Transient elastography for the noninvasive assessment of liver fibrosis: A multicentre Canadian study

Robert P Myers MD MSc1, Magdy Elkashab MD2, Mang Ma MD3, Pam Crotty MSc1, Gilles Pomier-Layrargues MD4

BACKGROUND: Liver stiffness measurement (LSM) using transient elastography (TE) is a promising tool for the noninvasive assessment of hepatic fibrosis.

OBJECTIVES: To determine the feasibility and performance of TE in a North American cohort of patients with chronic liver disease.

METHODS: LSMs were obtained using TE in 260 patients with chronic hepatitis B or C, or nonalcoholic fatty liver disease from four Canadian hepatology centres. The accuracy of TE compared with liver biopsy for the prediction of significant fibrosis (Metavir fibrosis score of F2 or greater), bridging fibrosis (Metavir fibrosis score of F3 or greater) and cirrhosis (Metavir fibrosis score of F4 ) was assessed using area under ROC curves (AUROCs), and compared with the aspartate aminotransferase-to-platelet ratio index. The influence of alanine aminotransferase (ALT) levels and other factors on liver stiffness was determined using linear regression analyses.

RESULTS: Failure of TE occurred in 2.7% of patients, while liver biopsies were inadequate for staging in 0.8%. Among the remaining 251 patients, the AUROCs of TE for Metavir fibrosis scores of F2 and F3 or greater, and F4 were 0.74 (95% CI 0.68 to 0.80), 0.89 (95% CI 0.84 to 0.94), and 0.94 (95% CI 0.90 to 0.97), respectively. LSM was more accurate than the aminotransferase-to-platelet ratio index for bridging fibrosis (AUROC 0.78) and cirrhosis (AUROC 0.88), but not significant fibrosis (AUROC 0.76). At a cut-off of 11.1 kPa, the sensitivity, specificity, and positive and negative predictive values for cirrhosis (prevalence 11%) were 96%, 81%, 39% and 99%, respectively. For significant fibrosis (prevalence 53%), a cut-off of 7.7 kPa was 68% sensitive and 69% specific, and had a positive predictive value of 70% and a negative predictive value of 65%. Liver stiffness was independently associated with ALT, body mass index and steatosis. The optimal LSM cut-offs for cirrhosis were 11.1 kPa and 11.5 kPa in patients with ALT levels lower than 100 U/L and 100 U/L or greater, respectively. For fibrosis scores of F2 or greater, these figures were 7.0 kPa and 8.6 kPa, respectively.

CONCLUSIONS: The major role of TE is the exclusion of bridging fibrosis and cirrhosis. However, TE cannot replace biopsy for the diagnosis of significant fibrosis. Because liver stiffness may be influenced by significant ALT elevation, body mass index and/or steatosis, tailored liver stiffness cut-offs may be necessary to account for these factors.

Key Words: Biopsy; Cirrhosis; Elastography; Fatty liver; Fibrosis; Hepatitis

Liver fibrosis is the final common pathway for most liver diseases including viral, autoimmune and metabolic conditions. Pathologically, liver fibrosis is characterized by an accumulation of extracellular matrix with bands of fibrous tissue, which may progress to cirrhosis (1,2). The ability to accurately determine the extent of liver fibrosis is important for prognostication and treatment.
decision making regarding the initiation of therapy (eg, antivirals for hepatitis B [HBV] or hepatitis C [HCV]) and surveillance strategies (eg, for hepatocellular carcinoma [HCC] and esophageal varices in patients with cirrhosis). The accepted gold standard for assessing liver fibrosis is biopsy. Although biopsy provides important information, it is limited by invasiveness, cost and difficulty of repetition to monitor changes in fibrosis over time (3). Moreover, potentially serious, albeit uncommon, complications including hemorrhage and death may occur (4). Finally, the accuracy of liver biopsy is influenced by subjectivity in pathological interpretation and sampling error, which may lead to discordance in fibrosis staging in up to 40% of patients (5,6).

In light of these limitations, many alternative, noninvasive means for staging liver fibrosis have been investigated (7,8). Routine laboratory tests and abdominal imaging may diagnose cirrhosis in overt cases, but are insensitive to lesser stages of fibrosis. Several biomarker panels including surrogate markers of fibrosis and/or components of the extracellular matrix have also been developed (eg, FibroTest [BioPredictive, France], Enhanced Liver Fibrosis test [iQur, United Kingdom] and FibroMeter [BioLiveScale, France]) (9-11), but their proprietary nature and concerns regarding their performance have impacted their use. Other tools including high-throughput technologies (eg, proteomics, glycomics and metabolomics) (12,13) and novel imaging modalities (eg, magnetic resonance elastography) (14) are still under development. Transient elastography (TE; FibroScan, Echosens, France) is a novel, noninvasive tool for measuring liver stiffness as a surrogate of liver fibrosis (15,16). First studied in patients with HCV (17,18), TE has now been validated in other populations including patients with HBV (19), nonalcoholic fatty liver disease (NAFLD) (20) and cholestatic liver disorders (21). To date, results suggest that liver stiffness measurement (LSM) using TE is most accurate for the diagnosis of cirrhosis and less sensitive for earlier stages of fibrosis (22,23).

Advantages of TE include its simplicity, short performance time and immediate results, high patient acceptance and ease of incorporation into an outpatient clinical setting. Disadvantages include the difficulty of obtaining valid LSM in patients who are obese or have narrow intercostal spaces, the expense of the equipment (approximately $100,000) and the potential influence of nonfibrotic histological features on LSM (16). For example, overestimation of liver stiffness values has been reported in patients with acute (24) and chronic hepatic inflammation (25-27), extrahepatic cholestasis (28) and hepatic venous congestion (29). Despite these disadvantages, TE has become an important tool in clinical practice in many regions, most notably Europe, where it has been available since 2003. The FibroScan received licensing approval in Canada in August 2009.

