Noninvasive diagnosis of nonalcoholic fatty liver disease

Robert P. Myers

Abstract

Nonalcoholic fatty liver disease (NAFLD) is the most common cause of chronic liver disease, affecting approximately 30% of Western populations and a frequent indication for liver transplantation. The histologic spectrum of NAFLD includes simple steatosis, which has a benign prognosis, and nonalcoholic steatohepatitis, a more aggressive form of liver injury that may progress to cirrhosis and its complications. At present, the only widely accepted means of differentiating these lesions, including the severity of hepatic fibrosis, is liver biopsy. However, due to the invasiveness of this procedure, the rising prevalence of NAFLD, and the expected availability of effective therapies for this condition, the identification of noninvasive tools for the diagnosis and staging of NAFLD has emerged as a major clinical and research priority. This review summarizes important advances in this field during the past decade, including the development of biomarkers of hepatic fibrosis, apoptosis, and inflammation; novel imaging techniques such as transient elastography; and high-throughput technologies including proteomics and genomics. Future studies must focus on the development of accurate, inexpensive, and reliable tools that can differentiate the major histologic determinants of NAFLD; are responsive to changes in NAFLD severity due to therapeutic intervention and time; and have prognostic significance. Until such tools are developed, liver biopsy remains an important tool in the assessment of patients with NAFLD.

Key words: Non-alcoholic fatty liver disease, nonalcoholic steatohepatitis, liver fibrosis, steatosis, non-invasive procedure, transient elastography (FibroScan), serum markers, liver biopsy.

Introduction

The development of noninvasive markers of nonalcoholic fatty liver disease (NAFLD) has emerged as a major clinical and research priority. The quest for accurate tools to stage NAFLD-related liver injury stems from a variety of recent developments. First, the prevalence of NAFLD has grown to epidemic proportions; it is currently the most common cause of abnormal liver biochemistry and cryptogenic cirrhosis, and a frequent indication for liver transplantation. In Western countries, the prevalence of NAFLD approaches 30%; in the United States alone, an estimated 80 to 100 million individuals are affected. Second, it is now well recognized that NAFLD exists as a spectrum consisting of two major phenotypes that have drastically different natural histories. While the majority of patients have simple steatosis, which has a benign clinical course, approximately 10-20% of individuals have nonalcoholic steatohepatitis (NASH), a potentially serious condition. In one study, patients with NASH had significantly reduced survival compared to the general population and a higher risk of liver-related (2.8% vs. 0.2%) and cardiovascular death (15.5% vs. 7.5%). End-stage liver disease occurred in 10% of patients during follow-up, including three cases of hepatocellular carcinoma. In contrast, survival of patients with simple steatosis was similar to that of the control population and none of these patients developed liver failure. Third, although there are currently no approved therapies for NAFLD, new treatments including insulin-sensitizing agents are under investigation in large-scale clinical trials. Noninvasive markers that will facilitate identification of patients most in need of these treatments (i.e. those with NASH) and serve as responsive and readily repeatable therapeutic endpoints are vital. Finally, currently available tools for assessing liver injury in patients with NAFLD are sub-optimal; none satisfy all of the crite-
ria for an ‘ideal’ tool (Table I). In my opinion, the major goals of any such tool are three-fold: 1) to diagnosis NAFLD; 2) to differentiate simple steatosis from NASH; and 3) to determine the severity of hepatic fibrosis. At present, only liver biopsy can achieve these goals. Nevertheless, liver biopsy has limitations; perhaps most important is its unsuitability as a screening tool for a condition that affects roughly one-third of the population. Other limitations include its cost, variability in pathological interpretation, the difficulty of performing repeated biopsies to track hepatic injury over time, and its small, but not inconsequential risk of complications including major hemorrhage (0.1%), and death (0.01%). Moreover, due to its ability to sample only 1:50,000th of the liver, biopsy carries an important risk of sampling error. Studies predominantly in viral hepatitis have suggested that a single liver biopsy will miss cirrhosis in 10-30% of patients and misclassify fibrosis by at least one stage in 20-30%. In an illustrative study of 51 patients with NAFLD by Ratziu et al., dual biopsies of the right and left hepatic lobes revealed significant discordance in measurements of steatosis (≥ 20% of hepatocytes affected), inflammation (≥ 1 grade), and fibrosis (≥ 1 stage) in 18%, 41%, and 43% of patients, respectively. This issue needs to be considered when evaluating the accuracy of noninvasive tools derived using liver biopsy as the reference standard.

