

The feasibility and reliability of transient elastography using Fibroscan®: A practice audit of 2335 examinations

Jack XQ Pang MD^{1,2}, Faruq Pradhan¹, Scott Zimmer BA³, Sophia Niu BSc MSc³, Pam Crotty MSc¹, Jenna Tracey BSc¹, Christopher Schneider MD¹, Steven J Heitman MD MSc^{1,2}, Gilaad G Kaplan MD MPH^{1,2}, Mark G Swain MD MSc¹, Robert P Myers MD MSc^{1,2}

JXQ Pang, F Pradhan, S Zimmer, et al. The feasibility and reliability of transient elastography using Fibroscan®: A practice audit of 2335 examinations. *Can J Gastroenterol Hepatol* 2014;28(3):143-149.

BACKGROUND: Liver stiffness measurement (LSM) using transient elastography is widely used in the management of patients with chronic liver disease.

OBJECTIVES: To examine the feasibility and reliability of LSM, and to identify patient and operator characteristics predictive of poorly reliable results.

METHODS: The present retrospective study investigated the frequency and determinants of poorly reliable LSM (interquartile range [IQR]/median LSM [IQR/M] >30% with median liver stiffness ≥ 7.1 kPa) using the FibroScan (Echosens, France) over a three-year period. Two experienced operators performed all LSMs. Multiple logistic regression analyses examined potential predictors of poorly reliable LSMs including age, sex, liver disease, the operator, operator experience (<500 versus ≥ 500 scans), FibroScan probe (M versus XL), comorbidities and liver stiffness. In a subset of patients, medical records were reviewed to identify obesity (body mass index ≥ 30 kg/m²).

RESULTS: Between July 2008 and June 2011, 2335 patients with liver disease underwent LSM (86% using the M probe). LSM failure (no valid measurements) occurred in 1.6% (n=37) and was more common using the XL than the M probe (3.4% versus 1.3%; P=0.01). Excluding LSM failures, poorly reliable LSMs were observed in 4.9% (n=113) of patients. Independent predictors of poorly reliable LSM included older age (OR 1.03 [95% CI 1.01 to 1.05]), chronic pulmonary disease (OR 1.58 [95% CI 1.05 to 2.37]), coagulopathy (OR 2.22 [95% CI 1.31 to 3.76]) and higher liver stiffness (OR per kPa 1.03 [95% CI 1.02 to 1.05]), including presumed cirrhosis (stiffness ≥ 12.5 kPa; OR 5.24 [95% CI 3.49 to 7.89]). Sex, diabetes, the underlying liver disease and FibroScan probe were not significant. Although reliability varied according to operator (P<0.0005), operator experience was not significant. In a subanalysis including 434 patients with body mass index data, obesity influenced the rate of poorly reliable results (OR 2.93 [95% CI 0.95 to 9.05]; P=0.06).

CONCLUSIONS: FibroScan failure and poorly reliable LSM are uncommon. The most important determinants of poorly reliable results are older age, obesity, higher liver stiffness and the operator, the latter emphasizing the need for adequate training.

Key Words: Biopsy; Diagnostic test; Fibrosis; Hepatitis; Stiffness

The evaluation of liver fibrosis is a vital component of the management of patients with chronic liver disease, both for guiding therapy and estimating prognosis. Traditionally, liver biopsy had been used for this purpose (1); however, biopsy is limited by invasiveness, cost, variability in histological interpretation, sampling error (2,3) and difficulty of repetition for monitoring changes in fibrosis over time.

La faisabilité et la fiabilité de l'élastographie transitoire par Fibroscan® : la vérification pratique de 2 335 examens

HISTORIQUE : La mesure de rigidité du foie (MRF) par élastographie transitoire est largement utilisée dans la prise en charge des patients atteints d'une maladie hépatique chronique.

OBJECTIFS : Examiner la faisabilité et la fiabilité de la MRF et déterminer les caractéristiques des patients et des opérateurs qui prédisent des résultats peu fiables.

MÉTHODOLOGIE : La présente étude rétrospective portait sur la fréquence et les déterminants d'une MRF peu fiable (plage interquartile [PIQ]/MRF médiane [PIQ/M] >30 % et rigidité hépatique médiane $\geq 7,1$ kPa) au moyen du FibroScan (Echosens, France) sur une période de trois ans. Deux opérateurs expérimentés ont effectué toutes les MRF. Par des analyses de régression logistique multiples, les chercheurs ont examiné les prédicteurs potentiels de MRF peu fiables, y compris l'âge, le sexe, la maladie hépatique, l'opérateur, l'expérience de l'opérateur (<500 ou ≥ 500 échographies), la sonde du FibroScan (moyenne plutôt que très grande), les comorbidités et la rigidité du foie. Ils ont passé en revue les dossiers médicaux d'un sous-groupe de patients pour déterminer la présence d'obésité (indice de masse corporelle ≥ 30 kg/m²).

RÉSULTATS : Entre juillet 2008 et juin 2011, 2 335 patients atteints d'une maladie hépatique ont subi une MRF (86 % au moyen de la sonde moyenne). Dans 1,6 % des cas (n=37), la MRF a échoué (aucune mesure valide), ce qui était plus courant avec la sonde très grande qu'avec la sonde moyenne (3,4 % par rapport à 1,3 %; P=0,01). En plus des échecs de MRF, 4,9 % des patients (n=113) ont obtenu des MRF peu fiables. Les prédicteurs indépendants de MRF peu fiables incluaient un âge plus avancé (RC 1,03 [95 % IC 1,01 à 1,05]), une maladie pulmonaire chronique (RC 1,58 [95 % IC 1,05 à 2,37]), une coagulopathie (RC 2,22 [95 % IC 1,31 à 3,76]) et une plus grande rigidité du foie (RC par kPa de 1,03 [95 % IC 1,02 à 1,05]), y compris une cirrhose présumée (rigidité $\geq 12,5$ kPa; RC 5,24 [95 % IC 3,49 à 7,89]). Le sexe, le diabète, la maladie hépatique sous-jacente et la sonde du FibroScan n'étaient pas significatifs. Même si la fiabilité dépendait de l'opérateur (P<0,0005), son expérience n'était pas significative. Dans une sous-analyse de 434 patients dont on connaissait l'indice de masse corporelle, l'obésité influait sur le taux de résultats peu fiables (RC 2,93 [95 % IC 0,95 à 9,05]; P=0,06).

