality. In a large population-based study with a 7-year follow-up, men with NAFLD and increased GGT had a two-fold increase in mortality [65]. A simple score, neutrophil to lymphocyte ratio, was suggested to predict NASH. For each unit increase in the ratio, the likelihood of having NASH increased by 70% [66].

Several scores using US, showed only modest accuracy for NASH, with AUROC below 0.8. The US-FLI score showed a 94% NPV for excluding NASH [32]. Another score incorporating echo amplitude attenuation and focal fat sparing [67] could distinguish NASH with 92% Se. A more promising parameter is the spleen longitudinal diameter, which presented 0.920 AUROC for NASH, 88% Se and 95% Sp for values above 116 mm [68]. The same rationale has been used in CT scan, in which spleen diameter controlled for body surface area could also predict NASH, though, with low accuracy [69]. Preliminary studies on contrast US using a microbubble contrast agent showed promising results in predicting NASH, correlating with pericellular fibrosis rather than steatosis [70,71]. For the time being, radiology is not accurate to diagnose NASH, and should be used carefully.

Multiple serum biomarkers have been evaluated for predicting NASH. Cytokines did not prove to be valuable. Adiponectin showed conflicting results in NASH [72–75]. Small studies tested formulas with adipokines, with AUROC around 0.8 that need validation [76,77]. Also, tumor necrosis factor-alpha (TNF-α) and interleukin-6 seem to be elevated in patients with NASH and fibrosis, but there are discrepancies among studies and its diagnostic value has not been clarified [73,74,76,78–80]. CC-chemokine ligand-2 (CCL2) is the major cytokine in macrophage recruitment. Two pilot studies showed increased levels in NAFLD patients vs. controls and in NASH vs. simple steatosis [81,82].

Acute phase proteins have also failed to consistently prove diagnostic value. C-reactive protein was evaluated for diagnosing NASH, namely through high-sensitive kits, in more than 1000 NAFLD patients. Most studies failed to demonstrate accuracy [82–85]. Pentraxin-3 was studied in 70 NAFLD patients, showing 0.755 AUROC for predicting NASH. With 1.61 ng/ml cut-off, it accomplished 67% Se and 70% Sp [86].

Oxidative stress has been implicated in NASH pathogenesis, and an increased hepatic lipid peroxidation is accepted [87]. However, plasma and hepatic levels do not correlate, leading to conflicting results when using oxidative stress markers to predict NASH.

Keratin 18 (CK18) fragments are a marker of hepatocyte apoptosis. It is the only biomarker validated in more than 10 studies and more than 1000 NAFLD patients [58,88–98]. A recent meta-analysis showed AUROC 0.82 with 78% Se and 96% Sp [5]. Total levels of CK18, a marker of necrosis, showed equal accuracy to identify NASH as CK fragments [89,98]. A two-step approach measuring CK18 following fibroblast growth factor-21 further increased the predictive value for NASH [97]. CK18 is the most consistent single parameter for differentiating steatosis from NASH.

Individual groups have suggested multiple biomarkers that warrant external validation, not being ready to use in clinical practice. A Turkish study on 71 NAFLD patients, showed that homocysteine levels can predict NASH with AUROC 0.948; using a cut-off of 11.935 ng/ml, 92% Se and 96% Sp were achieved [99]. A study on 54 NAFLD patients suggested that serum proline dipeptide enzyme activity (SPEA) can distinguish NASH from simple steatosis with 0.85 AUROC. This enzyme catalyses the final step of collagen breakdown [61]. The soluble receptor for advanced glycation end products (sRAGE) has been implicated in inflammatory response, insulin resistance and the metabolic syndrome [63]. In a pilot study on 57 NAFLD patients, sRAGE showed a 0.77 AUROC for predicting NASH [100].

Several complex models have been created to predict NASH (Table 2). The majority of them have not been externally validated and were obtained from specific populations, most frequently morbid obesity, and thus extrapolation for the whole NAFLD population is speculative. Dixon et al. [101] created a score with simple variables: HAIR – Hypertension, increased ALT and Insulin Resistance, that predicted NASH with great accuracy. However, external validation has not been done, and the study was performed in morbid obese patients.