**Routine studies for the diagnosis of NAFLD**

Information available to the practicing clinician when faced with a patient who may have NAFLD includes the history and physical examination, routine laboratory investigations, and typically, an ultrasound (US). Since most patients with NAFLD are asymptomatic and complaints, when present, are nonspecific (e.g. fatigue, right upper quadrant discomfort), the presence or absence of symptoms is an unreliable means of assessing NAFLD severity. Likewise, although various historical features have been associated with advanced NAFLD (e.g. older age, diabetes, obesity), their utility in staging disease when used in isolation is limited. Laboratory tests, including the alanine (ALT) and aspartate aminotransferase (AST), gamma-glutamyl-transpeptidase (GGT), albumin, bilirubin, international normalized ratio (INR), and platelets suffer the same limitation. For example, the most commonly used screening modality, the ALT, typically fluctuates in NAFLD and is normal in more than two-thirds of NASH patients at any given time. In one study, the sensitivity of serum ALT for NASH in a large bariatric surgery cohort was only 40%. For the identification of at least 5% steatosis, Poynard and colleagues reported that an ALT > 50 IU/L had a sensitivity and specificity of only 72% and 62%, respectively (area under the receiver operating characteristic curve [AUROC] 0.61). Although some have recommended a decrease in the reference ranges for ALT as a means of improving its sensitivity, the resultant high false positive rates are clearly unacceptable. Combining these measures is a potential, but imperfect means of overcoming these limitations. For example, an AST/ALT ratio above 1 may indicate advanced fibrosis; however, its sensitivity is poor (~50%). Similarly, the AST/platelet ratio index (APRI), a well-validated and readily available tool for staging hepatitis C (HCV)-related fibrosis - particularly for excluding advanced fibrosis - requires additional study in patients with NAFLD. Finally, measures of liver function such as albumin, bilirubin, and INR may be abnormal in patients with advanced NAFLD-related cirrhosis, but cannot reliably differentiate earlier stages of disease.

The most commonly used imaging modality for the diagnosis of NAFLD is US, in which steatosis manifests as a diffuse increase in hepatic echogenicity (the so-called ‘bright liver’). Additional sonographic features include hepatomegaly; decreased ability of the ultrasound beam to penetrate the liver causing posterior darkness and loss of definition of the diaphragm (‘posterior beam attenuation’); and decreased visualization of the portal and hepatic veins giving rise to a bland appearance of the liver because of compression by the surrounding fat-laden parenchyma. In various studies, the sensitivity and specificity of US for NAFLD have ranged from 60-95% and 84-100%, respectively. Although these are seemingly acceptable operating characteristics, it is important to note that the sensitivity of US decreases markedly in the setting of mild steatosis affecting less than 30% of hepatocytes. Although computed tomography (CT) and magnetic resonance imaging (MRI) can more reliably grade steatosis than US, they are not suitable for routine screening (e.g. due to cost and radiation exposure with CT), nor can any of these tools differentiate simple steatosis from NASH.

**Novel tools for the diagnosis of hepatic steatosis**

Despite the high prevalence of NAFLD, relatively few studies have examined novel, noninvasive measures for

---

**Table I. Characteristics of the ideal diagnostic marker of NAFLD.**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver-specific</td>
<td>Independent of metabolic alterations (e.g. impaired biliary or renal excretion and reticuloendothelial function)</td>
</tr>
<tr>
<td>Able to identify simple steatosis</td>
<td>Able to differentiate simple steatosis from NASH</td>
</tr>
<tr>
<td>Able to distinguish individual stages of fibrosis</td>
<td>Sensitive enough to distinguish individual stages of fibrosis (or at the very least, advanced from mild fibrosis)</td>
</tr>
<tr>
<td>Easy to perform and acceptable to patients and physicians</td>
<td>Easy to perform and acceptable to patients and physicians</td>
</tr>
<tr>
<td>Inexpensive</td>
<td>Inexpensive</td>
</tr>
<tr>
<td>Reproducible</td>
<td>Reproducible</td>
</tr>
<tr>
<td>Responsive to changes in disease severity</td>
<td>Responsive to changes in disease severity attributable to therapy or natural history of the disease</td>
</tr>
<tr>
<td>Correlated with clinical outcomes</td>
<td>Correlated with clinical outcomes (e.g. liver-related morbidity and mortality)</td>
</tr>
</tbody>
</table>
the prediction of hepatic steatosis. The identification of a simple, inexpensive tool for fatty liver would have enormous potential for population screening. Akin to the identification of biomarkers for the prediction of hepatic fibrosis (see below), Poynard et al. recently reported a combination of markers for steatosis referred to as the SteatoTest (BioPredictive, Paris, France). This proprietary index, which has yet to be externally validated, combines age, gender, body mass index (BMI), cholesterol, triglycerides, glucose, ALT, GGT, bilirubin, haptoglobin, alpha-2-macroglobulin, and apolipoprotein A1 in a logistic regression formula. For the prediction of steatosis ≥5%, the SteatoTest had an AUROC of 0.80 in a cohort of 811 patients with NAFLD, HCV, hepatitis B (HBV), and alcoholic liver disease. Although this AUROC is generally considered indicative of a ‘good’ diagnostic test, a substantial overlap between grades of steatosis will likely limit its widespread use. For example, the median SteatoTest value (which ranges between 0 and 1) was 0.14 in blood donors, 0.26 in patients without steatosis; 0.43 with 1-5% steatosis, 0.62 with 5-33% steatosis; 0.70 with 34-66% steatosis, and 0.75 with >66% steatosis. Moreover, it is debatable whether the diagnosis of 5% steatosis is of clinical significance.