CONCLUSIONS : Il est rare que le FibroScan échoue et que la MRF ne soit pas fiable. Les principaux déterminants de résultats peu fiables sont un âge plus avancé, l'obésité, une plus grande rigidité du foie et l'opérateur, cette dernière caractéristique faisant ressortir la nécessité d'une formation convenable.

Moreover, potentially serious, albeit uncommon, complications including hemorrhage and death may occur (4). Due to these limitations, numerous, noninvasive means for staging liver fibrosis have been developed, including transient elastography (TE) (5-8) using FibroScan (Echosens, France). TE is an ultrasound-based tool for measuring liver stiffness as a surrogate of liver fibrosis. It is widely used due to its high

¹Liver Unit, Division of Gastroenterology and Hepatology, Department of Medicine; ²Department of Community Health Sciences, University of Calgary; ³Medical Services, Alberta Health Services, Calgary, Alberta

Correspondence: Dr Robert P Myers, Liver Unit, University of Calgary, 6D22, Teaching, Research and Wellness Building, 3280 Hospital Drive Northwest, Calgary, Alberta T2N 4Z6. Telephone 403-592-5049, fax 403-592-5090, e-mail rjmyers@ucalgary.ca

Received for publication October 28, 2013. Accepted November 11, 2013

accuracy for the diagnosis of advanced fibrosis, thus reducing the use of biopsy for the initial evaluation of patients with liver disease (9,10). In addition, studies have demonstrated that liver stiffness is responsive to changes in fibrosis over time (eg, attributable to treatment or disease progression), thus supporting its use for serial patient monitoring. An emerging body of literature also supports the prognostic significance of liver stiffness, suggesting that TE may be useful for guiding management strategies including the intensity of follow-up (11-17). For example, because the incidence of variceal hemorrhage is low in patients with liver stiffness <21 kPa, restriction of endoscopic screening to cirrhotic patients with higher liver stiffness has been advocated (14).

Because numerous clinical decisions are made based on the results of liver stiffness measurement (LSM) using TE, it is vital to understand the reliability of this tool and factors that influence its accuracy. Traditionally, TE examinations with <10 valid measurements, a success rate <60% and/or a ratio of the interquartile range (IQR) of liver stiffness to the median value (IQR/M) >30% have been classified as unreliable (18). According to this definition, approximately 15% of LSMs are unreliable, with higher rates among older patients, women and patients with diabetes, hypertension, and higher body mass index (BMI) and waist circumference. Limited operator experience (<500 versus ≥500 examinations) may also influence reliability (18). However, this definition has been criticized because no studies have ever demonstrated that reliable results are more accurate than unreliable results. As such, Boursier et al (19) recently proposed a new definition of poorly reliable TE results (IQR/M >30% and median liver stiffness ≥7.1 kPa). In this study, poorly reliable LSMs – which occurred in 9% of patients – were less accurate than reliable LSMs according to multiple indicators of diagnostic test performance (19). The factors that influence the rate of poorly reliable results according to this definition have not been explored.

Accordingly, the objective of the present study was to examine the feasibility and reliability of TE using this revised definition of reliability in a large cohort of patients with various liver diseases and severities to reflect routine clinical practice. We aimed to identify patient and operator characteristics associated with poorly reliable results to better inform our interpretation of LSMs using TE.

METHODS

Study population

In the present retrospective study, 2437 consecutive patients who underwent LSM using TE (FibroScan) at the University of Calgary Liver Unit (UCLU; Calgary, Alberta), between July 2008 and June 2011 were identified. The UCLU is the major referral centre for patients with liver disease who reside in Southern Alberta and serves a catchment population of approximately 1.5 million. LSM using TE is performed routinely in all patients who attend the UCLU without overt evidence of hepatic decompensation. In patients with multiple examinations, only the first LSM was considered to eliminate selection bias (ie, repeated examinations on patients who are easier to scan). For our analysis of the prevalence of FibroScan failure and reliability (see definitions below), the entire study cohort was examined. To identify predictors of poorly reliable LSM, only patients for whom linkage with Alberta administrative databases was possible (n=2335) were included to permit identification of the patients' underlying liver diseases and comorbidities. Therefore, for these analyses, 36 nonresidents of Alberta, 32 patients with invalid provincial health numbers and 34 patients for whom linkage with the administrative data were unsuccessful were excluded. These patients had demographic characteristics and liver stiffness results similar to the remainder (data not shown). The Conjoint Health Research Ethics Board at the University of Calgary approved the study protocol.

LSM

Two experienced operators (OP1 and OP2) performed all FibroScan examinations as per the manufacturer's recommendations. OP2 was

employed after OP1 and performed her first LSM in September 2010. Between July 2008 and July 2009, the FibroScan M probe was used in all patients; thereafter, the FibroScan XL probe was used in obese patients (BMI ≥30 kg/m²). Briefly, with the patient lying in the dorsal decubitus position and the right arm in maximal abduction, the tip of the FibroScan transducer probe was placed on the skin between the ribs over the right lobe of the liver. Assisted by a sonographic image, a portion of the liver at least 6 cm thick and free of large vascular structures was identified, and an attempt was made to collect at least 10 valid LSMs. The median liver stiffness value (in kPa) was considered to be representative of the elastic modulus of the liver. As an indicator of LSM variability, the IQR/M was calculated.