Palekar et al. [102] predicted NASH by the combination of age, gender, obesity, hyaluronic acid, AST and AST/ALT ratio. It yielded only modest Se and Sp. Also, this study included only few patients and once again was not validated.

Poynard et al. proposed the NashTest [103], an undisclosed formula incorporating 13 parameters, which presented 94% Sp but only 33% Se. The authors advise that NashTest should only be done when SteatoTest is positive. The same considerations for Steatotest can be made for NashTest.

NASH Diagnostics™ [51] includes, in an undisclosed formula, cleaved and total CK18, adiponectin and resistin. It derived from morbid obese patients and showed AUROC 0.854. A re-evaluation of the score showed lower accuracy [104]. The authors proposed a different model, NASH Diagnostic Panel, which included DM, gender, BMI, triglycerides, cleaved and total CK18. This improved score accomplished a 0.81 AUROC, with good Se and Sp. Once again, it derives from a small cohort of obese patients, and has not been externally validated.

The “Nice model” incorporates CK18 fragments, ALT and the presence of metabolic syndrome [105]. In a large study with obese patients, it showed good NPV but low PPV. It is also not known if it adds value to CK18 determination alone.

The Apoptosis Panel [96] includes CK18 fragments, soluble Fas and Fas ligand, and showed good accuracy in predicting NASH in a training group, which was not reproducible in a validation group of morbid obese.

Respiratory tests have also been evaluated for NASH prediction. C-caffeine breath test is a non-invasive quantitative test of liver function. Caffeine has high oral bioavailability and its metabolism is almost exclusively done by cytochrome P450A12. It showed good accuracy for NASH and significant fibrosis [106,107]. Others found an increase in C-methacin demethylation (marker of microsomal liver function), decrease in ketoisocaproate decarboxylation and methionine transmethylation (markers of mitochondrial liver function) [108,109].

