UNIVERSITY OF BIRMINGHAM

Operator training requirements and diagnostic accuracy of Fibroscan in routine clinical practice

Armstrong, Matthew; Corbett, Christopher; Hodson, J.; Marwah, N.; Parker, R.; Houlihan, D. D.; Rowe, I. A.; Hazlehurst, J. M.; Brown, R.; Hubscher, S. G.; Mutimer, D.

Document Version
Publisher's PDF, also known as Version of record

Link to publication on Research at Birmingham portal

OPEN ACCESS

Operator training requirements and diagnostic accuracy of Fibroscan in routine clinical practice

M J Armstrong,[1,2] C Corbett,[1,2] J Hodson,[3] N Marwah,[4] R Parker,[1,2] D D Houlihan,[1,2] I A Rowe,[1,2] J M Hazlehurst,[5] R Brown,[6] S G Hübscher,[6,7] D Mutimer[1,2]

ABSTRACT

Background Fibroscan is a quick, non-invasive technique used to measure liver stiffness (kPa), which correlates with fibrosis. To achieve a valid liver stiffness evaluation (LSE) the operator must obtain all the following three criteria: (1) $\geq$10 successful liver stiffness measurements; (2) IQR/median ratio <0.30 and (3) $\geq$60% measurement success rate.

Objectives To assess the operator training requirements and the importance of adhering to the LSE validity criteria in routine clinical practice.

Methods We retrospectively analysed the LSE validity rates of 2311 Fibroscans performed (1 August 2008 to 31 July 2011) in our tertiary liver outpatients department at the University Hospital Birmingham, UK. The diagnostic accuracy of Fibroscan was assessed in 153 patients, by comparing LSE (valid and invalid) with the modified Ishak fibrosis stage on liver biopsy.

Results Learning curve analysis highlighted that the greatest improvement in validity of LSE rates occurs in the operator's first 10 Fibroscans, reaching 64.7% validity by the 50th Fibroscan. The correlation between LSE and the fibrosis stage on liver biopsy was superior in patients with a valid LSE (n=97) compared with those with an invalid LSE (n=56) (r_s 0.577 vs 0.259; p=0.022). Area under receiving operating characteristics for significant fibrosis was greater when LSE was valid (0.83 vs 0.66; p=0.048). Using an LSE cut-off of 8 kPa, the negative predictive value of valid LSE was superior to invalid LSE for the detection of significant (84% vs 71%) and advanced fibrosis (100% vs 93%).

Conclusions Fibroscan requires minimal operator training ($\geq$10 observed on patients), and when a valid LSE is obtained, it is an accurate tool for excluding advanced liver fibrosis. To ensure the diagnostic accuracy of Fibroscan it is essential that the recommended LSE validity criteria are adhered to in routine clinical practice.

INTRODUCTION

Chronic liver disease is now the third commonest cause of death in the UK in people under the age of 65 years. Early identification of people with significant liver fibrosis is therefore essential for ensuring the best outcomes from available treatments and preventing premature liver-related deaths. Due to the fact that most patients with chronic liver disease remain asymptomatic until their liver function is compromised, establishing the presence and severity of liver fibrosis remains a clinical challenge. Liver biopsy can accurately confirm the presence of liver fibrosis. However, its invasive nature, the risk of sampling error, interobserver variability and the understandable reluctance of patients to undergo repeat procedures make it an unsatisfactory approach.[1]

Consequently, over the last decade, non-invasive tools for identifying liver fibrosis have been developed,[2] with particular focus on the user-friendly technique of Fibroscan (Echosens, Paris, France).

Fibroscan, also called transient elastography, is a non-invasive technique used to provide a rapid measurement of liver stiffness (in kPa) at the bedside. A description of the Fibroscan technique and liver stiffness definitions are summarised in boxes 1 and 2.

To date, large meta-analyses of non-UK studies have shown that the liver stiffness evaluation (LSE) accurately correlates with histological fibrosis (in particular, cirrhosis) in several disease aetiologies.[3–5] The majority of these studies have focused on viral hepatitis (27 studies) and to a lesser extent non-alcoholic and alcoholic steatohepatitis (five studies). According to the manufacturer's criteria,[6] the LSE can be classified as 'valid', 'invalid' or an LSE 'failure' using the parameters that the Fibroscan machine provides (box 2). In research studies, the invalid LSEs are very often excluded from the statistical analyses.[7 8] Therefore, the clinical importance (diagnostic accuracy of Fibroscan) of complying with the LSE validity criteria has never been demonstrated in routine clinical practice. This question has significant implications in the UK and Ireland, as there are currently 134 Fibroscan machines in use in 70 hospitals (information provided by Echosens). Despite this widespread use, there has been a paucity of published Fibroscan data from routine clinical practice in the UK.