Administrative data sources

The present study used three Alberta administrative databases to identify the underlying liver disease etiologies and comorbidities of study participants via linkage using their unique personal health number. These databases have been used to examine the epidemiology (20-22), outcomes (22,23) and coding accuracy (20,24-28) of a variety of medical conditions.

1. Physician Claims Database. This database includes claims submitted for payment by Alberta physicians for services provided to registrants of the Alberta Health Care Insurance Plan, a universal plan that covers >99% of Alberta residents (29). Each record in the database includes the service provided, the date and up to three diagnosis fields. The database was queried from April 1, 2001 to March 31, 2011.
2. Inpatient Discharge Abstract Database. This database contains diagnosis, procedure and mortality information on all discharges from hospitals within Alberta. These data are routinely transmitted to the Canadian Institute for Health Information for aggregation with nationwide hospitalization data (29). Chart validation studies have shown rates of agreement exceeding 95% for demographic data and 75% to 96% for most responsible diagnosis codes (30). The database was queried from January 1, 1991 to January 31, 2012.
3. National Ambulatory Care Reporting System (NACRS)/ Ambulatory Care Classification System (ACCS) Database. This database contains information on facility-based ambulatory care including clinic and emergency department visits, same-day surgery, day procedures and rehabilitation services (29). The database was queried from July 1, 1996 to December 31, 2011.

Outcomes and predictor variables

The primary outcome variable was poorly reliable LSM, as defined by Boursier et al (19). Specifically, poorly reliable LSMs had an IQR/M >30% and median liver stiffness ≥7.1 kPa. Very reliable LSMs had an IQR/M ≤10%, whereas reliable LSMs had an IQR/M >10% and ≤30% (regardless of liver stiffness) or an IQR/M >30% with median liver stiffness <7.1 kPa. When compared with liver biopsy as the reference standard for staging fibrosis, poorly reliable LSMs are less accurate (ie, have lower areas under the ROC curves [AUROCs] and rates of accurate patient classification for fibrosis) than very reliable and reliable scans (19). As a secondary outcome measure, LSM failure, defined as no valid measurements after at least 10 attempts, was examined. The primary predictor variables included age, sex, the underlying liver disease, median liver stiffness (examined as a continuous variable and categorized as presumed cirrhosis [liver stiffness ≥12.5 kPa]) (31), comorbidities, the FibroScan operator (OP1 versus OP2), operator experience (first 499 scans versus ≥500 scans) (18) FibroScan probe (M versus XL) and the year of the LSM (2008 to 2009 versus 2010 to 2011). The hepatic diagnosis was categorized according to a hierarchy as follows: hepatitis B virus (HBV), hepatitis C virus, autoimmune (including primary biliary cirrhosis, primary sclerosing cholangitis and autoimmune hepatitis), hereditary hemochromatosis, alcoholic liver disease, nonalcoholic fatty liver disease and other (see Appendix 1 for list of relevant *International Classification of Diseases, Ninth Revision*,

TABLE 1
Characteristics of the study cohort according to the reliability of liver stiffness measurement

Variable	Entire cohort* (n=2335)	FibroScan			OR (95% CI) for poorly reliable FibroScan
		Very reliable (n=659)	Reliable (n=1526)	Poorly reliable (n=113)	
Male sex	56 (1310)	58 (381)	55 (835)	65 (73)	1.45 (0.98–2.16)
Age, years, median (IQR)	50 (40–58)	49 (40–57)	50 (40–58)	53 (48–61)	1.04 (1.02–1.05)
Body mass index†, kg/m ² , median (IQR)	25.5 (23–30)	25 (22–28)	25.5 (23–30)	32.6 (25–41)	1.06 (1.00–1.11)
<30	75 (332)	86 (107)	73 (211)	48 (10)	Reference
≥30	25 (112)	14 (18)	27 (77)	53 (11)	3.68 (1.52–8.94)
Hepatic diagnosis					
Hepatitis B	27 (638)	31 (201)	27 (413)	18 (20)	0.55 (0.34–0.90)
Hepatitis C	36 (840)	36 (234)	35 (541)	44 (50)	1.44 (0.99–2.11)
Autoimmune	5.5 (129)	4.3 (28)	6.0 (92)	7.1 (8)	1.31 (0.62–2.75)
Hemochromatosis	3.0 (69)	3.2 (21)	2.8 (43)	2.7 (3)	0.90 (0.28–2.92)
Alcohol	3.3 (78)	3.6 (24)	2.8 (42)	6.2 (7)	2.12 (0.95–4.73)
Nonalcoholic fatty liver disease	6.8 (158)	7.9 (52)	6.0 (92)	8.9 (10)	1.38 (0.70–2.69)
Other/unknown	18 (423)	15 (99)	20 (303)	13 (15)	0.68 (0.39–1.18)
Comorbidities‡					
Congestive heart failure	4.2 (97)	3.2 (21)	4.0 (61)	8.9 (10)	2.49 (1.25–4.94)
Cardiac arrhythmia	6.5 (152)	5.6 (37)	6.1 (93)	12.4 (14)	2.24 (1.24–4.02)
Hypertension	34 (803)	33 (218)	33 (507)	50 (56)	1.98 (1.35–2.89)
Paralysis	0.8 (18)	0.3 (2)	0.8 (12)	2.7 (3)	4.23 (1.20–14.9)
Chronic lung disease	30 (699)	29 (193)	29 (445)	40 (45)	1.60 (1.09–2.36)
Diabetes mellitus	12 (274)	9.4 (62)	11 (172)	23 (26)	2.49 (1.58–3.94)
Peptic ulcer	3.0 (69)	3.3 (22)	2.4 (37)	6.2 (7)	2.38 (1.06–5.34)
Coagulopathy	7.7 (179)	8.0 (53)	6.6 (100)	19 (21)	3.03 (1.84–5.01)
Fluid/electrolyte disorders	15 (340)	14 (94)	14 (210)	23 (26)	1.85 (1.17–2.91)
Alcohol abuse	13 (306)	14 (90)	12 (181)	22 (25)	2.01 (1.26–3.18)
Liver stiffness, kPa, median (IQR)	6.3 (4.7–10.1)	6.0 (4.6–8.8)	6.1 (4.6–9.6)	13.3 (9.0–27.0)	1.04 (1.03–1.05)
Cirrhosis (≥12.5 kPa)	19 (430)	17 (109)	17 (257)	57 (64)	6.49 (4.40–9.57)
M probe (versus XL probe)	86 (2009)	88 (582)	86 (1313)	78 (88)	0.54 (0.34–0.85)
Operator 2 (versus operator 1)	53 (1235)	53 (361)	54 (823)	35 (39)	0.45 (0.30–0.66)
Operator experience					
1–499 scans	35 (816)	38 (250)	33 (497)	37 (42)	Reference
≥500 scans	65 (1519)	62 (409)	67 (1029)	62 (71)	0.88 (0.59–1.30)
Year of FibroScan					
2008–2009	24 (550)	23 (154)	23 (334)	38 (43)	Reference
2010–2011	76 (1785)	77 (505)	78 (1192)	62 (70)	0.47 (0.32–0.69)