An emerging imaging modality for the quantitative assessment of hepatic steatosis is proton magnetic resonance spectroscopy (1H MRS). This technology grades hepatic triglyceride content (HTGC) by directly measuring protons in acyl groups of liver tissue triglycerides. In a recent illustrative study, Johnson et al. reported that compared with lean individuals, obese patients with and without hepatic steatosis had relative increases in hepatic lipid saturation and decreases in polyunsaturation. The accuracy and safety of this technique make it an ideal methodology to assess and monitor changes in HTGC in response to various therapeutic interventions. For example, in a randomized controlled trial of pioglitazone for NASH, Belfort et al. used 1H MRS to demonstrate a 54% reduction in hepatic fat content in patients treated with pioglitazone for 6 months; no difference was observed in controls. Although highly accurate and reproducible, the limited availability and cost of 1H MRS will undoubtedly limit its use to research applications.

Preliminary studies support a role for high-throughput or ‘omics’ technologies in the assessment of hepatic steatosis. For example, Younossi et al. examined hepatic gene expression and serum protein profiles in 98 bariatric surgery patients. In total, 7 genes/4 protein peaks were identified that were differentially expressed in patients with simple steatosis versus no steatosis; 3 genes/4 proteins for steatosis with nonspecific inflammation versus no steatosis; and 14 genes/4 proteins for NASH versus controls. The specific genes hold clues to the pathogenesis of this condition. For example, patients with steatosis alone had a down-regulation of Mu-class glutathione S-transferases, which are important enzymes in the cellular defense against oxidative stress. Conversely, the FGFR2 gene encoding fibroblast growth factor receptor 2 was significantly up-regulated. This gene likely plays a role in liver regeneration and hepatocyte survival. Similar studies employing a lipidomic approach have demonstrated changes in the lipid composition of the liver across the histologic spectrum of NAFLD patients. In another recent study employing a metabolomics approach, Subramanian et al. identified a small metabolite profile that was 100% sensitive and 96% specific in differentiating patients with hepatic steatosis versus controls. Specifically, principal components analysis showed that increased levels of the beta anomer of glucose and decreased levels of lactate differentiated these two groups.

As the availability and expense of these technologies improve and validation studies emerge, one might envision generation of a ‘hepatic steatosis profile’ that can identify the presence of fatty liver, estimate the severity of histologic features (e.g. simple steatosis versus NASH), and predict the risk of morbidity and mortality.

**Novel tools for the differentiation of simple steatosis from NASH**

NASH is characterized by hepatic steatosis with inflammation, evidence of liver cell injury (e.g. ballooning degeneration, necrosis), and fibrosis. As mentioned, the differentiation of NASH from simple steatosis has important prognostic implications and will be vital in the future to target effective therapies at individuals with the highest risk of complications. Despite extensive investigation, currently available tools for differentiating these important NAFLD phenotypes are not yet ready for the clinic. Different tools have been studied to identify patients with NASH; they can broadly classified as biomarkers of oxidative stress, inflammation, and hepatocyte apoptosis (markers of fibrosis will be reviewed later).

Oxidative stress – an excess of pro-oxidant compared with anti-oxidant mechanisms – is key to the pathogenesis of NAFLD. Both humans and experimental models of NASH have increased levels of oxidative stress within the liver. Thus, a number of investigators have examined markers of oxidative stress in the blood and/or breath of individuals with NAFLD in an attempt to differentiate simple fatty liver from NASH. In general, results have been disappointing and/or inconsistent, potentially because systemic levels of oxidative stress may or may not correlate with hepatic levels. Nonetheless, several interesting studies along this line of research deserve mention. For example, in a study of 21 patients with NASH and 19 age/gender/BMI-matched controls, Chalasani et al. measured circulating levels of lipid peroxidation products and their metabolic and nutritional correlates. Although subjects with NASH had higher serum levels of oxidized low-density lipoprotein and thiobarbituric acid-