The Fibroscan was first introduced at our liver outpatient department at the Queen Elizabeth University Hospital Birmingham (UK) in 2008. Our liver and transplant unit is a tertiary referral centre for populations with varied racial and socioeconomic background in the Midlands and West of England, Wales and Northern Ireland. Since 2008, there has been a gradual increase in the number of LSEs performed per month (2008–2009, 54/month; 2010–2011, 78/month). Between 2008 and 2011, 2311 LSEs were performed as part of the clinical assessment of patients attending the liver outpatients department. Prior to using the Fibroscan machine in clinical practice, the manufacturer (Echosens) provides a small group (2–3 trainees) 4-h training session. The following are discussed: indications for LSE; relevant anatomical landmarks; instructions on how to use the probe; and how to interpret the LSE. Each trainee performed three supervised LSEs on healthy volunteers in the initial training, prior to use on patients with suspected liver disease. There are currently no local, national or international guidelines on how experienced an operator needs to be to achieve consistent and valid LSE readings on patients in the clinic setting.

Original article

What is a Fibroscan?
▶ Painless, quick (5–10 min), and non-invasive ultrasound technique, manufactured by Echosens (Paris, France)
▶ Non-invasive measurement of liver stiffness, which in turn positively correlates with the degree of fibrosis.[3–5]

How does a Fibroscan work?
▶ It uses a modified ultrasound probe to measure the velocity of an elastic shear wave created by a vibratory source
▶ The velocity of transmission of the shear wave through the liver is affected by the liver stiffness (ie, the stiffer the tissue, the faster the shear wave propagates).

How is this performed?
▶ The ultrasound probe is pressed against the skin (intercostal space) overlying the liver with the patient lying supine
▶ The probe generates a vibration and then measures the velocity of the resultant shear wave as it propagates through the liver.
▶ The aim is to obtain 10 successful measurements
▶ The machine then calculates the median value and the IQR of all the successful measurements (in kilopascals, kPa).

The objectives of the current study are to: (1) use statistical modelling to evaluate how many Fibroscans an operator needs to have performed on patients to achieve consistent and valid LSE readings and (2) assess whether obtaining a valid LSE (vs an invalid LSE) affects the diagnostic accuracy of the Fibroscan in routine clinical practice.

METHODS
Study population
All adult patients with suspected chronic liver disease who underwent a Fibroscan as part of their clinical assessment in the 3 years between 1 August 2008 and 31 July 2011 were included

What is the LSE?
▶ LSE=the median of the successful stiffness measurements (target ≥10)
▶ The LSE ranges from 2.5 (lowest stiffness) to 75 kPa (highest stiffness)

What is an LSE failure?
▶ When no stiffness measurements are obtained with 10 attempts

How is the success rate calculated?
▶ Success rate=number of successful measurements/total number attempted (expressed as %)

What are the manufacturer's LSE validity criteria?
▶ A 'valid' LSE is classified as obtaining all three of the following[6]:
  – ≥10 successful measurements
  – success rate ≥60%
  – IQR/median ratio <0.30
▶ An 'invalid' LSE is when one or more of the criteria are not fulfilled

in the study to assess the operator training requirements (1st objective). For comparison of the results of Fibroscan with liver biopsy we included those from this group who had a liver biopsy within 12 months of their Fibroscan examination (objective 2). The decision to perform a Fibroscan and to refer for a liver biopsy was made by the specialist hepatologist in clinic (consultant or specialist registrar). Patients with suspected chronic liver disease of any aetiology were included.

Liver Stiffness Evaluation
Between the study dates, either a consultant hepatologist or a specialist trainee registrar performed the Fibroscan during the outpatient clinic visit. In our unit, all operators underwent a certified training session with an Echosens consultant prior to use in the clinical setting.

All Fibroscans were performed using either the M-probe (3.5 Hz frequency) or XL-probe (2.5 Hz frequency) with the Fibroscan 502 machine (Echosens, France). The manufacturer recommends that the XL-probe should be used in patients with a skin-liver capsule distance >2.5 cm (measured by sonographic imaging). Due to the time constraints in liver clinic, operators were advised to use the XL-probe in patients with a measured Body Mass Index (BMI) >30 kg/m².[9] In May 2011, our unit began using the Fibroscan 502 Touch (Echosens, France), which has a built-in automated indicator that recommends the probe best suited to the patient's morphology. In accordance with manufacturer's guidance, all Fibroscans are performed in our clinics with the patient lying in the dorsal decubitus position with the right arm extended. The tip of the ultrasound probe (covered with gel) is placed on the skin in an intercostal space overlying the right lobe of the liver. A time-motion ultrasound image allows the operator to locate a portion of liver at least 6 cm thick and free of large vascular structures or ribs. The median and IQR value of successful liver stiffness measurements (target ≥10) is calculated by the machine and recorded as the LSE. Each LSE was classified as 'valid' or 'invalid' based on the manufacturer's validity criteria[6] (box 2).