Data presented as % (n) unless otherwise indicated. Bolded values indicate statistical significance (ie, $P < 0.05$). *Entire cohort includes 37 FibroScan (Echosens, France) failures; †Body mass index data from medical record review available in 444 patients; ‡Only comorbidities with statistically significant associations with poorly reliable liver stiffness measurement are included. IQR Interquartile range

Clinical Modification [ICD-9-CM] [32] and ICD-10 [33] codes). For example, according to this approach, a patient with any record of a diagnosis code for HBV with or without other hepatic diagnoses would be categorized as having HBV. Comorbid conditions occurring before the FibroScan examination were defined using the Elixhauser list of 30 comorbidities (34), a well-validated algorithm for predicting outcomes in patients with hepatic (35) and nonhepatic disorders (34,36). Liver diseases were excluded from this algorithm. Also examined were the type of FibroScan probe because use of the XL probe could be considered a surrogate marker for obesity, which is not available in the administrative databases. In a subset of patients, medical records were reviewed to extract the BMI on the day of the FibroScan examination.

Statistical analyses

Between-group comparisons were made using Fisher's exact and χ^2 tests for categorical variables, and Wilcoxon rank-sum and Kruskal-Wallis tests for continuous variables. Univariate logistic regression was used to identify predictors of poorly reliable FibroScan examinations including age, sex, underlying liver disease, liver stiffness, comorbidities, the FibroScan probe, the operator, operator experience and year. Variables that were significant in univariate analyses were included in stepwise-forward, multiple logistic regression models in which

variables with $P < 0.1$ were retained in the models. In the subset of patients with available BMI data from medical record review, an additional multivariate analysis was performed including BMI. All analyses were performed using Stata version 11.0 (StataCorp, USA); a two-sided $P < 0.05$ was considered to be statistically significant.

RESULTS

Patient characteristics

A total of 2335 patients underwent LSM using TE at the UCLU between July 2008 and June 2011 and met the study inclusion criteria; their characteristics are outlined in Table 1. The median age was 50 years (IQR 40 to 58 years) and 56% were male. The majority of patients had chronic hepatitis C virus (36%) or HBV (27%) infection, while 7% had nonalcoholic fatty liver disease and 6% had autoimmune liver disease. Among 444 patients with BMI data available from medical record review, the median BMI was 25.5 kg/m²; 25% of patients were obese (BMI ≥30 kg/m²). Twelve percent of the cohort had a history of diabetes mellitus, 34% had hypertension, 35% had depression, and 13% and 10% had a history of alcohol and drug abuse, respectively.

LSM results

The majority of LSMs (86%) were obtained using the FibroScan M probe, 47% were performed by OP1 and 76% were performed during

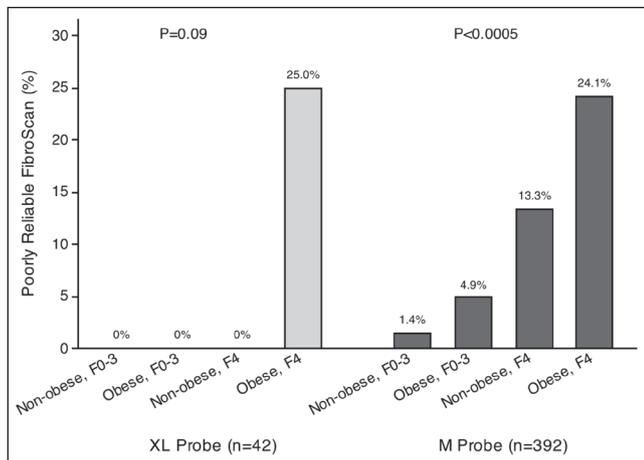


Figure 1 Prevalence of poorly reliable FibroScan (Echosens, France) examinations according to obesity (body mass index [BMI] ≥ 30 kg/m²), presumed cirrhosis (F4; liver stiffness ≥ 12.5 kPa) and FibroScan probe (M versus XL). This analysis is limited to 434 patients with available BMI data and successful liver stiffness measurement (ie, failures excluded)

the latter part of the study (2010 to 2011). In total, LSM failure was observed in 37 (1.6%) patients. The incidence of FibroScan failure was greater with the XL probe (3.4% [11 of 326]) than the M probe (1.3% [26 of 2009]; $P=0.01$) and for OP1 (2.3% [25 of 1100]) compared with OP2 (1.0% [12 of 1235]; $P=0.01$). Among the 2335 successfully measured patients, the median LSM was 6.3 kPa (IQR 4.7 kPa to 10.1 kPa). The majority (58%) of patients were estimated to have F0 to F1 fibrosis (liver stiffness <7.1 kPa), while 19% had presumed cirrhosis (liver stiffness ≥ 12.5 kPa) (31).