Before the FibroScan’s approval in Canada, we undertook the present study to confirm the clinical utility of LSM using TE in a multicentre Canadian cohort. We focused on patients with chronic HCV, HBV and NAFLD, who represent the majority of hepatology referrals. Our primary objectives were to validate the performance of LSM for the noninvasive assessment of liver fibrosis and to identify predictors of TE performance. As a secondary objective, we examined the influence of serum alanine aminotransferase (ALT) on LSM and assessed whether ALT-specific cut-offs of liver stiffness are necessary to improve the performance of TE.

**METHODS**

**Study population**

The study cohort included adult patients (18 years of age and older) with chronic HBV or HCV infection, or suspected NAFLD who had undergone or were scheduled to undergo percutaneous liver biopsy within six months of LSM. Patients were recruited between March 2008 and November 2009 from the three following Canadian academic centres: the University of Calgary (Calgary, Alberta), the University of Alberta (Edmonton, Alberta) and Hôpital St Luc (Montréal, Quebec), and a community-based hepatology clinic (the Liver Centre [Toronto, Ontario]). Chronic HBV was defined as positivity for HBV surface antigen for at least six months and positive HBV DNA according to polymerase chain reaction assay. Chronic HCV was defined as positivity for anti-HCV antibodies and HCV RNA using polymerase chain reaction assay. All patients with viral hepatitis were treatment naive or had discontinued antiviral therapy for at least three months before LSM and liver biopsy. Suspected NAFLD was based on the presence of abnormal liver biochemistry, an echogenic liver on ultrasound and negative serological work-up for other causes of chronic liver disease (eg, autoimmune hepatitis, primary biliary cirrhosis, etc). Patients meeting any of the following exclusion criteria were ineligible: contraindications to LSM (eg, pregnancy, ascites, implantable cardiac devices, etc [15,16]); active extrahepatic infectious or inflammatory disorders or malignancy; significant alcohol consumption (20 g/day or more for women, and 30 g/day or more for men); body mass index (BMI) of greater than 40 kg/m²; and overt cirrhosis based on routine clinical and radiographic features. All patients in the present study provided informed consent, which was approved by Health Canada and the research ethics boards of the participating institutions.

**Clinical data**

Using a structured data collection instrument, demographic details (age, sex and race), BMI and history of diabetes mellitus were recorded. The following laboratory parameters, collected within six months of liver biopsy and LSM, were recorded where available: ALT, aspartate aminotransferase (AST), platelet count, HCV genotype and HBV e-antigen. The AST-to-platelet ratio index (APRI), an alternative and readily available noninvasive means of assessing liver fibrosis (30), was calculated as follows:

\[
\text{APRI} = \frac{100 \times \text{AST (upper limit of normal for AST} \times \text{platelets, } < 10^9/L)}{\text{ALT (upper limit of normal for ALT), } < 10^9/L} \]

**LSM using TE**

Operators who had completed at least 50 previous TE examinations performed all LSMs. Liver stiffness was evaluated using the FibroScan as per the manufacturer’s recommendations, with the patient lying in the dorsal decubitus position with the right arm in maximal abduction. Briefly, the tip of the transducer probe was covered with coupling gel and placed on the skin between the ribs over the right lobe of the liver. The operator, assisted by a time-motion sonographic image, located a portion of the liver at least 6 cm thick and free of large vascular structures. When a suitable target area had been located, an attempt was made to collect at least 10 valid measurements from a depth of 25 mm to 65 mm below the skin surface. Examinations that yielded no successful measurements after at
Liver biopsy

All liver biopsies were performed under ultrasound guidance. Liver tissue was fixed, paraffin embedded and stained, at minimum, with hematoxylin and eosin and saffron, and Masson’s trichrome or picrosirius red for collagen. Local pathologists at the participating institutions analyzed the liver biopsy specimens without knowledge of the TE results. Fibrosis in patients with HBV or HCV was staged according to the Metavir classification, which ranges from F0 to F4 (F0 = no fibrosis; F1 = portal fibrosis without septa; F2 = portal fibrosis with few septa; F3 = portal fibrosis with many septa; and F4 = cirrhosis) (33). In patients with NAFLD, fibrosis was staged according to the NAFLD Activity Score as per Kleiner et al. (34) (F0 = no fibrosis; F1 = perisinusoidal or portal fibrosis; F2 = perisinusoidal and portal/periportal fibrosis; F3 = septal or bridging fibrosis; and F4 = cirrhosis). To aggregate fibrosis data across diseases, Metavir stages were considered to be equivalent to those of the NAFLD Activity Score. Both fibrosis classification systems have demonstrated excellent intra- and interobserver agreement (33,34).

Hepatic steatosis was graded from S0 to S3 (S0, fewer than 5% of hepatocytes affected; S1, 5% to 33% of hepatocytes affected; S2, 34% to 66% of hepatocytes affected; and S3, more than 66% of hepatocytes affected) (34). In patients with HBV or HCV, necroinflammatory activity was graded according to the Metavir classification, which ranges from A0 to A3 (A0 = none; A1 = mild; A2 = moderate; and A3 = severe necroinflammation) (33). This classification system is not applicable to patients with NAFLD. Finally, the total length of liver biopsy specimens (in cm) was recorded as a measure of the quality of the reference standard for staging liver fibrosis.