Data collection
Data were retrospectively obtained from all three Fibroscan machines in our unit to form a database of the study cohort for assessment of operator training requirements. The Fibroscan parameters that were recorded included: patient identification number, date of Fibroscan, operator, probe, number of successful measurements, success rate and median value (IQR) of successful measurement (known as LSE).

Histopathology reports were then reviewed to identify those patients who had an ultrasound-guided liver biopsy within 12 months of the Fibroscan examination to assess the diagnostic accuracy of the Fibroscan. Demographics, anthropometric measurements (weight, height, BMI), liver enzymes and liver disease aetiology at the time of fibroscan examination were obtained for these cases. The definitive disease aetiology was determined by a combination of the clinical and histological findings and was categorised into fatty liver disease (non-alcoholic or alcoholic), viral hepatitis (hepatitis B, C), autoimmune (autoimmune hepatitis, primary biliary cirrhosis, primary sclerosing cholangitis), post-transplant and other, for purposes of statistical analysis.

Liver biopsy
Fibrosis staging was used to assess the accuracy of fibroscan for the diagnosis of significant and advanced fibrosis. In our centre, liver biopsies are routinely reported using the appropriate disease-specific liver fibrosis staging (ie, Ishak for hepatitis C;

Kleiner for non-alcoholic fatty liver disease). For purposes of this study, however, each biopsy was reassessed independently by two liver pathologists (NM and RB or NM and SGH) without knowledge of LSE results or other clinical data. In cases of disagreement, a consensus was reached by a joint review. To take account of the diverse aetiologies of liver disease, liver fibrosis was staged using a modified version of the Ishak scoring system,[10] as previously described by Rosenberg et al[11] (see online supplementary table S1). Significant fibrosis was defined as a modified Ishak score >2 and advanced fibrosis as a modified Ishak score of 5 or 6. The length of biopsy specimens and the number of portal tracts sampled were recorded as measures of biopsy quality. Biopsies specimens that were deemed not adequate by the pathologists for fibrosis staging were excluded from the analysis.

Statistical analysis

The demographics and characteristics of patients were summarised according to the validity criteria of the LSE (as defined above). Continuous variables were compared with independent sample t tests and Mann–Whitney tests (as applicable), and categorical variables were compared with Fisher's exact test.

Operator experience

Binary logistic regression was used to consider the effect of the number of Fibroscan examinations performed on the likelihood of a valid LSE reading. Prior to the analysis, the scan number was $\log_{10}$ transformed, in order that the model was based on the shape of curve generally observed in a learning curve analysis. The results of the analysis were only reported for the first 100 Fibroscan examinations, as some operators had performed fewer than 25 scans. This was in order to maximise their usefulness, while minimising the amount of extrapolation required. However, all the data (n=2311) was used in the production of the statistical model.

Diagnostic accuracy of Fibroscan

The strength of the relationship between the LSE and the modified Ishak score was analysed using Spearman's rho correlation coefficients. Separate coefficients were produced for those measurements where each of the three LSE validity criteria were met (ie, 'valid' LSE), and those where the criteria were contravened (ie, 'invalid' LSE). The coefficients were then compared to test whether non-compliance with the LSE validity criteria is detrimental to the ability of Fibroscan to predict the histological severity of liver fibrosis. The modified Ishak score was then converted into two binary outcomes indicating the presence of significant fibrosis (Ishak 3–6) and of advanced fibrosis/cirrhosis (Ishak 5 or 6). Receiver operating characteristic (ROC) curves were produced to test the accuracy of LSEs in the prediction of significant and advanced fibrosis. Separate ROC curves were produced for LSEs that were deemed 'valid' by each of the validity criteria, and those that were 'not valid', with comparisons made between the resulting areas under the ROC curves (AUROC). A LSE cut-off value of 8 kPa was used to determine the presence of significant fibrosis, above which further investigation is deemed appropriate.[12]

Analyses were performed using IBM SPSS V.19 and Microsoft Excel, with p values less than 0.05 deemed to be indicative of significance.

RESULTS

Effect of operator experience on obtaining a valid LSE (objective 1).