---

**References:**

- Poynard et al., *Noninvasive diagnosis of nonalcoholic fatty liver disease*.
- Chalasani et al., *Noninvasive diagnosis of nonalcoholic fatty liver disease*.
reacting substance (TBARS), these differences were not significant in multivariate analyses and ‘total antioxidant status’ did not differ between groups. There were also no differences in the intake of nutritional determinants of oxidative stress and defense (e.g. vitamin E, beta-carotene, selenium). In another study of 22 patients with NASH and 22 controls, the total plasma antioxidant response (TAR), total plasma peroxide concentration, and an oxidative stress index (OSI; the ratio of total peroxide to TAR) were significantly higher in NASH patients than controls. Although fibrosis scores in those with NASH were significantly correlated with total peroxide levels, OSI (both positively correlated), and TAR (inversely correlated), no significant associations were observed with the grade of necroinflammatory activity. Finally, in a recent French study including 64 patients with NAFLD and viral hepatitis, there was no significant association between the presence of steatosis and blood levels of markers of oxidative stress (TBARS, superoxide dismutase activity, plasma and erythrocyte glutathione peroxidase activity, vitamin E, and selenium). Clearly, based on these discrepant results, additional studies are necessary before markers of oxidative stress can be used clinically to differentiate simple steatosis from NASH.

A variety of biomarkers of inflammation have also been studied for this purpose. It is well recognized that an imbalance between pro- and anti-inflammatory cytokines (e.g. tumor necrosis factor-α [TNF-α], interleukin-6 [IL-6]) and adipokines (e.g. adiponectin) are pathogenic in the progression from simple steatosis to NASH. As with the aforementioned studies regarding markers of oxidative stress, an inflammatory biomarker that has been validated and is ready for clinical use is not available. However, interesting preliminary findings have been reported. For example, in a study of 109 patients with NAFLD who underwent biopsy (80 of whom had NASH) and 82 controls, TNF-α levels were significantly higher and adiponectin levels lower in NAFLD patients. Although TNF-α levels did not differ between patients with simple steatosis and NASH, adiponectin levels were significantly lower in the latter subgroup. In this study, 84% of patients with NASH had adiponectin levels less than 10 µg/mL versus only 52% of those with simple steatosis. An index including adiponectin with the homeostasis model assessment of insulin resistance (HOMA-IR) had an AUROC of 0.79 (95% confidence interval [CI] 0.68-0.89) for the differentiation of simple steatosis from NASH. Jarrar and colleagues also reported lower adiponectin levels in patients with NASH versus simple steatosis, but in this study, TNF-α levels were higher in the former group. In comparison to simple steatosis, four factors were independently associated with the presence of NASH: age, ALT, interleukin-8, and adiponectin.

Another inflammatory biomarker that has been studied in NAFLD is C-reactive protein (CRP), an acute phase reactant synthesized in the liver that is elevated in chronic inflammatory conditions and associated with central and visceral adiposity. Somewhat surprisingly, several studies have failed to show an association between CRP levels measured using highly sensitive assays and the presence of NASH. In one of these studies; however, CC-chemokine ligand-2 (CCL2) was higher in patients with NASH than those with simple steatosis. This and other inflammatory markers including cytokines (e.g. IL-6) and adipokines such as resistin and visfatin, deserve further study as potential biomarkers of NASH.

Another exciting development in the field of NAFLD biomarkers is the recognition that hepatocyte apoptosis plays an important role in the progression of NAFLD-related liver injury. Both humans and animal models of NASH have increased hepatocyte cell death by apoptosis that is not seen in simple steatosis. This finding has been exploited for diagnostic purposes via the development of assays for cleavage products of various substrates for effector caspases (notably caspase-3). A promising assay measures caspase-generated fragments of cytokeratin-18 (CK-18), the major intermediate protein filament in the liver. In a study by Wieckowska and colleagues, plasma CK-18 fragments were markedly increased in patients with NASH compared with simple steatosis or normal biopsies. Moreover, CK-18 fragment levels independently predicted NASH (odds ratio [OR] 1.95 for every 50 U/L increase; 95% CI 1.18-3.22). A cutoff value of 395 U/L was 99.9% specific and 85.7% sensitive for NASH. In another study also restricted to bariatric surgery patients, Diab et al. observed higher cleaved CK-18 levels in patients with NASH compared with simple steatosis and controls; the AUROC for NASH was 0.88 (95% CI 0.77-0.99). Interestingly, a significant decrease in CK-18 levels was observed in most patients six months postoperatively, suggesting that this tool is responsive to histologic changes attributable to effective therapy for NAFLD. Ultimately, this test is unlikely to be used in isolation since most studies would suggest that combinations of biomarkers have the highest predictive utility. In one such study by Younossi and colleagues that included 101 patients undergoing bariatric surgery, cleaved CK-18 was combined with a marker of necrosis (CK-18 minus cleaved CK-18), adiponectin, and resistin in an index now referred to as the NASH Diagnostics™ Biomarker Panel. For the differentiation of simple steatosis from NASH, this index had AUROCs of 0.91 (95% CI 0.81-0.96) and 0.73 (95% CI 0.55-0.87) in the derivation and validation cohorts, respectively. If these findings are validated in large-scale, multicenter studies that include more diverse patient populations, markers of apoptosis may prove useful for the identification of patients with NASH.