Outcome measures

The primary outcome measures were the diagnosis of cirrhosis (F4 versus F0 to F3) and significant fibrosis (F2 to F4 versus F0 to F1). The identification of cirrhosis is important for decision making regarding surveillance for HCC and esophageal varices, while significant fibrosis is often considered to be a threshold for the initiation of antiviral therapy. As a secondary outcome, the identification of bridging fibrosis or cirrhosis was examined (F3 to F4 versus F0 to F2).

Statistical analyses

Patient characteristics and laboratory values were descriptively summarized and are reported as medians with IQRs and proportions. Patient subgroups were compared using Mann-Whitney and Kruskal-Wallis tests for continuous variables, and Fisher’s exact and χ² tests for categorical variables. Patient and histological factors potentially associated with liver stiffness (age, sex, BMI, underlying liver disease, ALT and AST levels, platelets and hepatic fibrosis, steatosis and necroinflammation) were examined using linear regression analyses. For these analyses, liver stiffness was logarithmically transformed due to its skewed distribution. Only variables that were significant in univariate analyses were included in multivariate models. To determine the diagnostic utility of TE for the noninvasive assessment of fibrosis, areas under ROC (AUROC) curves were calculated. The AUROC curve ranges between 0.5 and 1.0, with 0.5 equivalent to chance and 1.0 indicating a perfect test (35). In general, tests with an AUROC exceeding 0.80 are considered to be clinically useful (36). To determine factors associated with the accuracy of LSM for the primary histological outcomes, AUROCs in patient subgroups were compared according to the following a priori-defined patient-, biopsy- and TE-related factors: study centre, age (younger than 50 years versus 50 years or older [approximately the median]), sex, disease etiology, BMI (less than 25 kg/m², 25 kg/m² to 29.9 kg/m², and 30 kg/m² or greater), ALT (lower than 100 U/L versus 100 U/L or greater [approximately twice the upper limit of normal]), biopsy length (less than 1.5 cm, 1.5 cm to 2.5 cm, and greater than 2.5 cm), hepatic steatosis (S0 to S1 versus S2 to S3), necroinflammation (A2 to A3 versus A0 to A1 [in viral hepatitis patients only]), SR of LSM (less than 60% versus 60% or greater) and LSM IQR/M (less than 30% versus 30% or greater). AUROCs were compared using the method of DeLong et al. (37).

Because the AUROC curve is difficult to translate into clinical practice, we also calculated the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), and overall accuracy of TE with exact binomial 95% CI. For these analyses, optimal liver stiffness cut-off values in the study cohort that maximized the sum of sensitivity and specificity were determined. The results obtained with these thresholds were compared with those obtained using thresholds previously published by Castera et al. (18), Ziol et al. (17) and Friedman-Rust et al. (22). All analyses were conducted using Stata/SE version 11.0 (StataCorp LP, USA). A two-sided P<0.05 was considered to be statistically significant.

RESULTS

Patient characteristics

In total, 260 patients were enrolled at the four participating centres. Nine patients (3.4%) were excluded due to inadequate biopsies for histological interpretation in two patients (0.8%) and failure of LSM in seven (2.7%). All patients with LSM failure were overweight or obese. LSM failure occurred in 0% of patients with a BMI of less than 25 kg/m², 2.4% of overweight patients (BMI 25 kg/m² to 29.9 kg/m²), 4.0% of mildly obese patients (BMI 30 kg/m² to 34.9 kg/m²) and 10% of severely obese patients (BMI of greater than 35 kg/m²) (P=0.02). The characteristics of the final study cohort (n=251) are summarized in Table 1. The majority of the patients (66%) were men, the median age of the cohort was 49 years (IQR 42 to 55 years) and 60% were Caucasian. Fifty-three per cent of patients had chronic HCV, 27% had chronic HBV and 20% had NAFLD. Among patients with HCV, 72% were infected with genotype 1, while 21% had HCV genotypes 2 or 3. A single HCV patient was HIV coinfected and another had previously undergone a liver transplant. The majority (67%) of HBV patients (n=63 with available serology) were HBV e-antigen negative. The median BMI was 26 kg/m² (IQR 23 kg/m² to 30 kg/m²); 26% of patients...
TABLE 1
Characteristics of the study cohort according to study centre