FibroScan reliability

According to the reliability criteria of Boursier et al (19) and excluding LSM failures, FibroScan examinations were classified as very reliable in 29% ($n=659$), reliable in 66% ($n=1526$) and poorly reliable in 4.9% ($n=113$) of patients (Table 1). According to previously recommended criteria for reliability (18), 15% of these examinations ($n=343$) would have been classified as unreliable (valid shots <10 , success rate $<60\%$ and/or IQR/M $>30\%$). Based on the updated definitions of Boursier et al (19), 41% of these 'unreliable' examinations would have been classified as very reliable, 22% as reliable and 37% as poorly reliable.

Predictors of poorly reliable FibroScan results

Table 1 includes patient and procedural characteristics according to the reliability of the FibroScan examination. In univariate analysis, older age (OR per year 1.04 [95% CI 1.02 to 1.05]) was associated with an increased risk of poorly reliable LSM; male sex was of borderline significance (OR 1.45 [95% CI 0.98 to 2.16]; $P=0.06$). Patients with HBV were less likely to have a poorly reliable LSM (OR 0.55 [95% CI 0.34 to 0.90]). Several comorbid conditions were associated with an increased risk of poorly reliable LSM including congestive heart failure, cardiac arrhythmias, hypertension, chronic pulmonary disease, diabetes mellitus, paralysis, peptic ulcer disease, coagulopathy, fluid and electrolyte disorders, and alcohol abuse (all $P<0.05$; Table 1).

Patients with poorly reliable LSM had higher median liver stiffness (13.3 kPa) than those with very reliable (6.0 kPa) and reliable scans (6.1 kPa; $P<0.0005$). The unadjusted odds of a poorly reliable scan increased by 4% per 1 kPa increase in liver stiffness (OR 1.04 [95% CI 1.03 to 1.05]). Similarly, presumed cirrhosis was more common in patients with poorly reliable LSM (57% versus 17% in patients with very reliable and reliable LSM; $P<0.0005$). Among 430 patients with cirrhosis, poorly reliable scans were recorded in 15% versus only 2.6% among noncirrhotic patients (OR 6.49 [95% CI 4.40 to 9.57]). In terms of procedural characteristics, poorly reliable LSM was less frequent with the M probe than the XL probe (4.1% versus 7.6%; OR

TABLE 2
Independent predictors of poorly reliable FibroScan examinations*

Variable	OR (95% CI)	P
Age [†] , per year	1.03 (1.01–1.05)	<0.0005
Male sex [†]	1.32 (0.87–1.99)	0.19
Obesity [‡] (body mass index ≥ 30 kg/m ²)	2.93 (0.95–9.05)	0.06
Chronic pulmonary disease	1.58 (1.05–2.37)	0.03
Coagulopathy	2.22 (1.31–3.76)	0.003
Liver stiffness [†] , per kPa	1.03 (1.02–1.05)	<0.0005
Cirrhosis [§]	5.24 (3.49–7.89)	<0.0005
M probe (versus XL probe)	0.64 (0.38–1.06)	0.08
Operator (OP1 versus OP2)	2.58 (1.70–3.93)	<0.0005
Operator experience [†] (<500 versus ≥ 500 scans)	0.86 (0.67–1.10)	0.23

*Data from a stepwise-forward logistic regression model (bolded values represent statistical significance) including age, sex, liver stiffness, FibroScan (Echosens, France) probe, operator (OP), OP experience, year of scan (2008 to 2009 versus 2010 to 2011), liver disease (hepatitis B versus other) and all comorbidities with $P<0.05$ in Table 1; [†]Variable forced into the model based on pre-existing literature suggesting an association with FibroScan reliability. If not shown in the Table, variables were dropped from the model due to $P>0.10$ (eg, liver disease, year and other comorbidities); [‡]OR for obesity from a separate model including the same variables as described above in the subset of patients with available body mass index data and excluding FibroScan failures ($n=434$); [§]OR for cirrhosis from a separate model including the same variables as described above

0.54 [95% CI 0.34 to 0.85]) and in scans performed by OP2 compared with OP1 (OR 0.45 [95% CI 0.30 to 0.66]). However, operator experience (<500 versus ≥ 500 scans) was not a significant predictor of poorly reliable LSM. Examinations performed in the latter half of the study period were less likely to be poorly reliable than earlier scans (2010 to 2011 versus 2008 to 2009 [4.0% versus 8.1%, respectively]; OR 0.47 [95% CI 0.32 to 0.69]).

Among 434 patients with available BMI data (excluding LSM failures), the median BMI was greater in patients with poorly reliable scans (30.0 kg/m² [IQR 25.5 to 32.0 kg/m²] versus 25.4 kg/m² [IQR 22.9 to 29.0 kg/m²]; $P=0.004$). Poorly reliable scans were observed in 10% (11 of 106) of obese patients (BMI ≥ 30 kg/m²) compared with only 3.1% (10 of 328) of nonobese individuals (OR 3.68 [95% CI 1.52 to 8.94]). Using the XL probe, 5.6% (two of 36) of obese patients had poorly reliable results versus 13% (nine of 70) measured using the M probe ($P=0.33$). The combined influence of obesity and presumed cirrhosis on the risk of poorly reliable LSM according to FibroScan probe is illustrated in Figure 1. In patients with obesity and cirrhosis, the risk of poorly reliable results was high (24% to 25%) with both probes compared with only 0% to 13.3% in patients with none or only one of these risk factors.