Patients

In the 3-year study period, 2311 LSEs were performed and included in the assessment of operator experience. Of these, 127 (5.5%) were LSE failures, 625 (27.0%) were invalid LSEs and 1559 (67.5%) were valid LSEs (figure 1).

Operator experience

Totally, nine consultants and eight specialist training registrars performed over 25 LSEs each, while a further 29 operators performed less than 25 LSEs each. The most experienced operator performed 670 LSEs, whereas the least experienced performed one in clinical practice (excluding the three performed on healthy volunteers during the initial training day).

Analysis

Binary logistic regression model (figure 2) was used to consider the effect of the number of LSEs performed on the likelihood of a valid LSE, as determined by obtaining the manufacturer's validity criteria. The model shows that a 10-fold increase in the number of LSEs that an operator has performed significantly improves their odds of obtaining a valid LSE (OR 1.57, 95% CI 1.39 to 1.78; p<0.001). Figure 2 shows that only 46% of the initial clinical LSEs performed by an operator were valid, whereas the validity rate rises to 57% by 10 LSEs. After 10 LSEs, the rate at which the operator achieves a valid LSE slows, reaching 64.7% by 50 LSEs and 67.7% by 100 LSEs. In order to obtain a valid LSE, 80% of the time the model forecasts that approximately 2500 LSE would be required.

Importance of the LSE validity criteria for the diagnostic accuracy of Fibroscan (Objective 2).

Patients

In the 3-year study period, 153 (6.6%) patients had a LSE (valid or invalid) that could be compared with liver biopsy (table 1). Of these, 56 (36.6%) patients had an invalid LSE; of which 21 patients (37.5%) had <10 successful measurements, 36 (64.3%) had IQR/median ratio >0.30, and 33 (58.9%) had a success rate <60% (figure 3). The mean age of this group (valid and invalid LSEs) was 48.4 (SE 1.1) years, 68.6% were male, and the mean BMI was 28.4 kg/m^2 (95% CI 27.3 to 29.5). The
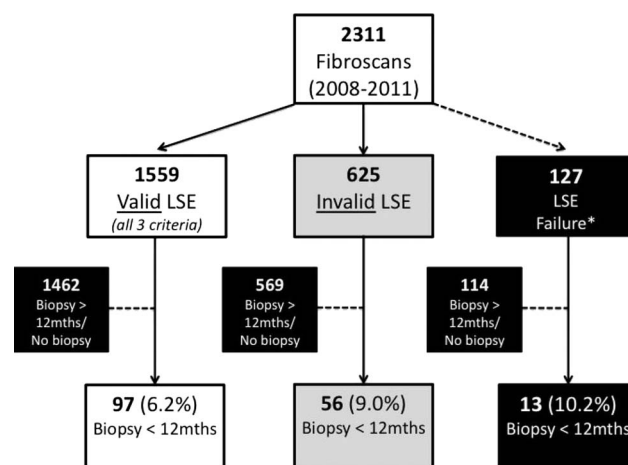

Figure 1 Flow diagram of the entire study; 2311 were included in the operator experience analysis. Of these, 153 patients were selected after exclusion (*black shading*) of patients in which the operator failed to get a single liver stiffness evaluation (LSE) reading (*defined as LSE failure) and/or when biopsy wasn't performed within 12 months of LSE.

Figure 2 Statistical model (learning curve) to highlight the number of liver stiffness evaluation (LSE) that need to be performed by an operator to achieve a consistent valid LSE. The black line represents the model produced from the binary logistic regression analysis. For scans 1–25, the rates of validity across all operators are plotted at each scan number (the red line). Since the number of operators drops off sharply (n=46 to 17 operators) after this point, the subsequent scans are summarised as a nine-point moving average, in order to isolate the trend from variability in the data. The model seems to be a reasonable fit to the observed data, suggesting that it is a valid summary of the general trend.
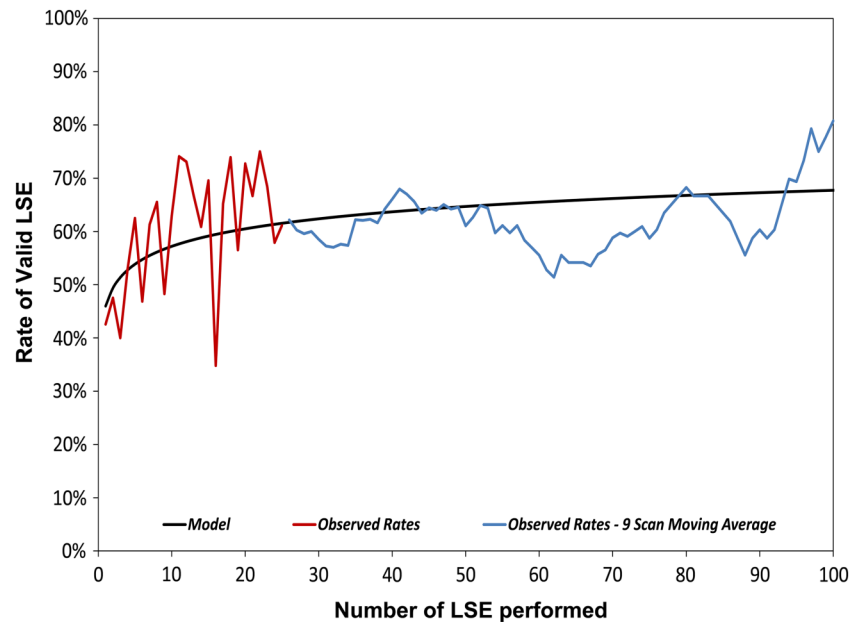
disease aetiologies (confirmed on biopsy) were fatty liver disease in 37.9% (n=58), viral hepatitis in 32.0% (n=49), autoimmune in 8.5% (n=13), post-transplant in 9.8% (n=15) and miscellaneous/other in 11.8% (n=18).