Finally, two methods using liver scintigraphy with 99mTc-phytate and 99mTc-HIBI seem promising in detecting NASH. The first showed a decrease in liver/spleen uptake indicative of decreased
<table>
<thead>
<tr>
<th>Author, [Ref.]</th>
<th>Score’s name</th>
<th>N</th>
<th>Formula</th>
<th>Results</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dixon JB et al., 2001, [101]</td>
<td>HAIR</td>
<td>105 morbid obese</td>
<td>1. Hypertension, 2. Increased ALT (&gt;40 IU/L), 3. Insulin Resistance (index &gt;5)</td>
<td>≥2 parameters: AUROC 0.90</td>
<td>Se 80%, Sp 89%,</td>
</tr>
<tr>
<td>Palekar NA et al., 2006, [102]</td>
<td>Palekar’s Score</td>
<td>80 NAFLD patients</td>
<td>Score calculated by the sum of 6 risk factors: 1. Age ≥50 yr, 2. Female sex, 3. Elevated AST (≥45 IU/L), 4. BMI ≥30 kg/m², 5. AST/ALT ratio ≥0.8, 6. Plasma levels of hyaluronic acid ≥55 µg/L</td>
<td>≥3 factors: AUROC of 0.763</td>
<td>Se 74%, Sp 66%, PPV 68%, NPV 71%</td>
</tr>
<tr>
<td>Poynard T et al., 2006, [103]</td>
<td>NASH Test</td>
<td>160 - training group 97 - validation group 383 - controls</td>
<td>Undisclosed formula incorporating: α2-MG, haptoglobin, apolipoprotein A1, total bilirubin, GGT, ALT, AST, triglycerides, cholesterol, age, gender, height, weight</td>
<td>AUROC 0.79 Se 33%, Sp 94%, PPV 66%, NPV 81%</td>
<td>In other French cohorts (total 494 patients) - similar results</td>
</tr>
<tr>
<td>Gholam PM et al., 2007, [64]</td>
<td>Gholam’s model</td>
<td>97 patients</td>
<td>-2.627 * lnAST + 2.13 if diabetes mellitus</td>
<td>AUROC 0.90 cut-off 6.6 - Se 83%, Sp 82%</td>
<td></td>
</tr>
<tr>
<td>Younossi ZM et al., 2008, [91]</td>
<td>NASH Diagnostics</td>
<td>Morbid obese 69 - training group 32 - validation group</td>
<td>Undisclosed formula incorporating: cleaved and total CK18 (M30 and M65 antigens, respectively), adiponectin, resistin</td>
<td>combined AUROC 0.90 cut-off 0.432 - Se 72%, Sp 91%</td>
<td>The same group re-evaluated in 79 patients: AUROC 0.70 [104]</td>
</tr>
<tr>
<td>Younossi ZM et al., 2011, [104]</td>
<td>NASH Diagnostic Panel</td>
<td>79 patients</td>
<td>Undisclosed formula incorporating: diabetes mellitus, sex, BMI, triglycerides, M30 and M53 antigens</td>
<td>AUROC 0.81 2 cut-offs, 0.221 and 0.6183 - Se 91%, Sp 92%, PPV 83%, NPV 96%</td>
<td></td>
</tr>
<tr>
<td>Tamimi TI et al., 2011, [96]</td>
<td>Apoptosis Panel</td>
<td>95 - training group 82 - validation group (morbid obese)</td>
<td>Includes: CK18 fragments, soluble Fas and Fas ligand</td>
<td>AUROC 0.93 Cut-off -0.5509 - Se 88%, Sp 89%, PPV 86%, NPV 91%</td>
<td></td>
</tr>
<tr>
<td>Arty R et al., 2010, [105]</td>
<td>Nice model</td>
<td>464 morbid obese patients</td>
<td>Model = -5.654 + 3.780e-02 * ALT (IU/L) + 2.215e-03 * CK 18 fragments (IU/L) + 1.825 * (metabolic syndrome: yes = 1, no = 0) Logarithmic transformation = 1/[1 + Exp(-Nice Model)]</td>
<td>AUROC 0.83-0.88 Cut-off 0.14 - Se 84%, Sp 86%, PPV 44%, NPV 98%</td>
<td></td>
</tr>
</tbody>
</table>

NASH, non-alcoholic steatohepatitis; ALT, Alanine aminotransferase; AUROC, area under receiver operating characteristic; Se, sensitivity; Sp, specificity; NAFLD, non-alcoholic fatty liver disease; AST, aspartate aminotransferase; BMI, body mass index; PPV, positive predictive value; NPV, negative predictive value; α2-MG, alpha-2 macroglobulin; GGT, gamma-glutamyltransferase; CK18, cytokeratin 18.
phagocytosis by Kupffer cells [110]. The latter showed a decrease in liver/heart ratio, reflecting mitochondrial dysfunction [111].

In summary, the most validated tool to predict NASH is the determination of CK18 fragments. Unfortunately, it is not widely available in clinical practice. Several complex scores have been tried, but lack external validation and the majority were only tested in morbid obese populations.

**Prediction of significant or advanced fibrosis**

Facing a particular patient, the presence and severity of fibrosis are probably the most informative factors regarding prognosis. Also, recognizing cirrhosis allows the inclusion in screening protocols for hepatocellular carcinoma and portal hypertension. Significant fibrosis has been considered when at least F2 and advanced when F3 or F4 (Brunt’s classification).

**Serum tests**

Aminotransferases levels are not helpful in identifying fibrosis, since they can in fact decrease with histological improvement of steatosis and inflammation, despite fibrosis progression [112]. Also, the whole spectrum of liver disease, including cirrhosis, can be found in patients with normal aminotransferases [57]. The AST/ALT ratio can be of more value. In fact, several studies have found that a ratio above 1 is predictive of advanced fibrosis [2,113], probably due to an impaired AST clearance by sinusoidal liver cells. Regarding GGT, a small study suggested modest ability to predict advanced fibrosis, with a cut-off of 96.6 IU/L presenting 83% Se and 69% Sp [114].