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall (n=251)</th>
<th>Calgary, Alberta (n=74)</th>
<th>Edmonton, Alberta (n=35)</th>
<th>Montreal, Quebec (n=97)</th>
<th>Toronto, Ontario (n=45)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female sex</td>
<td>34 (65)</td>
<td>38 (28)</td>
<td>40 (14)</td>
<td>29 (28)</td>
<td>33 (19)</td>
<td>0.53</td>
</tr>
<tr>
<td>Age, years</td>
<td>49 (42–55)</td>
<td>50 (42–55)</td>
<td>52 (39–55)</td>
<td>49 (43–55)</td>
<td>49 (39–56)</td>
<td>1.00</td>
</tr>
<tr>
<td>Race*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>60 (150)</td>
<td>49 (36)</td>
<td>40 (14)</td>
<td>78 (76)</td>
<td>53 (24)</td>
<td>&lt;0.0005</td>
</tr>
<tr>
<td>Asian</td>
<td>30 (76)</td>
<td>43 (32)</td>
<td>46 (16)</td>
<td>19 (18)</td>
<td>22 (10)</td>
<td></td>
</tr>
<tr>
<td>Other/unknown</td>
<td>10 (25)</td>
<td>8 (6)</td>
<td>14 (5)</td>
<td>3 (3)</td>
<td>24 (11)</td>
<td></td>
</tr>
<tr>
<td>Etiology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hepatitis C</td>
<td>53 (133)</td>
<td>53 (39)</td>
<td>40 (14)</td>
<td>72 (70)</td>
<td>22 (10)</td>
<td>&lt;0.0005</td>
</tr>
<tr>
<td>Hepatitis B</td>
<td>27 (68)</td>
<td>36 (23)</td>
<td>43 (15)</td>
<td>16 (16)</td>
<td>22 (10)</td>
<td></td>
</tr>
<tr>
<td>Nonalcoholic fatty liver disease</td>
<td>20 (50)</td>
<td>11 (8)</td>
<td>17 (6)</td>
<td>11 (11)</td>
<td>56 (25)</td>
<td></td>
</tr>
<tr>
<td>BMI, kg/m²‡</td>
<td>26 (23–30)</td>
<td>24 (22–29)</td>
<td>26 (24–30)</td>
<td>25 (23–28)</td>
<td>29 (26–33)</td>
<td>0.0008</td>
</tr>
<tr>
<td>Obesity (BMI ≥30 kg/m²)†</td>
<td>26 (65)</td>
<td>23 (17)</td>
<td>26 (9)</td>
<td>19 (18)</td>
<td>47 (21)</td>
<td>0.005</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>8 (20)</td>
<td>8 (6)</td>
<td>6 (2)</td>
<td>5 (5)</td>
<td>16 (7)</td>
<td>0.19</td>
</tr>
<tr>
<td>ALT, U/L†</td>
<td>61 (39–92)</td>
<td>68 (44–93)</td>
<td>61 (37–77)</td>
<td>65 (45–114)</td>
<td>49 (35–76)</td>
<td>0.17</td>
</tr>
<tr>
<td>AST, U/L‡</td>
<td>44 (32–68)</td>
<td>45 (30–69)</td>
<td>40 (30–65)</td>
<td>48 (34–76)</td>
<td>40 (32–56)</td>
<td>0.23</td>
</tr>
<tr>
<td>Platelets, × 10⁹/L †</td>
<td>222 (176–266)</td>
<td>223 (162–266)</td>
<td>202 (178–247)</td>
<td>215 (163–254)</td>
<td>253 (201–281)</td>
<td>0.02</td>
</tr>
<tr>
<td>APRI†</td>
<td>0.53 (0.35–0.90)</td>
<td>0.56 (0.37–1.07)</td>
<td>0.60 (0.35–0.97)</td>
<td>0.60 (0.36–1.12)</td>
<td>0.40 (0.31–0.60)</td>
<td>0.02</td>
</tr>
<tr>
<td>Liver fibrosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F0</td>
<td>14 (36)</td>
<td>19 (14)</td>
<td>23 (8)</td>
<td>2 (2)</td>
<td>27 (12)</td>
<td>&lt;0.0005</td>
</tr>
<tr>
<td>F1</td>
<td>33 (82)</td>
<td>39 (29)</td>
<td>31 (11)</td>
<td>27 (256)</td>
<td>36 (16)</td>
<td></td>
</tr>
<tr>
<td>F2</td>
<td>33 (82)</td>
<td>27 (20)</td>
<td>23 (8)</td>
<td>48 (47)</td>
<td>16 (7)</td>
<td></td>
</tr>
<tr>
<td>F3</td>
<td>9 (23)</td>
<td>8 (6)</td>
<td>14 (5)</td>
<td>7 (7)</td>
<td>11 (5)</td>
<td></td>
</tr>
<tr>
<td>F4</td>
<td>11 (28)</td>
<td>7 (5)</td>
<td>9 (3)</td>
<td>15 (15)</td>
<td>11 (5)</td>
<td></td>
</tr>
<tr>
<td>Biopsy length, cm§</td>
<td>2.4 (1.7–2.8)</td>
<td>2.2 (1.7–2.8)</td>
<td>2.4 (2.1–2.8)</td>
<td>2.7 (2.2–3.0)</td>
<td>1.6 (1.4–1.7)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Liver stiffness, kPa</td>
<td>7.7 (5.3–11.6)</td>
<td>6.9 (5.1–13.6)</td>
<td>9.9 (6.0–13.8)</td>
<td>7.2 (5.4–10.3)</td>
<td>7.6 (5.9–10.4)</td>
<td>0.23</td>
</tr>
<tr>
<td>Success rate, %</td>
<td>100 (91–100)</td>
<td>100 (91–100)</td>
<td>91 (67–100)</td>
<td>100 (92–100)</td>
<td>100 (91–100)</td>
<td>0.0003</td>
</tr>
<tr>
<td>IQR/median stiffness</td>
<td>0.16 (0.10–0.24)</td>
<td>0.20 (0.12–0.26)</td>
<td>0.18 (0.13–0.26)</td>
<td>0.14 (0.08–0.21)</td>
<td>0.16 (0.10–0.22)</td>
<td>0.008</td>
</tr>
</tbody>
</table>

Data presented as % (n) or median (interquartile range [IQR]) unless indicated otherwise. P values for comparisons of patient variables across study centres calculated using χ² tests for categorical variables and Kruskal-Wallis rank tests for continuous variables. *As reported by health care providers at study centres. Race status missing in three patients; †Body mass index (BMI) data missing in one patient; ‡Alanine aminotransferase (ALT), aspartate aminotransferase (AST), platelets and AST to platelet ratio index (APRI) data missing in one, 22, two and 24 patients, respectively; §Biopsy specimen length data missing in five patients.

were obese (BMI of 30 kg/m² or greater). There were significant differences in race, liver disease etiology, BMI, liver biochemistry, fibrosis stage and liver biopsy length between study centres (Table 1).