The final biomarker of inflammation in NAFLD that deserves mention is the NashTest (Biopredictive; Paris, France) developed by Poynard and colleagues. This proprietary tool includes the components of the Stea-
to Test in addition to AST. In a study that included 257 patients with NAFLD who underwent liver biopsy, this panel was 71% sensitive and 94% specific for the diagnosis of NASH versus no NASH or borderline NASH according to the NAFLD Activity Score. The AUROC for this outcome was 0.75 (95% CI 0.67-0.82). As this index has not been externally validated, additional study is necessary prior to its widespread dissemination.

**Noninvasive assessment of hepatic fibrosis**

An important endpoint for a biomarker in NAFLD is the ability to identify advanced fibrosis because patients with this degree of liver damage are at risk of major complications including end-stage liver disease and hepatocellular carcinoma. A large number of biochemical markers have been assessed for the evaluation of hepatic fibrosis across a diverse spectrum of liver diseases. These assays can be broadly divided into 2 groups: 1) indirect markers of fibrosis that reflect alterations in hepatic function, but do not directly reflect extracellular matrix (ECM) metabolism (e.g. liver biochemistry, platelets); and 2) direct markers of fibrosis that reflect the dynamics of ECM turnover (e.g. matrix metalloproteinases and their inhibitors, collagens, hyaluronic acid). Most studies suggest that combinations of direct and indirect markers, with in some cases, routine clinical variables (e.g. age, sex, and BMI), are most useful. In fact a growing number of fibrosis marker panels have been developed and many commercialized. Although most panels were derived in patients with chronic HCV, validation studies in other conditions, including NAFLD, are emerging. Table II includes summary data for fibrosis marker panels that have been specifically examined in patients with NAFLD. Although I will briefly discuss each of these panels, interested readers are directed to excellent general reviews of liver fibrosis markers.42,43

The first index developed to predict NAFLD-related fibrosis, the BAAT score, was reported by Ratziu and colleagues in a cohort of 93 overweight patients. According to this score, 1 point is given for each of BMI ≥ 28 kg/m2, age ≥ 50 years, ALT ≥ twice normal, and triglycerides ≥ 1.7 mmol/L. In patients with a BAAT score of 0 or 1, significant fibrosis (≥ F2) was excluded with 100% certainty (sensitivity 100%, specificity 47%). At the opposite end of the spectrum, a BAAT score of 4 was 100% specific for significant fibrosis, although the sensitivity was poor (14%). In a recently reported study employing a

### Table II. Overview of selected serum fibrosis marker panels evaluated in patients with NAFLD.*

<table>
<thead>
<tr>
<th>Test (Author, publication)</th>
<th>Components</th>
<th>Proprietary Fibrosis (95% CI)</th>
<th>AUROC for F2-4 vs. F0-1</th>
<th>Test threshold</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAAT score (Ratziu, Gastroenterology 2000)†</td>
<td>BMI ≥ 28 kg/m², age ≥ 50 yrs, ALT ≥ 2x ULN, triglycerides ≥ 1.7 mmol/L (1 point each)</td>
<td>No</td>
<td>0.84</td>
<td>0 or 1</td>
<td>100%</td>
<td>47%</td>
</tr>
<tr>
<td>FibroTest/FibroSure (Ratziu, BMC Gastro 2006)†</td>
<td>Age, sex, α2-macroglobulin, haptoglobin, apolipoprotein A1, GGT, bilirubin</td>
<td>Yes (BioPredictive, LabCorp)</td>
<td>0.86 (0.77-0.91)</td>
<td>0.30</td>
<td>77%</td>
<td>77%</td>
</tr>
<tr>
<td>NAFLD Fibrosis score (Angulo, Hepatology 2007)†</td>
<td>Age, glucose, BMI, platelets, albumin, AST/ALT ratio</td>
<td>No</td>
<td>0.82 (0.76-0.88)</td>
<td>-1.455</td>
<td>77%</td>
<td>71%</td>
</tr>
<tr>
<td>BARD score (Harrison, Gut 2008)†</td>
<td>BMI &gt; 28 kg/m² (1 point), AST/ALT ratio &gt; 0.8 (2 points), diabetes (1 point)</td>
<td>No</td>
<td>0.81†</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>European liver fibrosis (ELF) panel (Guha, Hepatology 2008)†</td>
<td>Hyaluronic acid, TIMP-1, PIII-NP</td>
<td>Yes (iQuor)</td>
<td>0.90 (0.84-0.96)</td>
<td>-0.6746</td>
<td>90%</td>
<td>50%</td>
</tr>
<tr>
<td>FibroMeter (Cales, J Hepatol 2009)†</td>
<td>Age, glucose, AST, ALT, ferritin, platelets, weight</td>
<td>Yes (BioLiveScale)</td>
<td>0.95</td>
<td>N/A</td>
<td>79%</td>
<td>96%</td>
</tr>
</tbody>
</table>

AUROC, area under the receiver operating characteristic curve; GGT, gamma-glutamyltransferase; N/A, not available; PIII-NP, N-terminal peptide of procollagen III; TIMP-1, tissue inhibitor of metalloproteinase-1; ULN, upper limit of normal.