Extracellular matrix components are obvious candidates for the evaluation of fibrosis. Hyaluronic acid (HA) production is increased when collagen synthesis is accelerated and, in advanced liver disease, sinusoidal endothelial dysfunction decreases its clearance. Small studies have shown it to predict advanced fibrosis with AUROC 0.75–0.97. Different cut-offs were used among studies [102,115–117], making it difficult to pool the data. Type IV collagen 7S domain has also been studied in small cohorts with similar results as HA [116,118]. Pentraxin-3 showed an AUROC of 0.850 for advanced fibrosis. A 2.45 ng/ml cut-off accomplished 71% Se, 94% Sp, PPV 81% and NPV 91% [86]. It is a small study that requires external validation.

**Scores**

Several models attempt predicting fibrosis, with reasonable accuracies for advanced fibrosis, but not for mild/intermediate stages (Table 3). Ratziu et al. proposed a simple score, BAAT incorporating age, BMI, triglycerides and ALT [119]. Although not sensitive, it yielded 100% Sp.

The European Liver Fibrosis Study group proposed a score, ELF, after evaluation of more than 1000 patients, including 61 with NAFLD. It combines age, HA, amino-terminal propeptide of type III collagen and tissue inhibitor of metalloproteinases (TIMP-1), and showed excellent AUROC for advanced fibrosis in NASH, 0.87. However, it performed poorly in early and intermediate stages [120]. A simplified score without age, in 196 patients showed an even better AUROC, 0.9 [121]. Although this score has been shown very valuable in fibrosis assessment in large populations of viral hepatitis, the NAFLD cohort is still very small to be used with confidence in this field.

FibroTest [122] is an undisclosed formula incorporating: age, α2-macroglobulin, bilirubin, GGT, and apolipoprotein A1. It was first developed for viral hepatitis and afterwards extended to NAFLD, where it has shown AUROC for significant fibrosis 0.75–0.86, with excellent Sp. This score was unable to distinguish mild from moderate fibrosis and one third of the patients were attributed intermediate values and therefore could not be classified. Additional causes of failure were Gilbert syndrome, cholestasis, acute inflammation and abnormal lipoprotein A1 related to lipid abnormalities frequent in these patients.

NAFLD Fibrosis Score is the most studied score [123], with external validation in 13 studies, including more than 3000 patients [5]. It incorporates age, glycemia, BMI, platelet count, albumin, and AST/ALT ratio and presents great accuracy for advanced fibrosis. A meta-analysis confirmed those results [5]. This score can already be applied in clinical practice, with the drawback that it will not be useful for one fourth of the patients that fall in indeterminate values [123].

BARD includes BMI, AST/ALT ratio and presence of DM [124]. In a large convenience sample of NAFLD obese patients, the presence of at least 2 factors increased 17-fold the risk for advanced fibrosis, with high NPV. Subsequent studies showed lower accuracy, performing worse than NAFLD Fibrosis Score [125–127]. However, it has the advantage of being easier to estimate and without indeterminate results.

FIB-4, described to evaluate advanced fibrosis in chronic liver disease, incorporating age, aminotransferases and platelet count, was compared with several other scores in NAFLD, showing AUROC most often above 0.80 and always performing the best, even when compared with NAFLD Fibrosis Score and FibroTest [127,128]. It is a promising test, maybe using in combination with other scores.

Fibrometer, which includes glucose, platelet count, aminotransferases, ferritin, body weight and age, performed better than NAFLD Fibrosis Score for significant fibrosis [129]. It needs external validation.

The NAFLD Diagnostic Panel [104] is a score for advanced fibrosis that includes DM, triglycerides, TIMP-1 and AST. In a small study, it performed better than ELF, NAFLD Fibrosis Score and APRI. Those findings need to be reproduced by other groups and in larger studies.