Liver histology and LSE
The median time difference between LSE and liver biopsy was 70 days (IQR 22.0–127.0). The mean number of portal tracts and length of biopsy was 14.9 (95% CI 13.9 to 16.1) and 15.7 (95% CI 14.9 to 16.5) mm, respectively. The liver pathologists deemed all 153 liver biopsies adequate for fibrosis staging. Seventy patients (45.7%) had significant fibrosis (Ishak stage 3–6), of which 25 had advanced fibrosis (Ishak 5–6). Seventy-eight (51.0%) of the LSEs were performed by consultant hepatologists, with the remainder by specialist registrars in training. One hundred and six (69.2%) of the LSEs were performed using the M-probe versus 47 (30.8%) with the XL-probe. Overall, the median LSE for the population of readings was 10.2 kPa (IQR 6.8–17.1).

Analysis
LSEs were significantly higher in patients with an invalid scan compared with those with a valid scan (14.1 vs 9.4 kPa; p=0.011). This was most pronounced in patients without fibrosis on biopsy (12.9 vs 5.6 kPa; p=0.008). There was no significant difference in age, sex, disease type, BMI, aspartate transaminase and histological parameters between patients with a valid LSE and those with an invalid LSE (table 1). The correlation between LSE and modified Ishak fibrosis stage was significantly superior in patients with a valid LSE compared to those with an invalid LSE (r_s 0.577 vs 0.259; p=0.022) (figure 4).

The accuracy of LSE (valid vs invalid) in predicting significant and advanced fibrosis was analysed using AUROC. The AUROC for significant fibrosis (Ishak 3–6) was significantly greater with a valid LSE than an invalid LSE (0.83 vs 0.66; p=0.048). There was no significant difference in the AUROC for advanced fibrosis (Ishak 5–6) between a valid LSE and an invalid LSE (0.87 vs 0.76; p=0.361).

The published[12 13] LSE cut-off of 8 kPa was used to determine the sensitivity, specificity, negative predictive value (NPV) and positive predictive value for the presence of significant (Ishak 3–6) and advanced fibrosis (Ishak 5–6) (table 2). A valid LSE produced a sensitivity of 86% (95% CI 71% to 95%) and specificity of 58% (95% CI 44% to 71%), whereas an invalid LSE resulted in a sensitivity of 84% (95% CI 64% to 95%) and a specificity of 42% (95% CI 22% to 63%). Subsequently, the NPV for the presence of significant fibrosis was 84% for a valid LSE compared with 71% for an invalid LSE. Furthermore, the NPV for presence of advanced fibrosis was 100% for a valid LSE versus 93% for an invalid LSE (table 2).

DISCUSSION
Our large retrospective, single-centre study (n=2311) highlights that Fibroscan requires minimal operator training (≥10 observed on patients). The greatest improvement in ability to achieve a valid LSE occurs in the operator's first 10 scans (46–57%), and thereafter the validity rate progressively increases, albeit very slowly. Second, our subgroup analysis of patients who underwent a liver biopsy (n=153) highlights that importance of adhering to the manufacturer's recommended LSE validity criteria. Obtaining a valid LSE (vs invalid LSE) resulted in a significantly greater correlation with liver fibrosis stage and greatly enhanced the accuracy of a negative LSE in ruling out significant (using LSE cut-off >8 kPa, NPV 84%), and advanced liver fibrosis (NPV 100%).