Liver histology and liver stiffness using TE
The median length of biopsy specimens was 2.4 cm (IQR 1.7 cm to 2.8 cm); 87% of biopsies were at least 1.5 cm long. In total, 53% of patients had significant fibrosis (F2 to F4), 20% had bridging fibrosis or cirrhosis (F3 to F4) and 11% had cirrhosis (F4) (Table 1). The severity of steatosis (data available for 232 patients [92%]) was mild (S0 to S1: 33% of hepatocytes or fewer affected) in 83% of patients, moderate (S2: 34% to 66%) in 13% and severe (S3: greater than 66%) in 4%. Nécroinflammatory activity was moderate to severe (A2 to A3) in 65% of patients with HCV and 46% with HBV.

The median interval between liver biopsy and LSM was 18 days (range zero to 183 days). The SR of TE was less than 10 valid LSMs and the IQR/M was 30% or greater in 11.1% of cases (n=28). Therefore, after excluding patients with TE failure, unreliable LSM results were observed in 17% (n=42). Liver stiffness ranged from 2.5 kPa to 75 kPa, with a median of 7.7 kPa (IQR 5.3 kPa to 11.6 kPa). As illustrated in Figure 1, there was a substantial overlap of liver stiffness values, particularly at lower fibrosis stages (F0 to F2). The median (IQR) liver stiffness among patients with Metavir F0, F1, F2, F3 and F4 fibrosis were 5.5 kPa (4.0 to 7.7), 6.3 kPa (4.7 to 9.0), 7.7 kPa (5.7 to 10.4), 12.0 kPa (8.3 to 17.3) and 24.3 kPa (13.7 to 34.6), respectively. In univariate analysis, liver stiffness was associated with older age, male sex, HCV (versus HBV) infection, lower platelet count and higher BMI, ALT, AST, and hepatic fibrosis and steatosis (P<0.05 for all variables). Liver stiffness was not associated with necroinflammatory activity (P=0.35). In a multivariate model that adjusted for fibrosis stage (P<0.05 for each of F2, F3 and F4 versus F0) and excluded AST due to collinearity with ALT, moderate to severe steatosis (P=0.01) and higher BMI (P<0.0005) were significantly associated with liver stiffness. Serum ALT was also significant (P=0.009). The relationships between ALT and steatosis with LSM are illustrated in Figure 2. After controlling for fibrosis stage, differences in liver stiffness according to ALT were small except for patients with bridging fibrosis (F3; Figure 2A). In these patients, the median (IQR) liver stiffness was 16.8 kPa (13.6 kPa to 25.4 kPa) in those with ALT levels of 100 U/L or greater (n=6) versus 10.3 kPa (8.1 kPa to 14.0 kPa) in patients with ALT levels of lower than 100 U/L (n=17; P=0.046; Figure 2A). Steatosis-related differences in LSM were most marked in patients with F0 to F2 fibrosis (P<0.05; Figure 2B).
Diagnostic accuracy of TE for the prediction of liver fibrosis

Figure 3 shows ROC curves of the diagnostic accuracy of TE for staging fibrosis compared with liver biopsy. The AUROCs for significant fibrosis (F2 or greater), bridging fibrosis or cirrhosis (F3 or greater) and cirrhosis (F4) were 0.74 (95% CI 0.68 to 0.80), 0.89 (95% CI 0.84 to 0.94) and 0.94 (95% CI 0.90 to 0.97), respectively. Table 2 summarizes the operating characteristics of LSM for these outcomes according to optimal cut-offs derived in the cohort, both overall and according to disease etiology. For cirrhosis, a cut-off of 11.1 kPa or greater was highly sensitive (96%; 95% CI 82% to 100%) and specific (81%; 75% to 86%). At a cirrhosis prevalence of 11%, liver stiffness values below this threshold excluded cirrhosis (NPV) with 99% (95% CI 97% to 100%) certainty. Only one cirrhotic patient, who had an LSM of 9.8 kPa, would have been misdiagnosed with noncirrhotic fibrosis according to this cut-off. However, liver stiffness values of 11.1 kPa or greater were poorly predictive of cirrhosis (PPV 39%; 95% CI 28% to 52%).

The diagnostic performance of a cut-off of 7.7 kPa was similar to that of thresholds of 7.1 kPa and 8.8 kPa, as recommended by Castera et al (18) and Ziol et al (17), respectively (Table 3).

Optimal liver stiffness cut-offs according to serum ALT

Because liver stiffness was significantly associated with serum ALT, sensitivity analyses were conducted to determine whether the optimal liver stiffness cut-offs varied according to ALT. For the diagnosis of cirrhosis in patients with ALT levels below 100 U/L (n=196), the optimal cut-off was 11.1 kPa or greater, as observed in our primary analysis. In patients with ALT levels of 100 U/L or greater, the optimal threshold was slightly higher at 11.5 kPa or greater (n=54). The diagnostic performance of these thresholds did not differ substantially (Table 4). For the diagnosis of significant fibrosis in patients with ALT levels of below 100 U/L, the optimal cut-off was 7.0 kPa or greater, similar to our primary analysis. In patients with ALT levels of 100 U/L or greater, the optimal threshold was 8.6 kPa or greater. The specificity (92%) and PPV (97%) of the latter cut-off were excellent, while the NPV was suboptimal (44%).