Table 2 includes the results of a multivariate model for predictors of poorly reliable FibroScan examinations with adjustment for age, sex, hepatic diagnosis (HBV versus other), liver stiffness, comorbidities, FibroScan probe and operator, operator experience and year. Independent predictors of poorly reliable LSM included older age (OR 1.03 per year [95% CI 1.01 to 1.05]), OP1 (OR 2.58 [95% CI 1.70 to 3.93]), higher liver stiffness (OR 1.03 [95% CI 1.02 to 1.05]), and a history of coagulopathy (OR 2.22 [95% CI 1.31 to 3.76]) and chronic pulmonary disease (OR 1.58 [95% CI 1.05 to 2.37]). While use of the M probe was of borderline significance (OR 0.64 [95% CI 0.38 to 1.06]; $P=0.08$), operator experience, year of LSM and patient sex were not statistically significant. In a supplementary analysis, in which liver stiffness was categorized rather than examined as a continuous variable, presumed cirrhosis was associated with a fivefold higher risk of poorly reliable results (OR 5.28 [95% CI 3.50 to 7.95]); the remainder of the results were largely unchanged (data not shown). Finally, in an analysis including BMI from medical record review, obesity was associated with nearly threefold odds of poorly reliable results (OR 2.93 [95% CI 0.95 to 9.05]; $P=0.06$) after adjustment for FibroScan probe and the other factors described above.

DISCUSSION

In the present large practice audit of >2300 FibroScan examinations, poorly reliable results – as defined by the revised definition of Boursier et al (19) – were uncommon, occurring in approximately 5% of individuals. In contrast, unreliable results according to a previously recommended definition (valid shots <10, success rate <60% and/or IQR/M >30% [18]) were observed in 15% of patients. This difference is relevant because the latter ‘unreliable’ results have never been shown to be less accurate than reliable results. In fact, use of this outdated definition of reliability may have led to the needless discarding of approximately 10% of FibroScan results, with potentially important implications in clinical practice and in research studies. Importantly, approximately two-thirds of these results would have been classified as very reliable or reliable according to the revised definitions. The prevalence of poorly reliable results observed in our study is slightly lower than that of Boursier et al (19) (4.9% versus 9.1%). This likely reflects differences in the study populations, particularly the lower prevalence of advanced fibrosis in our study (median liver stiffness 6.3 kPa versus 8.1 kPa in Boursier et al [19]) because higher liver stiffness (≥ 7.1 kPa) is a criterion in the definition for reliability (see below). Moreover, liver stiffness was measured using the XL probe in 25% of our patients, whereas only the M probe was used in the French study. Because liver stiffness measured with the XL probe is consistently lower than with the M probe, a lower rate of poorly reliable results could also be anticipated in our study.

Because TE is increasingly being used in clinical decision making, it is important to understand factors that influence its reliability because poorly reliable results are less accurate than reliable examinations. Specifically, in the study by Boursier et al (19), poorly reliable results were only 70% accurate for the diagnosis of cirrhosis compared with 86% and 90% in patients with reliable and very reliable results, respectively. Corresponding AUROCs for cirrhosis were 0.82, 0.90 and 0.97, respectively. With these facts in mind, we examined several patient- and operator-related characteristics as potential predictors of unreliable examinations. As previously reported, older age was associated with an increased risk of unreliable LSM (18). The exact reasons for this finding have never been identified; however, we speculate that age-related alterations in the chest wall are involved. It is known that chest wall compliance decreases with age due to structural changes of the intercostal muscles, intercostal joints and rib-vertebral articulations. In addition, age-associated osteoporosis may increase kyphosis, resulting in changes in the geometry of the thorax (37). On a related note, we identified a 1.6-fold increase in the risk of poorly reliable results among patients with chronic lung disease, predominantly chronic obstructive pulmonary disease (data not shown). This novel finding may also relate to structural changes in the chest wall (eg, pulmonary hyperinflation) or technical difficulties with the FibroScan procedure due to deep respirations in these patients. Because an increased risk of LSM failure or unreliable results has not been reported in cohorts with cystic fibrosis (38), additional studies are necessary to confirm this finding and to elucidate potential mechanisms.

Although previous studies have reported that women have a higher rate of unreliable FibroScan examinations compared with men (18), we did not observe a significant impact of sex using the updated definition of reliability. As previously reported (18), obesity was associated with a nearly threefold risk of unreliable results in a subanalysis of patients with available BMI data. The importance of obesity as a predisposing factor for poorly reliable results is supported by the borderline effect of XL probe use ($P=0.08$), considered a surrogate marker for obesity in the absence of BMI data in all patients. Presumably, subcutaneous and prehepatic adipose tissue in obese patients interferes with transmission of the mechanical shear wave and/or the measurement of its propagation by the FibroScan device (7).

Previous studies by Lucidarme et al (39) and Myers et al (40) have demonstrated an impact of fibrosis stage on the rate of discordance between fibrosis estimated by LSM and liver biopsy; however, an

impact on poorly reliable results has not been reported. In the current study, elevated liver stiffness was an independent predictor of poorly reliable LSM. Presumed cirrhosis was the most important risk factor, with fivefold higher odds of poorly reliable results in cirrhotic patients. This finding is not surprising because an LSM ≥ 7.1 kPa is one criterion in the definition of poor reliability (19). The other factor, IQR/M, also likely played a role because LSM variability tends to be greater in patients with cirrhosis (data not shown). The reason for this is unclear, but may relate to the broader range of potential LSMs in cirrhotic patients (ie, approximately 12.5 kPa to 75 kPa) compared with those who have lower liver stiffness (ie, 2.5 kPa to 12.4 kPa) (31). Interestingly, coagulopathy was associated with a twofold risk of poorly reliable results. Because there is no clear physical explanation for this finding, we suspect it reflects more severe liver disease and, therefore, higher liver stiffness in coagulopathic patients. The majority of these patients had diagnosis codes for thrombocytopenia as opposed to hereditary or acquired coagulation defects (data not shown). Because the effect of coagulopathy on poorly reliable results was independent of liver stiffness, it likely reflects the imperfect sensitivity of FibroScan for the diagnosis of cirrhosis.