* Where possible, test performance characteristics extracted from validation cohorts in original publications.
† Performance characteristics for differentiation of F3-4 vs. F0-2 fibrosis.
similar approach, Harrison et al. derived the BARD score, which includes a combination of BMI > 28 kg/m² (given 1 point), AST/ALT ratio > 0.8 (2 points), and the presence of diabetes mellitus (1 point), in a cohort of 827 patients with NAFLD. The score, which ranges from 0 to 4, had an AUROC of 0.81 for bridging fibrosis or cirrhosis. With a BARD score of 0 or 1, simple steatosis or NASH with mild fibrosis (F0-2) was confirmed with 96% certainty, obviating liver biopsy in these patients. On the contrary, BARD scores of 3-4 were associated with a 43% probability of NASH with bridging fibrosis or cirrhosis (F3-4). In these patients, a liver biopsy would be recommended to confirm or exclude advanced disease due to this suboptimal predictive value. Although additional studies are necessary to validate this simple index, it is promising due its inclusion of readily available clinical data and the ability to exclude significant disease with very high accuracy.

In another study that included routinely available parameters, Angulo and colleagues from multiple international centers derived and validated the NAFLD Fibrosis Score, which consists of age, hyperglycemia, BMI, platelets, albumin, and the AST/ALT ratio combined in a logistic regression formula. For the diagnosis of bridging fibrosis or cirrhosis (F3-4), the AUROCs of the NAFLD Fibrosis Score were 0.88 (95% CI 0.85-0.92) and 0.82 (95% CI 0.76-0.88) in the derivation and validation groups, respectively. Scores less than a low threshold with 93% certainty (sensitivity 82%, specificity 77%). On the contrary, BARD scores of 3-4 were associated with a 43% probability of NASH with bridging fibrosis or cirrhosis (F3-4). In these patients, a liver biopsy would be recommended to confirm or exclude advanced disease due to this suboptimal predictive value. Although additional studies are necessary to validate this simple index, it is promising due its inclusion of readily available clinical data and the ability to exclude significant disease with very high accuracy.

The most widely validated fibrosis marker panel is the FibroTest (Biopredictive; Paris, France), originally described by Imbert-Bismut et al. in patients with chronic HCV. This index includes age, gender, bilirubin, GGT, haptoglobin, apolipoprotein A1, and alpha-2-macroglobulin combined in a logistic regression formula that is available on a proprietary basis. Among 170 patients with NAFLD, the AUROC of the FibroTest for septal fibrosis (≥ F2) and bridging fibrosis (F3-4) were 0.86 (95% CI 0.77-0.91) and 0.92 (95% CI 0.83-0.96), respectively. These figures are similar to those reported in patients with other liver conditions including chronic HCV. Importantly, there was a discordance of at least 2 fibrosis stages estimated by the FibroTest and biopsy in 10% of patients; half were attributable to failure of the fibrosis marker panel. An important cause of discordance – specifically, a false negative FibroTest measurement - was a highly elevated apolipoprotein A1 concentration, potentially attributable to the dyslipidemia frequently seen in patients with NAFLD.

The most recently developed serum marker panel for the diagnosis of NAFLD-related fibrosis is referred to as the NAFLD Fibrometer, developed by Cales and colleagues in a French study of 235 patients. This index, which includes age, glucose, AST, ALT, ferritin, platelets, and body weight, had AUROCs of 0.94 and 0.90 for septal fibrosis (≥ F2) and cirrhosis (F4), respectively. Its performance was significantly better than both the NAFLD Fibrosis Score and APRI. If validated, this index may prove very useful clinically because it includes readily available parameters in a formula that has been published.