Predictive factors of TE performance

The impact of patient-, biopsy- and TE-related factors on the performance of LSM for the identification of cirrhosis (F4) and

Figure 2) Influence of serum alanine aminotransferase (ALT) level (lower than 100 U/L versus 100 U/L or higher) (A) and steatosis grade on liver stiffness stratified according to Metavir fibrosis (F) stage (B)

Figure 3) Receiver operating characteristic curves illustrating the performance of liver stiffness measurement using transient elastography for the diagnosis of significant fibrosis (Metavir F2 or greater), bridging fibrosis or cirrhosis (Metavir F3 or greater), and cirrhosis (Metavir F4). AUROC Area under ROC curve

Can J Gastroenterol Vol 24 No 11 November 2010
TABLE 2
Diagnostic accuracy of transient elastography at optimal cut-off values for the staging of liver fibrosis

<table>
<thead>
<tr>
<th>Histological outcome/cut-off value</th>
<th>AUROC (95% CI)</th>
<th>Sensitivity, % (95% CI)</th>
<th>Predictive value, % (95% CI)</th>
<th>Accuracy, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Significant fibrosis (F2)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥7.7 kPa (Optimal cut-off)†</td>
<td>0.74 (0.68–0.80)</td>
<td>68 (59–76)</td>
<td>71 (62–79)</td>
<td>68 (62–74)</td>
</tr>
<tr>
<td>≥7.1 kPa (Castera et al [18])</td>
<td>0.73 (0.64–0.82)</td>
<td>65 (54–75)</td>
<td>71 (62–79)</td>
<td>65 (57–73)</td>
</tr>
<tr>
<td>≥8.8 kPa (Ziol et al [17])</td>
<td>0.79 (0.63–0.87)</td>
<td>61 (42–78)</td>
<td>78 (62–90)</td>
<td>71 (58–81)</td>
</tr>
<tr>
<td>Nonalcoholic fatty liver disease</td>
<td>0.84 (0.73–0.95)</td>
<td>94 (71–100)</td>
<td>55 (36–74)</td>
<td>72 (58–84)</td>
</tr>
<tr>
<td>bridging fibrosis or cirrhosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥10.3 kPa (Optimal cut-off)‡</td>
<td>0.89 (0.80–0.94)</td>
<td>82 (69–92)</td>
<td>82 (77–88)</td>
<td>83 (77–87)</td>
</tr>
<tr>
<td>≥9.5 kPa (Castera et al [18])</td>
<td>0.92 (0.83–0.95)</td>
<td>88 (72–97)</td>
<td>81 (72–88)</td>
<td>83 (75–89)</td>
</tr>
<tr>
<td>≥9.6 kPa (Ziol et al [17])</td>
<td>0.99 (0.92–1.00)</td>
<td>100 (29–100)</td>
<td>50 (21–79)</td>
<td>88 (79–95)</td>
</tr>
<tr>
<td>Cirrhosis (F4)</td>
<td>0.82 (0.67–0.97)</td>
<td>70 (35–93)</td>
<td>64 (40–70)</td>
<td>76 (62–87)</td>
</tr>
</tbody>
</table>

*Number in brackets after disease etiologies indicate the number of patients with the specified histological outcome, not the total number of patients. †One-sided 97.5% CI for point estimates of 100%. AUROC Area under ROC curve

TABLE 3
Comparison of the diagnostic accuracy of transient elastography at optimal and published cut-off values for the staging of liver fibrosis

<table>
<thead>
<tr>
<th>Histological outcome/cut-off value</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive</th>
<th>Negative</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant fibrosis (F2)*</td>
<td>68 (59–76)</td>
<td>69 (59–77)</td>
<td>71 (62–79)</td>
<td>65 (56–74)</td>
<td>68 (62–74)</td>
</tr>
<tr>
<td>bridging fibrosis or cirrhosis</td>
<td>82 (69–92)</td>
<td>83 (77–88)</td>
<td>55 (43–67)</td>
<td>95 (90–98)</td>
<td>83 (77–87)</td>
</tr>
<tr>
<td>Cirrhosis (F4)</td>
<td>84 (71–93)</td>
<td>75 (68–81)</td>
<td>46 (35–57)</td>
<td>95 (90–98)</td>
<td>77 (71–82)</td>
</tr>
</tbody>
</table>

Data presented as % (95% CI). *For significant fibrosis (F2), Friedrich-Rust et al (22) also reported an optimal cut-off value of 7.7 kPa (22); †Optimal cut-off values determined in study cohort; ‡For bridging fibrosis or cirrhosis (F3), Friedrich-Rust et al (22) did not report an optimal cut-off value

Comparison of TE with the APRI and the Fib-4 test for the prediction of liver fibrosis

In total, 227 patients (90% of the cohort) had a valid LSM and complete data for calculation of the APRI. TE was more accurate than the APRI for the diagnosis of cirrhosis (AUROCs 0.94 [95% CI 0.90 to 0.97] versus 0.88 [95% CI 0.83 to 0.94]; P=0.046) and bridging fibrosis or cirrhosis (AUROCs 0.89 [95% CI 0.84 to 0.94] versus 0.78 [95% CI 0.70 to 0.86]; P=0.007). However, the AUROCs did not differ between TE and the APRI for the prediction of significant fibrosis (F2 or greater) (0.73 [95% CI 0.67 to 0.80] versus 0.76 [95% CI 0.69 to 0.82]; P=0.55).

DISCUSSION

The burden of chronic liver disease in North America is growing due to, among other factors, aging of the HCV-infected population (38) and an epidemic of NAFLD due to the rising...
Transverse elastography for assessing liver fibrosis

Can J Gastroenterol Vol 24 No 11 November 2010 667

prevalence of obesity and diabetes mellitus (39). To manage this expanding pool of patients, accurate, noninvasive tools for staging liver fibrosis are needed. Our study is the first to describe the feasibility and diagnostic performance of TE for the noninvasive assessment of liver fibrosis in a Canadian population. In our cohort, failure to obtain a valid LSM was uncommon, occurring in only 3% of patients. This finding is consistent with previous reports that documented TE failure in 3% to 5% of cases (31,32,40). The major causes of TE failure in the present and other studies was overweight or obesity. Whereas valid LSMs were obtained in all patients with a BMI of less than 25 kg/m², failures occurred in 2.4% of overweight patients (BMI 25 kg/m² to 29.9 kg/m²) and 6% of the obese patients (BMI 30 kg/m² or greater). Because approximately one-half of North Americans are overweight, the development of a specific ‘obese probe’ will hopefully reduce the incidence of TE failure (41). A multicentre Canadian trial (42) is underway to validate the utility of this probe.