A strength of our study was our analysis of the impact of specific liver conditions and comorbidities on FibroScan reliability using administrative data. In addition to uncovering a novel association between chronic lung disease and poorly reliable results, this approach revealed several other associations. First, patients with HBV had a lower likelihood of unreliable results. Because HBV-infected patients in our practice tend to be Asian and of smaller stature, this finding is not unexpected. In fact, we previously reported no FibroScan failures with the M probe in a study examining the value of the pediatric (S2) probe in this patient population (41). Second, hypertension and diabetes, components of the metabolic syndrome that have been associated with unreliable LSMs in previous studies (18), were significant in unadjusted, but not adjusted, analyses. Similarly, patients with congestive heart failure or arrhythmias had a twofold higher risk of poorly reliable LSM in univariate analyses. These findings may relate to hepatic congestion due to cardiac dysfunction, a well-described cause of liver stiffness overestimation (42). Because the power of our multivariate analysis may have been limited due to a small number of poorly reliable results ($n=113$), studies that use this novel definition of reliability in larger patient populations will be necessary to confirm or refute these findings.

In addition to patient-related predictors of FibroScan reliability, we examined the impact of procedural characteristics including the operator and operator experience. Castera et al (18) reported a lower risk of unreliable results among seven operators who had performed at least 500 examinations, challenging previous assertions that a novice can consistently obtain reliable results after a short training period of only 50 examinations. In our study, however, there was no difference in the proportion of poorly reliable results between the first 500 and subsequent examinations by our two operators. On the contrary, OP1 was twice as likely to produce a poorly reliable LSM as OP2. This finding supports the importance of adequate operator training and ongoing quality control when using the FibroScan in clinical decision making. It is important to note, however, that these results were confounded by the availability of the FibroScan XL probe only during the final two years of the study. Therefore, OP1 – who was employed before OP2 – scanned many obese patients using the M probe during the early part of the study. This likely led to an overestimation of her true rate of unreliable results. In fact, OP1 was twice as likely to have FibroScan failure as OP2, presumably for the same reason.

Our study has several limitations that warrant discussion. First, we did not have histological data to confirm whether poorly reliable LSMs were less accurate for staging fibrosis than reliable examinations. Second, because this was a retrospective study, we did not prospectively collect data regarding BMI and other anthropometric measures (eg, waist circumference, thoracic perimeter, skin-capsular distance)

that may have influenced FibroScan reliability. Similarly, we relied on administrative data to define patient comorbidities and hepatic diagnoses. Although many of these codes have been validated by medical record review (20,24-28), additional validation is necessary.

CONCLUSION

Poorly reliable FibroScan results are uncommon, occurring in approximately 5% of individuals. The most important determinants of poorly reliable LSM are older age, obesity, higher liver stiffness and the operator, the latter emphasizing the need for adequate training and quality control. A novel association between chronic pulmonary disease and poorly reliable results requires confirmation. Additional studies should be conducted to identify means of improving the reliability of FibroScan examinations so as to improve the accuracy of this valuable tool.

DISCLOSURES: Dr Myers was supported by a salary support award from the Canadian Institutes for Health Research (CIHR). Dr Kaplan is supported by salary support awards from the CIHR and Alberta Innovates-Health Solutions (AI-HS). Dr Swain is supported by the Cal Wenzel Family Foundation Chair in Hepatology. This study was supported, in part, by grants from AI-HS, CIHR (#84371) and the Canadian Liver Foundation. This study was based, in part, on data provided by Alberta Health. The interpretation and conclusions contained herein are those

REFERENCES

- Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD. Liver biopsy. *Hepatology* 2009;49:1017-44.
- Regev A, Berho M, Jeffers LJ, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. *Am J Gastroenterol* 2002;97:2614-8.
- Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology* 2005;128:1898-906.
- Myers RP, Fong A, Shaheen AA. Utilization rates, complications and costs of percutaneous liver biopsy: A population-based study including 4275 biopsies. *Liver Int* 2008;28:705-12.
- Sandrin L, Fourquet B, Hasquenoph JM, et al. Transient elastography: A new noninvasive method for assessment of hepatic fibrosis. *Ultrasound Med Biol* 2003;29:1705-13.
- Myers RP, Elkashab M, Ma M, Crotty P, Pomier-Layrargues G. Transient elastography for the noninvasive assessment of liver fibrosis: A multicentre Canadian study. *Can J Gastroenterol* 2010;24:661-70.
- Myers RP, Pomier-Layrargues G, Kirsch R, et al. Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients. *Hepatology* 2012;55:199-208.
- Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. *J Hepatol* 2008;48:835-47.
- Friedrich-Rust M, Ong MF, Martens S, et al. Performance of transient elastography for the staging of liver fibrosis: A meta-analysis. *Gastroenterology* 2008;134:960-74.
- Steadman R, Myers RP, Leggett L, et al. A health technology assessment of transient elastography in adult liver disease. *Can J Gastroenterol* 2013;27:149-58.
- Vergniol J, Foucher J, Terrebonne E, et al. Noninvasive tests for fibrosis and liver stiffness predict 5-year outcomes of patients with chronic hepatitis C. *Gastroenterology* 2011;140:1970-9.
- Klibansky DA, Mehta SH, Curry M, Nasser I, Challies T, Afdhal NH. Transient elastography for predicting clinical outcomes in patients with chronic liver disease. *J Viral Hepat* 2012;19:e184-e193.
- Merchante N, Rivero-Juarez A, Tellez F, et al. Liver stiffness predicts clinical outcome in human immunodeficiency virus/hepatitis C virus-coinfected patients with compensated liver cirrhosis. *Hepatology* 2012;56:228-38.
- Robic MA, Procopet B, Metivier S, et al. Liver stiffness accurately predicts portal hypertension related complications in patients with chronic liver disease: A prospective study. *J Hepatol* 2011;55:1017-24.
- Kim SU, Lee JH, Kim do Y, et al. Prediction of liver-related events using fibroscan in chronic hepatitis B patients showing advanced liver fibrosis. *PLoS One* 2012;7:e36676.
- Masuzaki R, Tateishi R, Yoshida H, et al. Prospective risk assessment for hepatocellular carcinoma development in patients with chronic hepatitis C by transient elastography. *Hepatology* 2009;49:1954-61.
- Jung KS, Kim SU, Ahn SH, et al. Risk assessment of hepatitis B virus-related hepatocellular carcinoma development using liver stiffness measurement (FibroScan). *Hepatology* 2011;53:885-94.
- Castera L, Foucher J, Bernard PH, et al. Pitfalls of liver stiffness measurement: A 5-year prospective study of 13,369 examinations. *Hepatology* 2010;51:828-35.
- Boursier J, Zarski JP, de Ledinghen V, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. *Hepatology* 2013;57:1182-91.
- Myers RP, Liu M, Shaheen AA. The burden of hepatitis C virus infection is growing: A Canadian population-based study of hospitalizations from 1994 to 2004. *Can J Gastroenterol* 2008;22:381-7.
- Kaplan GG, Gregson DB, Laupland KB. Population-based study of the epidemiology of and the risk factors for pyogenic liver abscess. *Clin Gastroenterol Hepatol* 2004;2:1032-8.
- Myers RP, Shaheen AA, Fong A, et al. Epidemiology and natural history of primary biliary cirrhosis in a Canadian health region: A population-based study. *Hepatology* 2009;50:1884-92.
- Myers RP, Shaheen AA, Li B, Dean S, Quan H. Impact of liver disease, alcohol abuse, and unintentional ingestions on the outcomes of acetaminophen overdose. *Clin Gastroenterol Hepatol* 2008;6:918-25.
- Myers RP, Leung Y, Shaheen AA, Li B. Validation of ICD-9-CM/ICD-10 coding algorithms for the identification of patients with acetaminophen overdose and hepatotoxicity using administrative data. *BMC Health Serv Res* 2007;7:159.
- Quan H, Parsons GA, Ghali WA. Assessing accuracy of diagnosis-type indicators for flagging complications in administrative data. *J Clin Epidemiol* 2004;57:366-72.
- Quan H, Parsons GA, Ghali WA. Validity of information on comorbidity derived from ICD-9-CCM administrative data. *Med Care* 2002;40:675-85.
- Quan H, Parsons GA, Ghali WA. Validity of procedure codes in International Classification of Diseases, 9th revision, clinical modification administrative data. *Med Care* 2004;42:801-9.
- Quan H, Sundarajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005;43:1130-9.
- Data Disclosure Handbook. In: Alberta Health and Wellness; 2003:1-15.