Clinical findings and implications compared to previous studies
It is important to understand what defines adequate Fibroscan training prior to widespread incorporation into UK clinical practice (including the potential for community-based assessment[13]). Previous hospital-based studies have reported contrasting degrees of operator experience that are required to achieve consistent and valid LSE readings (range 20 to >500 LSEs required [6 13–15]). Our statistical model highlights that the initial supervised period should incorporate a minimum of 10 Fibroscans on patients with suspected liver disease, to ensure that the trainee has the expected improvement in validity rate. Thereafter, 50 Fibroscans should achieve a stable degree of consistency in valid LSE rates. Furthermore, LSE validity rates were not affected by the grade of

Table 1 Demographics and characteristics of patients who underwent Fibroscan and liver biopsy

	Validity of LSE		
	Valid (n=97)	Not valid (n=56)	p Value
Characteristics			
Age (years)†	47.4 (1.4)	50.1 (1.5)	0.232
Gender			0.718
Male	68 (70.1%)	37 (66.1%)	
Female	29 (29.9%)	19 (33.9%)	
Disease type			0.678
Fatty liver disease (NAFLD/ALD)	39 (40.2%)	19 (33.9%)	
Viral (HBV/HCV)	28 (28.9%)	21 (37.5%)	
Autoimmune (AIH/PSC/PBC)	7 (7.2%)	6 (10.7%)	
Post-transplant	10 (10.3%)	5 (8.9%)	
Other	13 (13.4%)	5 (8.9%)	
BMI (kg/m^2)	27.9 (26.7 to 29.2)	29.2 (27.3 to 31.2)	0.263
AST (U/L)	49.9 (44.3 to 56.3)	48.4 (40.2 to 58.1)	0.762
Liver biopsy			
Time difference between biopsy and Fibroscan (days)‡	70.0 (23.5 to 122.5)	69.0 (12.8 to 195.0)	0.953
Portal tracts (n)	15.1 (13.8 to 16.6)	14.6 (12.7 to 16.7)	0.630
Length of biopsy (mm)	15.4 (14.5 to 16.5)	16.1 (14.7 to 17.6)	0.477
Modified Ishak Stage of Fibrosis (0–6)			0.387
0	14 (14.4%)	10 (17.9%)	
1	24 (24.7%)	13 (23.2%)	
2	17 (17.5%)	5 (8.9%)	
3	18 (18.6%)	9 (16.1%)	
4	8 (8.2%)	10 (17.9%)	
5	11 (11.3%)	4 (7.1%)	
6	5 (5.2%)		
Fibroscan			
Operator			0.738
Consultant	49 (50.5%)	30 (53.6%)	
Specialist registrar	48 (49.5%)	26 (46.4%)	
Probe			0.856
M-probe	68 (70.1%)	38 (67.9%)	
XL-probe	29 (29.9%)	18 (32.1%)	
LSE (kPa)‡	9.4 (6.6 to 14.5)	14.1 (7.3 to 26.1)	0.011§,*
LSE per modified Ishak Stage (kPa)‡			
0	5.6 (4.7 to 6.8)	12.9 (6.8 to 17.1)	0.008§,*
1–2	8.6 (6.5 to 10.9)	8.5 (6.1 to 18.8)	1.000§
3–4	11.4 (8.7 to 20.0)	16.0 (8.9 to 18.6)	1.000§
5–6	17.3 (12.1 to 26.0)	48.9 (11.9 to 68.2)	0.612§

Continuous data displayed as: Geometric Mean (95% CI) Categorical data displayed as: N (%).
*Significant=p<0.05.
†Data displayed as: Mean (SE).
‡Data displayed as: Median (Quartiles).
§p Values Bonferroni-Adjusted for 4 comparisons.
AIH, autoimmune hepatitis; ALD, alcoholic liver disease; AST, aspartate transaminase; BMI, body mass index; HBV, hepatitis B virus; HCV, hepatitis C virus; LSE, liver stiffness evaluation; NAFLD, nonalcoholic fatty liver disease; PBC, primary biliary cirrhosis; PSC, primary sclerosing cholangitis.

the doctor (consultant vs specialist registrar, p=0.738). This is in keeping with previous studies that recommend that a novice, of any medical professional status, can be trained to use Fibroscan.[13] [15]

By contrast with previous research studies,[7] [8] [16] our study highlights for the first time in UK clinical practice that complying with the recommended LSE validity criteria (box 2) provides better diagnostic accuracy than invalid LSE. Our data suggests that failure to meet the LSE validity criteria increases the risk of overinterpreting an LSE >8 kPA and incorrectly labelling a patient as having significant fibrosis, in those without fibrosis (figure 4). Furthermore, after obtaining an invalid LSE <8 kPA the clinician runs the risk of falsely reassuring 7% patients who have underlying advanced fibrosis (Ishak