We also evaluated the frequency of unreliable TE results, characterized by large variability in liver stiffness values (IQR/M 30% or greater) or a limited number of successful acquisitions (fewer than 10 valid measurements or SR of less than 60%), according to the FibroScan manufacturer’s quality criteria (31,32). Unreliable TE results were recorded in 17% of patients, as reported by Castera et al (31) in more than 13,000 examinations. Interestingly, the diagnostic performance of TE, as assessed by the AUROC, did not differ according to these criteria, either individually or collectively. In a study that examined predictors of agreement between TE and liver biopsy, Lucidarme et al (32) observed an independent association between IQR/M and discordance of at least two fibrosis stages. However, the AUROCs for significant fibrosis and cirrhosis in this study did not differ between patients with IQR/M values above and below the most discriminative IQR/M threshold (0.21 or greater). Although confirmation in larger cohorts is necessary, our data suggest that TE results in patients who do not satisfy these criteria should not be automatically regarded as invalid. However, because TE is a new technique – particularly in North America – these unreliable TE results should be interpreted with caution.

The primary objective of our study was to examine the diagnostic performance of LSM using TE for the staging of liver fibrosis. In our cohort, the AUROCs for bridging fibrosis (F3 or greater) or a limited number of successful acquisitions (fewer than 10 valid measurements or SR of less than 60%), according to the FibroScan manufacturer’s quality criteria (31,32). Unreliable TE results were recorded in 17% of patients, as reported by Castera et al (31) in more than 13,000 examinations. Interestingly, the diagnostic performance of TE, as assessed by the AUROC, did not differ according to these criteria, either individually or collectively. In a study that examined predictors of agreement between TE and liver biopsy, Lucidarme et al (32) observed an independent association between IQR/M and discordance of at least two fibrosis stages. However, the AUROCs for significant fibrosis and cirrhosis in this study did not differ between patients with IQR/M values above and below the most discriminative IQR/M threshold (0.21 or greater). Although confirmation in larger cohorts is necessary, our data suggest that TE results in patients who do not satisfy these criteria should not be automatically regarded as invalid. However, because TE is a

### Table 4: Optimal liver stiffness cut-off values according to serum alanine aminotransferase (ALT) activity

<table>
<thead>
<tr>
<th>Histological outcome</th>
<th>ALT activity</th>
<th>Optimal cut-off, kPa</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Positive predictive value (95% CI)</th>
<th>Negative predictive value (95% CI)</th>
<th>Accuracy (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant fibrosis (≥F2)</td>
<td>&lt;100 U/L (n=196)</td>
<td>≥7.0</td>
<td>70 (60–79)</td>
<td>64 (54–73)</td>
<td>63 (53–72)</td>
<td>71 (61–80)</td>
<td>67 (60–73)</td>
</tr>
<tr>
<td></td>
<td>≥100 U/L (n=54)</td>
<td>≥8.6</td>
<td>67 (50–80)</td>
<td>92 (62–100)</td>
<td>97 (82–100)</td>
<td>44 (24–65)</td>
<td>72 (58–84)</td>
</tr>
<tr>
<td>Cirrhosis (F4)</td>
<td>&lt;100 U/L (n=196)</td>
<td>≥11.1</td>
<td>94 (73–100)</td>
<td>84 (77–89)</td>
<td>37 (23–52)</td>
<td>98 (96–100)</td>
<td>85 (79–89)</td>
</tr>
<tr>
<td></td>
<td>≥100 U/L (n=54)</td>
<td>≥11.5</td>
<td>100 (70–100)</td>
<td>73 (57–85)</td>
<td>45 (24–68)</td>
<td>100 (89–100)</td>
<td>78 (64–89)</td>
</tr>
</tbody>
</table>

Data presented as % (95% CI) unless otherwise indicated. *Optimal cut-off values are those with the maximal sum of sensitivity and specificity. ALT data missing in one patient. †One-sided 97.5% CI for point estimates of 100%
liver stiffness above this threshold (39%), a liver biopsy is necessary to ‘rule in’ cirrhosis in the absence of a strong clinical suspicion (e.g., thrombocytopenia or suggestive radiographic evidence). The optimal cut-off value of 11.1 kPa in our cohort is within the range recommended by other authors (17,18,22); diagnostic performance was generally similar between these thresholds (Table 3). Discrepancies among thresholds reported in different studies likely reflect differences in patient populations, underlying prevalence of fibrosis (‘spectrum bias’) and histological interpretation.