APPENDIX 1 Diagnosis codes used to define underlying liver disease etiologies

Hepatic diagnoses	Diagnosis code(s) (reference)	
	ICD-9-CM (32)	ICD-10 (33)
Hepatitis B	070.2, 070.3	B16, B18.0, B18.1, B19.1
Hepatitis C	070.41, 070.44, 070.51, 070.54, 070.7	B17.1, B18.2, B19.2
Nonalcoholic fatty liver disease	571.8	K75.81, K76.0
Alcoholic liver disease	571.0, 571.1, 571.2, 571.3	K70
Primary biliary cirrhosis	571.6	K74.3
Primary sclerosing cholangitis	576.1	K83.0
Autoimmune hepatitis	571.42	K75.4
Hemochromatosis	275.0	E83.1

ICD International Classification of Diseases

of the researchers and do not necessarily represent the views of the Government of Alberta. Neither the Government nor Alberta Health express any opinion in relation to this study. The authors have no financial disclosures or conflicts of interest to declare.

30. Williams JI, Young W. A summary of studies on the quality of health care administrative databases in Canada. In: Goel V, Williams JI, Anderson GM, Blackstein-Hirsch P, Fooks C, Naylor CD, eds. *Patterns of Health Care in Ontario: The ICES Practice Atlas*, 2nd edn. Ottawa: The Canadian Medical Association, 1996:339-345.
 31. Castera L, Vergniol J, Foucher J, et al. Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. *Gastroenterology* 2005;128:343-50.
 32. International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM). Los Angeles: Practice Management Information Corporation, 2001.
 33. WHO. *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10)* Geneva: World Health Organization, 1992.
 34. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care* 1998;36:8-27.
 35. Myers RP, Quan H, Hubbard JN, Shaheen AA, Kaplan GG. Predicting in-hospital mortality in patients with cirrhosis: Results differ across risk adjustment methods. *Hepatology* 2009;49:568-77.
 36. Li B, Evans D, Faris P, Dean S, Quan H. Risk adjustment performance of Charlson and Elixhauser comorbidities in ICD-9 and ICD-10 administrative databases. *BMC Health Serv Res* 2008;8:12.
 37. Sprung J, Gajic O, Warner DO. Review article: Age related alterations in respiratory function – anesthetic considerations. *Can J Anaesth* 2006;53:1244-57.
 38. Witters P, De Boeck K, Dupont L, et al. Non-invasive liver elastography (Fibroscan) for detection of cystic fibrosis-associated liver disease. *J Cyst Fibros* 2009;8:392-9.
 39. Lucidarme D, Foucher J, Le Bail B, et al. Factors of accuracy of transient elastography (Fibroscan) for the diagnosis of liver fibrosis in chronic hepatitis C. *Hepatology* 2009;49:1083-9.
 40. Myers RP, Crotty P, Pomier-Layrargues G, Ma M, Urbanski SJ, Elkashab M. Prevalence, risk factors and causes of discordance in fibrosis staging by transient elastography and liver biopsy. *Liver Int* 2010;30:1471-80.
 41. Pradhan F, Ladak F, Tracey J, Crotty P, Myers RP. Feasibility and reliability of the FibroScan S2 (pediatric) probe compared with the M probe for liver stiffness measurement in small adults with chronic liver disease. *Ann Hepatol* 2013;12:100-7.
 42. Millonig G, Friedrich S, Adolf S, et al. Liver stiffness is directly influenced by central venous pressure. *J Hepatol* 2010;52:206-10.
-



Hindawi
Submit your manuscripts at
<http://www.hindawi.com>