Several conclusions can be made when reviewing this and additional literature regarding serum fibrosis marker panels. First, none of the available tools meet all of the requirements for the ‘ideal’ fibrosis marker (Table I). Second, none is clearly superior to the others; most have AUROCs for septal fibrosis (≥ F2) of 0.80-0.90, considered in the ‘good’ range for a diagnostic test. Third, all are limited by overlap in values between adjacent stages of fibrosis - partly attributable to the limitations of liver biopsy - that may hinder accurate identification of individual stages. Thus, roughly 25-50% of estimates are in a so-called ‘intermediate range’ in which liver biopsy may be necessary to accurately stage fibrosis. Third, prospective data regarding changes in these indices over time (e.g. with disease progression or therapy) or their prognostic value is virtually non-existent in NAFLD. Most experts would agree that the most rational approach is to use these markers as a complement to liver biopsy on a case-by-case basis. It would be unrealistic to expect any index to completely replace liver biopsy, which offers a wealth of additional information including the severity of necroinflammatory activity and the presence of coexistent liver conditions (e.g. iron overload). It is important to remember that up to a third of patients suspected of having NAFLD have another cause of liver enzyme elevations identified by liver biopsy.

The most widely validated fibrosis marker panel is the FibroTest (Biopredictive; Paris, France), originally described by Imbert-Bismut et al. in patients with chronic HCV. This index includes age, gender, bilirubin, GGT, haptoglobin, apolipoprotein A1, and alpha-2-macroglobulin combined in a logistic regression formula that is available on a proprietary basis. Among 170 patients with NAFLD, the AUROC of the FibroTest for septal fibrosis (≥ F2) and bridging fibrosis (F3-4) were 0.86 (95% CI 0.77-0.91) and 0.92 (95% CI 0.83-0.96), respectively. These figures are similar to those reported in patients with other liver conditions including chronic HCV. Importantly, there was a discordance of at least 2 fibrosis stages estimated by the FibroTest and biopsy in 10% of patients; half were attributable to failure of the fibrosis marker panel. An important cause of discordance – specifically, a false negative FibroTest measurement - was a highly elevated apolipoprotein A1 concentration, potentially attributable to the dyslipidemia frequently seen in patients with NAFLD.

The most recently developed serum marker panel for the diagnosis of NAFLD-related fibrosis is referred to as the NAFLD Fibrometer, developed by Cales and colleagues in a French study of 235 patients. This index, which includes age, glucose, AST, ALT, ferritin, platelets, and body weight, had AUROCs of 0.94 and 0.90 for septal fibrosis (≥ F2) and cirrhosis (F4), respectively. Its performance was significantly better than both the NAFLD Fibrosis Score and APRI. If validated, this index may prove very useful clinically because it includes readily available parameters in a formula that has been published.

Several conclusions can be made when reviewing this and additional literature regarding serum fibrosis marker panels. First, none of the available tools meet all of the requirements for the ‘ideal’ fibrosis marker (Table I). Second, none is clearly superior to the others; most have AUROCs for septal fibrosis (≥ F2) of 0.80-0.90, considered in the ‘good’ range for a diagnostic test. Third, all are limited by overlap in values between adjacent stages of fibrosis - partly attributable to the limitations of liver biopsy - that may hinder accurate identification of individual stages. Thus, roughly 25-50% of estimates are in a so-called ‘intermediate range’ in which liver biopsy may be necessary to accurately stage fibrosis. Third, prospective data regarding changes in these indices over time (e.g. with disease progression or therapy) or their prognostic value is virtually non-existent in NAFLD. Most experts would agree that the most rational approach is to use these markers as a complement to liver biopsy on a case-by-case basis. It would be unrealistic to expect any index to completely replace liver biopsy, which offers a wealth of additional information including the severity of necroinflammatory activity and the presence of coexistent liver conditions (e.g. iron overload). It is important to remember that up to a third of patients suspected of having NAFLD have another cause of liver enzyme elevations identified by liver biopsy.
Transient elastography (TE; Fibroscan®, Echosens, Paris, France) is a relatively novel approach to measuring liver stiffness - a surrogate for liver fibrosis – that has gained increasing clinical use. The technique utilizes an US transducer probe mounted on the axis of a vibrator which transmits a low frequency (50 Hz), elastic shear wave through the liver. The velocity of the wave, measured using pulse-echo US acquisition, is proportional to liver stiffness; stiffer, fibrotic livers are associated with faster wave propagation. TE is entirely noninvasive, takes less than 5 minutes to complete, can be reliably performed following a short training period (≥50 examinations are necessary for competency), is highly reliable (coefficients of variation ~3%), and can be readily integrated into an outpatient hepatology clinic. Moreover, since TE measures liver stiffness in a volume roughly 100-times that of the typical liver biopsy, it is likely more representative of the entire hepatic parenchyma. Since originally described in 2003, numerous publications have assessed the performance of TE across a spectrum of liver diseases. In a systematic review of 50 such studies, Friedrich-Rust et al. reported summary AUROCs (95% CI) for the diagnosis of septal fibrosis (≥ F2), bridging fibrosis (≥ F3) and cirrhosis (F4) of 0.84 (0.82-0.86), 0.89 (0.88-0.91), and 0.94 (0.93-0.95), respectively. Thus far, studies describing the performance of TE in patients with NAFLD are limited. However, in a recent study including 97 NAFLD patients, Yoneda et al. reported similar operating characteristics. For the diagnosis of septal fibrosis (≥ F2), liver stiffness exceeding 6.65 kPa was 88% sensitive and 74% specific (AUROC 0.87). For cirrhosis, a cutoff of 17.5 kPa was 100% sensitive and 97% specific (AUROC 0.99). Similarly, Kelleher reported that a TE cutoff of 8.7 kPa was 81% sensitive and 78% specific for septal fibrosis among 129 patients with NASH.