5–6). By contrast, when a valid LSE is performed, clinicians (consultant or registrar level) can exclude significant fibrosis, and to a greater extent advanced fibrosis, with a high degree of confidence. In our unit, using the cut-off of 8 kPa with a valid LSE, we could reliably exclude significant and advanced fibrosis (sensitivities Ishak >2=86%; Ishak 5–6=100% and NPVs Ishak >2=84%; Ishak 5–6=100%) as effectively as data reported in large prospective studies,[14] [17] [18] nurse-based studies[13] and recent meta-analyses.[3–5] In order to reduce the number of false positive LSEs in our centre, while ultimately maintaining the ability to exclude advanced fibrosis (ie, NPV 100%), a cut-off of 10 kPa could be adopted (data not shown). Even by increasing the valid LSE cut-off further, our reported

Original article

< 10 successful measurements

Number successful measurements	N
1-3	10
4-6	9
7-9	2
Total	**21**
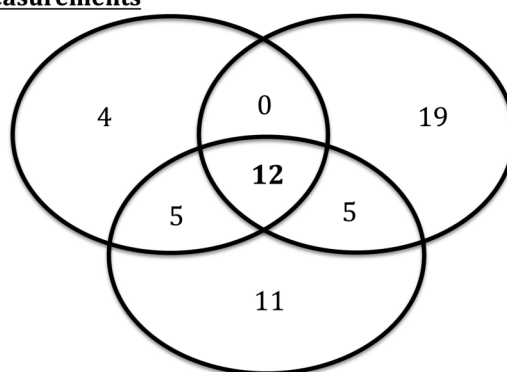
IQR/median ≥0.3

IQR/median ratio	N
0.3-0.49	31
0.5-0.69	4
0.7-0.90	1
Total	**36**

(Venn diagram values: 4, 0, 19, 5, **12**, 5, 11)

Success rate < 60%

Success rate %	N
1-19	3
20-39	9
40-59	21
Total	**33**

Figure 3 Venn diagram highlighting the reason(s) for an invalid liver stiffness evaluation (LSE) in the biopsied patients only. 56/153 patients had an invalid LSE. Tables highlight the distribution of spread for each of the three reasons that the LSE was classed as invalid.

number of false positives (LSE over estimates) outweigh those recently reported in a large 40 study meta-analysis by Tsochatzis *et al*.[3]

Limitations and strengths of the study:
Only a small percentage of the patients (153/2311; 6.6%) included in the study had a liver biopsy within 12 months, to

Figure 4 Relationship between the liver stiffness evaluation (LSE) and the modified Ishak stage of liver fibrosis for both valid LSE (*black circle*) and invalid LSE (*white box*). Median LSE values for each modified Ishak stage represented by horizontal bar.
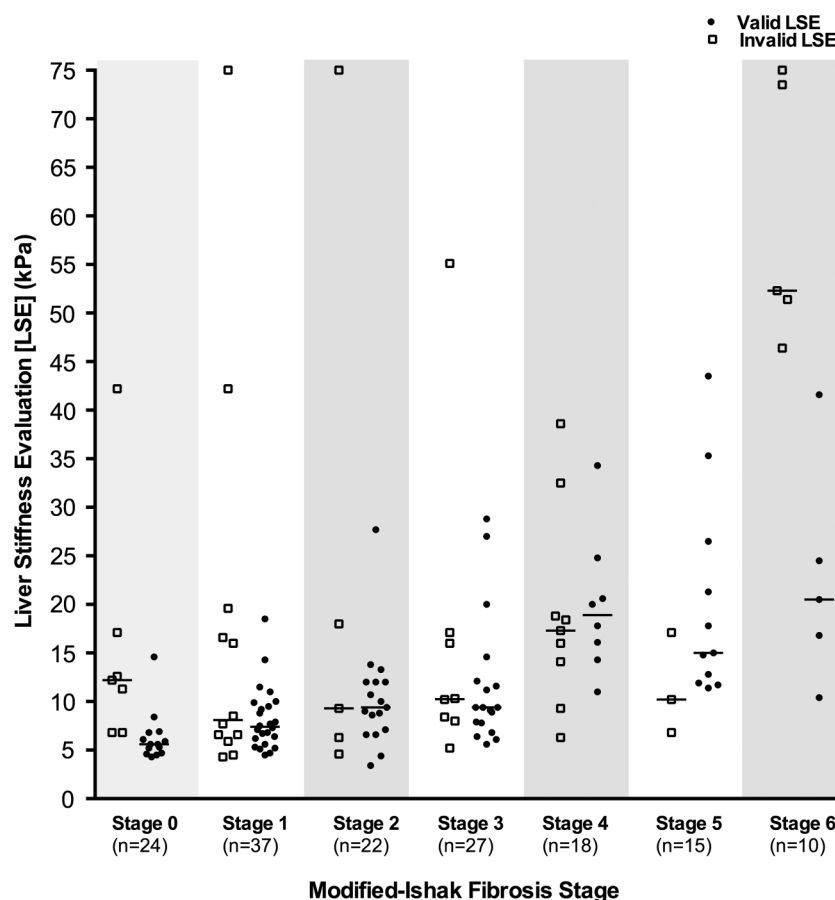
Table 2 Importance of the LSE validity criteria (cut-off > 8 kPa) for diagnostic accuracy of (a) significant and (b) advanced fibrosis