The utility of TE for diagnosing significant fibrosis was suboptimal, as demonstrated by an AUROC of only 0.74 (95% CI 0.68 to 0.80), not significantly different from that of the APRI. The optimal liver stiffness cut-off value (7.7 kPa) was 68% sensitive and 69% specific for this end point. This performance is slightly lower than that reported in the meta-analysis by Friedrich-Rust et al (22) (summary AUROC 0.84; 95% CI 0.82 to 0.86). Interestingly, the optimal threshold of both studies was 7.7 kPa, although heterogeneity between studies precluded the authors from reporting the sensitivity and specificity of this cut-off. The major limitation of TE for identifying significant fibrosis relates to the large overlap in liver stiffness among patients with stages F0 to F2 fibrosis (Figure 1). Median liver stiffness was very similar between these groups (F0, 5.5 kPa versus F1; 6.3 kPa versus F2, 7.7 kPa). Only at F3 (12.0 kPa) or F4 (24.3 kPa) fibrosis was a substantial increase in median liver stiffness observed. Although a limitation of TE from a diagnostic viewpoint, this finding suggests that TE is actually more reflective of liver fibrosis than conventional histological staging because fibrous tissue accumulation has a nonlinear relationship with the Metavir stage. As described by Bedossa et al (43), the areas of fibrosis measured by digital image analysis in patients with F2, F3 and F4 fibrosis were three-, seven- and 12-fold greater, respectively, than in patients with F0 fibrosis. Whether LSM is actually more effective than histology for the assessment of disease progression and prognosis is being addressed in longitudinal studies. Nevertheless, based on our results, TE is not yet ready to replace liver biopsy for differentiating F2 or greater fibrosis from lesser stages.

A potential limitation of our study that may have contributed to the suboptimal results of TE for predicting significant fibrosis is our inclusion of local pathologists rather than a centralized, expert pathologist. This approach was taken to ensure the feasibility of this nonfunded, investigator-initiated study. In addition to sampling error, this variability in histological interpretation may have contributed to an ‘imperfect gold standard bias’ (44,45). As illustrated in Figure 1, liver stiffness values exceeding 20 kPa – clearly within the cirrhotic range – were observed among several outliers with F0 to F2 fibrosis. Although these discrepancies are suggestive of liver biopsy error, further analyses are necessary to attribute the cause of these discordances to biopsy or TE. Interestingly, the AUROC of TE for cirrhosis (but not significant fibrosis) among patients with liver biopsies shorter than 1.5 cm was greater than among those with longer biopsy specimens (Figure 4). Previous studies have suggested that smaller biopsy specimens are associated with reduced accuracy for surrogate markers when liver biopsy is used as the reference standard (20,46,47). Although our finding may represent a type 1 (false positive) error, it justifies our decision not to exclude patients with biopsies shorter than 1.5 cm (13% of the cohort) to maximize our sample size.

Lever stiffness may be influenced by factors other than fibrosis. Indeed, spursiously elevated LSMs have been reported in patients with acute flares of viral hepatitis (24), necroinflammation in chronic viral hepatitis (25-27), extrahepatic cholestasis (28) and hepatic congestion (29). In this regard, we observed significant associations between liver stiffness and serum ALT, BMI and steatosis, but not necroinflammatory activity, after controlling for fibrosis stage (Figure 2). According to our regression models, for every 100 unit increase in ALT, liver stiffness increased by 1.1 kPa. Similarly, on average, liver stiffness was 1.3 kPa higher in patients with steatosis exceeding 33% than patients with lesser grades. Although controversial in that most studies have not found an association between liver stiffness and steatosis or BMI (17,20,25), these findings are potentially clinically relevant with respect to the choice of optimal LSM cut-off values and the estimation of
fibrosis in an individual patient. For example, the optimal cut-off values for significant fibrosis and cirrhosis were 1.4 kPa and 0.6 kPa higher among patients with ALT levels of 100 U/L or greater versus patients with ALT levels lower than 100 U/L, respectively. Although an ALT-specific threshold had a small impact on the operating characteristics of TE for cirrhosis, significant improvements in specificity and PPV were observed for significant fibrosis (Table 4). However, because the overall AUROCs of TE did not differ according to ALT, larger studies are needed to define optimal LSM cut-off values according to this factor.

Our study had several limitations. First, only patients with HBV, HCV and NAFLD were included. Although the performance of TE appears to be similar in other hepatic conditions (e.g., cholestatic liver disease [21], post-transplantation [48], etc.), additional studies in other diseases are necessary. Second, due to the recent availability of the FibroScan in Canada, TE examinations were performed by relatively novice operators. Although some studies suggest that a novice can obtain reliable TE results after a minimal training period of approximately 50 examinations (as required for participation in the present study) (49,50), emerging data are contradictory. For example, Castera et al (31) reported that operators who had performed fewer than 50 TE examinations had a 2.5-fold increased odds of obtaining unreliable LSM results compared with more experienced operators. Counter to this point, performance of TE was similar between the first and latter one-half of the study (data not shown). Finally, some of our statistical comparisons were likely underpowered including our analyses of predictors of TE performance and the identification of optimal liver stiffness thresholds in specific patient subgroups.

CONCLUSION

The major strength of LSM using TE is the exclusion of bridging fibrosis and cirrhosis. However, TE is not yet ready to replace liver biopsy for the confirmation of advanced fibrosis or the diagnosis of significant fibrosis. Liver stiffness may be influenced by hepatic steatosis, BMI and significant ALT elevation. Additional studies in larger cohorts are necessary to define the optimal liver stiffness thresholds that account for these potential influences. The prognostic significance of liver stiffness and the role of TE in the longitudinal follow-up of patients with chronic liver disease (e.g., as a screening tool for esophageal varices) should also be addressed.

DISCLOSURE: Dr Myers is supported by a Clinical Investigator Award from the Alberta Heritage Foundation for Medical Research, and a New Investigator Award from the Canadian Institutes for Health Research. Dr Myers has received research support from Echosens, the makers of the FibroScan, and lecture fees from KNS Canada, the distributors of the FibroScan in Canada.

CONFLICTS OF INTEREST: The authors have no conflicts of interest to declare.

REFERENCES