**Measurement of liver stiffness**

Transient elastography (FibroScan®) evaluates liver stiffness using pulse-echo ultrasound. It has demonstrated great value in assessing fibrosis in chronic hepatitis C, and it might also be useful in NAFLD patients, although with less accuracy [130]. The first report in NAFLD was in 67 Japanese patients [131], describing a stepwise increase in liver stiffness with increase severity of liver fibrosis, AUROC 0.876, 0.914, and 0.997 for significant, advanced fibrosis and cirrhosis, respectively. An NPV of 100% for excluding cirrhosis was shown. The first larger study on FibroScan® in
<table>
<thead>
<tr>
<th>Author, [Ref.]</th>
<th>Score’s name</th>
<th>N</th>
<th>Formula</th>
<th>Results</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratziu V et al., 2000, [119]</td>
<td>BAAT</td>
<td>93 NAFLD patients</td>
<td>Score calculated by the sum of 4 risk factors: 1. Age ≥50 yr, 2. BMI ≥28 kg/m², 3. Triglycerides ≥1.7 mmol/L, 4. ALT ≥5 times upper normal limit</td>
<td>AUROC 0.86</td>
<td>The 4 parameters: Se 14%, Sp 100%, PPV 100%, NPV 73%</td>
</tr>
<tr>
<td>Gholam PM et al., 2007, [64]</td>
<td>Gholam’s model</td>
<td>97 patients</td>
<td>2.45 * lnALT - 38.55 (1/HbA1c) + 5</td>
<td>AUROC 0.822</td>
<td>cut-off 6.6 - Se 76%, Sp 66%</td>
</tr>
<tr>
<td>Rosenberg W et al., 2004, [120]</td>
<td>Original European Liver Fibrosis score (OELF)</td>
<td>1021 patients, 61 with NAFLD</td>
<td>OELF = -6.38 - (Ln(age) * 0.14) - [Ln(hyaluronic acid) * 0.616] - [Ln(aminoterminal propeptide of type III collagen) * 0.586] - [Ln(TIMP-1) * 0.472]</td>
<td>AUROC 0.87 for advanced fibrosis, in NASH cut-off 0.375 - Se 89%, Sp 96%, PPV 80%, NPV 90%</td>
<td></td>
</tr>
<tr>
<td>Guha IN et al., 2008, [121]</td>
<td>Simplified ELF</td>
<td>196 NAFLD patients</td>
<td>ELF = -7.412 - [Ln(hyaluronic acid) * 0.681] - [Ln(aminoterminal propeptide of type III collagen) * 0.755] - [Ln(TIMP-1) * 0.494]</td>
<td>AUROC 0.87 for advanced fibrosis, in NASH cut-off -2.3824 - Se 91%, Sp 59%, PPV 42%, NPV 95%</td>
<td></td>
</tr>
<tr>
<td>Ratziu V et al., 2006, [122]</td>
<td>FibroTest</td>
<td>170 - training group 97 - validation group</td>
<td>Undisclosed formula incorporating: age, α2-macroglobulin, total bilirubin, GGT and apolipoprotein A1</td>
<td>AUROC 0.75-0.86 for significant fibrosis 2 cut-off: 0.3 and 0.7 - Se 77%, Sp 98%, PPV 90%, NPV 73%</td>
<td></td>
</tr>
<tr>
<td>Angulo P et al., 2007, [123]</td>
<td>NAFLD Fibrosis Score</td>
<td>480 - training group 253 - validation group</td>
<td>-1.675 + 0.037 * age (yr) + 0.094 * BMI (kg/m²) + 1.13 x IFG/diabetes mellitus (yes = 1, no = 0) + 0.99 * AST/ALT ratio - 0.013 * platelet (* 10⁹/L) - 0.66 x albumin (g/dl)</td>
<td>AUROC 0.84 for advanced fibrosis 2 cut-off: -1.455 and 0.676 - Se 82%, Sp 98%, PPV 90%, NPV 93%</td>
<td></td>
</tr>
<tr>
<td>Harrison SA et al., 2008, [124]</td>
<td>BARD</td>
<td>827 NAFLD patients</td>
<td>Includes 3 variables: 1. BMI ≥28 kg/m² (1 point), 2. AST/ALT ratio ≥0.8 (2 points), 3. Diabetes mellitus (1 point)</td>
<td>AUROC 0.81 for advanced fibrosis Score 2-4 - PPV 43%, NPV 96%</td>
<td></td>
</tr>
<tr>
<td>Calès P et al., 2009, [129]</td>
<td>Fibrometer</td>
<td>235 NAFLD patients</td>
<td>0.4184 glucose (mmol/L) + 0.0701 AST (IU/L) + 0.0008 ferritin (mg/L) - 0.0102 platelet (G/L) - 0.0260 ALT (IU/L) + 0.0459 body weight (kg) + 0.0842 age (yr) + 11.6226</td>
<td>AUROC 0.943 for significant fibrosis 2 cut-off: 0.611 and 0.715 - Se 79%, Sp 96%, PPV 88%, NPV 92%</td>
<td></td>
</tr>
<tr>
<td>McPherson S et al., 2010, [127]</td>
<td>FIB-4</td>
<td>145 NAFLD patients</td>
<td>age (yr) * AST (IU/L)/platelet count (10⁹/L) * ALT¹/² (IU/L)</td>
<td>AUROC 0.86 for advanced fibrosis cut-off: 1.3 - Se 85%, Sp 65%, NPV 95%</td>
<td></td>
</tr>
<tr>
<td>Younossi ZM et al., 2011, [104]</td>
<td>NAFLD Diagnostic Panel</td>
<td>79 NAFLD patients</td>
<td>Undisclosed formula incorporating: diabetes mellitus, triglycerides, TIMP-1 and AST</td>
<td>AUROC 0.81 for advanced fibrosis 2 cut-off: 0.0816 and 0.364 - Se 93%, Sp 91%, PPV 58%, NPV 95%</td>
<td></td>
</tr>
</tbody>
</table>