Several caveats warrant mention when discussing the utility of TE in patients with NAFLD. First, technical failure occurs in approximately 5% of patients, predominantly those with obesity/thick chest walls, which is almost universal in this patient population. In one study, the only independent risk factor for failure to obtain a TE measurement was a BMI greater than 28 kg/m² (odds ratio 10.0; 95% CI 5.7-17.9). Specific probes under development will hopefully minimize this limitation by sampling an acoustic window that is deeper within the abdomen. Second, the best TE thresholds for the delineation of fibrosis stages are unclear, and may be disease-dependent. For example, the optimal cutoffs reported in Friedrich-Rust and colleagues’ meta-analysis, which included predominantly patients with HCV, were 7.7 kPa and 13.0 kPa for septal fibrosis and cirrhosis, respectively. These figures are clearly different from those described above in studies restricted to NAFLD patients. Although these discrepancies may relate to differences in study populations and fibrosis classification systems between diseases, additional investigation is necessary to clarify this issue in patients with NAFLD (and all liver diseases). Finally, it is unknown whether the severity of hepatic steatosis has an impact on liver stiffness, as has been described for hepatic inflammation in the setting of acute and chronic viral hepatitis. In studies of potential liver donors and a recent description of NAFLD patients, steatosis was not associated with liver stiffness. On the contrary, among 324 patients with chronic HCV, the severity of hepatic steatosis was an independent predictor of liver stiffness. In light of these unresolved issues, and despite the widespread use of TE in patients with NAFLD, much further investigation is necessary to guide the optimal incorporation of this promising technology into routine clinical practice.

A technique related to TE, known as magnetic resonance elastography (MRE), also holds promise for the noninvasive assessment of liver fibrosis, including in patients with NAFLD. This tool uses a passive pneumatic driver placed over the abdomen to transmit acoustic pressure waves (60 Hz) through the liver. Using a standard MRI system and specialized software, an elastogram is produced, akin to that with TE, that quantifies liver stiffness. Advantages of MRE over conventional TE include: 1) the ability to scan the entire liver and avoid sampling error; 2) insensitivity to body habitus (i.e. thickness of the chest wall) for acquisition of an acoustic window; and 3) the ability to obtain conventional MRI at the same sitting. In a recent study describing the performance of MRE in 54 patients with NAFLD, Talwalkar et al. reported a significant association between liver stiffness values and fibrosis that was independent of BMI. For the detection of septal fibrosis (≥ F2), a threshold value ≥ 4.2 kPa had a sensitivity of 78%, specificity of 94%, and AUROC of 0.90. For cirrhosis, corresponding figures were 90%, 87%, and 0.94, respectively. Interestingly, ten individuals with hepatic steatosis on histology had liver stiffness values consistent with healthy controls (< 2.9 kPa), suggesting that steatosis itself has a minimal impact on liver stiffness assessed using MRE. Future studies will be necessary to validate these findings and confirm the cost-effectiveness of MRE before it gains widespread clinical use.

Conclusions

The growing prevalence of NAFLD has been likened to an epidemic affecting nearly one-third of Western populations. Although liver biopsy remains the gold standard for diagnosing NAFLD, differentiating simple steatosis from NASH, and staging the degree of hepatic fibrosis, it is clearly not an acceptable diagnostic tool for the vast majority of affected individuals. Thankfully, remarkable progress - particularly the development of serum fibrosis marker panels and transient elastography - has been made in the field of biomarkers of NAFLD dur-
ing the past decade. As effective therapies for NAFLD are under development, it is vital that simple, inexpensive, widely available, and accurate tools that are responsive to changes in NAFLD severity and correlated with clinical outcomes are developed. To achieve these goals, large, multicenter studies that include well-defined patient cohorts with adequate liver biopsy specimens and prospective follow-up will be critical.

Acknowledgements

Dr. Myers is supported by a Clinical Investigator Award from the Alberta Heritage Foundation for Medical Research (AHFMR) and a New Investigator Award from the Canadian Institutes of Health Research (CIHR).

References

RP Myers et al. Noninvasive diagnosis of nonalcoholic fatty liver disease