NAFLD, non-alcoholic fatty liver disease; BMI, body mass index; ALT, alanine aminotransferase; AUROC, area under receiver operating characteristic; Se, sensitivity; Sp, specificity; PPV, positive predictive value; NPV, negative predictive value; TIMP-1, tissue inhibitor of metalloproteinases; NASH, non-alcoholic steatohepatitis; GGT, gamma-glutamyltranspeptidase; IFG, increased fasting glucose; AST, aspartate aminotransferase.
NAFLD gathered 246 patients from two ethnic groups, a French and Chinese cohort, showing similar accuracy [132]. The best cut-offs were 7.0 kPa, 8.7 kPa and 10.3 kPa for significant, advanced fibrosis and cirrhosis. It has been validated by several other groups in Romania, Japan and India, with similar results and cut-offs [133–135]. A recent meta-analysis showed a pooled AUROC of 0.94, with 94% Se and 95% Sp for advanced fibrosis [5]. Limitations are considerable failure in obtaining measurements in obese patients [132], however, a new XL probe [136] decreases the failure rate in obese subjects from 35% to 6% [137]. A caveat with XL probe is that median measurements were 1.68 kPa lower as compared to M probe, and therefore different cut-offs will be needed. Besides obesity, ascites and narrow intercostals spaces may limit the acquisition of measurements. Several conditions can falsely increase values of liver stiffness, namely acute hepatitis, cholestasis, and hepatic congestion. Regarding the latter, cardiac failure decompensation, timing of previous hemodialysis in renal failure and even fasting vs. postprandial state may influence the readings. Finally, the technique itself has specific variability such as the probe position/inclination and respiratory movements, and the standards described by the manufacturer are probably not yet optimized.

An alternative to FibroScan® is acoustic radiation force impulse (ARFI) sonoelectrography that is based on the principle of mechanical excitation of tissue by short duration acoustic pulses. It has the advantage of being integrated in a conventional US system; therefore it can be performed during standard liver US [138]. A study in 131 patients with chronic liver disease, of whom 20 with NAFLD, showed a similar accuracy between FibroScan® and ARFI for fibrosis [139]. Studies on NAFLD are still preliminary with few patients recruited [140,141].

Real-time elastography is a new method that uses a B-mode US machine, incorporating elastography into the conventional US scanner. Relative hardness of the tissue is calculated and displayed as real-time color images, presented simultaneously as the B-mode images. It was evaluated in 181 NAFLD patients, and using hepatic and splenic elastic ratios, diagnostic accuracy for fibrosis was 82.6–96.0% in all stages [142].

MR elastography allows evaluation throughout the whole liver. A preliminary study on 50 patients with chronic liver disease, 20% of whom with NAFLD, found that a 2.93 kPa cut-off presented 98% Se and 99% Sp for diagnosing fibrosis [143]. It also showed good accuracy in differentiating mild from moderate/severe fibrosis. Another small study, on chronic liver disease, of which 8% NASH, suggested MR elastography to be superior to FibroScan® concerning success rate and diagnostic/staging accuracy [144]. Its major limitations are the costs, small populations studied and not being broadly available [5].

Summing up for fibrosis, the multitude of scores may lead to tremendous confusion in what to use in clinical practice. For the time being, NAFLD Fibrosis Score is the most validated. Also, it relies on simple parameters, being a good tool for selecting patients for LB. Elastography techniques seem very accurate, and are gaining a place in selecting a minority of patients that really need LB, while avoiding it in a large number of patients, as it already happens in chronic hepatitis C.
Review

Key Points

- In the clinical setting, active search for hepatic steatosis in asymptomatic individuals is not recommended, except in very precise situations, such as living donors for liver transplantation. Methods for screening hepatic steatosis are more appropriate in the context of epidemiological studies.

- FLI score has been extensively used in epidemiological studies, as an attempt to avoid LB, however, its accuracy is not perfect, and it was designed having US and not LB as gold standard. MRS has demonstrated an excellent accuracy for detecting and quantifying steatosis and may even perform better than LB. However, it is expensive and not widely available, CAP seems promising in assessing and quantifying steatosis, but should be further validated.

- Concerning the prediction of NASH, most biomarkers and complex scores showed suboptimal AUROC. CK18 fragments, a marker of apoptosis, is the only biomarker extensively validated. However, it is not broadly available and as such it has not yet been introduced in clinical practice.

- Regarding fibrosis prediction, the only model adequately validated is NAFLD Fibrosis Score that correctly identifies and excludes advanced fibrosis in most patients and can avoid LB in a substantial number of patients. Elastography, either by FibroScan® or by ARFI, can reliably predict advanced fibrosis and may help in the decision to perform LB. ARFI has been less studied, but has the advantage of being incorporated in a regular US set.

Conclusions

The search for the ideal non-invasive test has not been accomplished yet, which explains the vast number of tests available. Imaging techniques, biomarkers and complex models have been studied as tools to predict steatosis, NASH and fibrosis. New methods are usually compared with LB that has accuracy far from perfect for NASH and even for fibrosis. For assessing steatosis, US may also be used as comparison, and US cannot detect mild steatosis. In that way, new methods can perform better than the gold standard leading to underestimation of the power of the test.

Most methods have been evaluated in small pilot studies and have not been externally validated. The majority showed only suboptimal accuracy for NASH. Although discrimination for advanced fibrosis is usually reasonable, no test detects confidently mild/moderate fibrosis. Also, standardization of cut-offs is difficult and most methods lack reproducibility.

For the time being, non-invasive tests do not replace LB, but may avoid it in a large number of cases with low probability or high-risk for having advanced fibrosis/cirrhosis. A suggestion of a decision algorithm of when to perform LB, according to current guidelines and a meta-analysis on non-invasive tools [5,14,145], is presented in Fig. 1.

Conflict of interest

The authors declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

References

Review