LSE validity criteria	Sensitivity %	Specificity %	PPV %	NPV %
a. Significant fibrosis (Ishak 3–6)				
Invalid LSE	84 (64 to 95)	42 (22 to 63)	60 (42 to 76)	71 (42 to 92)
Valid LSE	86 (71 to 95)	58 (44 to 71)	61 (47 to 73)	84 (69 to 94)
b. Advanced fibrosis (Ishak 5–6)				
Invalid LSE	88 (47 to 100)	32 (18 to 48)	20 (8 to 37)	93 (66 to 100)
Valid LSE	100 (79 to 100)	47 (36 to 58)	27 (16 to 40)	100 (91 to 100)

Sensitivity, specificity and positive/NPVs of LSE for 153 patients.
95% CIs in brackets. LSE validity criteria as per Castera et al.[6]
LSE, liver stiffness evaluation; NPV, negative predictive value; PPV, positive predictive value.

enable the diagnostic accuracy of Fibroscan to be determined. This introduces an unavoidable selection-bias that accompanies routine clinical decision-making, and provides a possible explanation for the high false positive rate of LSE in our study. For instance, in the event that the clinician has a low clinical suspicion of advanced liver fibrosis, he/she is more likely to proceed to a liver biopsy in the event of an unexpected high LSE compared with an expected (confirmatory) low LSE. A clinically relevant question would be to investigate how many patients in routine practice avoided liver biopsy as a direct result of the Fibroscan result. After the introduction of Fibroscan in our unit in August 2008, we saw a reduction in the number of outpatient liver biopsies (134 biopsies between 1 February 2008 and 1 July 2008; 89 biopsies between 1 February 2009 and 1 July 2009) (personal communication, Dr D Freshwater). This would imply that the Fibroscan influenced the clinicians' decision making, but to answer this question accurately would require prospective cohort study (ie, using questionnaires).

Due to time constraints it was routine practice in our centre to use the measured BMI (cut-off 30 kg/m² [9]) to determine the correct probe to use, rather than measure the skin to liver capsule distance (as used in previous research studies). This may have resulted in the inappropriate use of the M-probe in cases of >2.5 cm subcutaneous adipose tissue (despite a BMI <30 kg/m²) and therefore overestimates of LSE, as previously reported with the M-probe.[9] The sample size prior to and following the introduction of Fibroscan 502 Touch, which automatically informs the operator of which size probe to use, is too small to determine the impact of the new model of Fibroscan on the diagnostic accuracy in our centre. As in all studies that use liver biopsy to evaluate the performance of Fibroscan, interobserver agreement and sampling error in fibrosis staging must be considered.[19] In order to minimise these limitations, three liver pathologists (RB, SGH, NM) restaged liver fibrosis and reached a consensus in cases of disagreement (<10% cases). Given that the median time delay between Fibroscan and biopsy was 70 days (IQR 22–127) it is unlikely that progression of fibrosis could have contributed to discordance. Furthermore, time delay between Fibroscan and biopsy was not a predictor of false positives/negatives in our study (data not shown). Due to the sample size of our heterogeneous cohort, the employed LSE cut-offs for significant/advanced fibrosis were generic[5] [12] [13] and not specific to individual disease aetiology and/or probe use.

Outstanding research questions
Disease-specific and probe-specific LSE cut-offs for advanced fibrosis still require validation in UK clinical practice with

prospective study. The methodological challenges of comparing Fibroscan with histological fibrosis in clinical practice are well documented.[19] Future studies should, therefore, focus on investigating the influence of the LSE validity criteria (and modified versions[20]) and the accuracy of Fibroscan in predicting clinical events (ie, liver failure, hepatocellular carcinoma, death, etc).

SUMMARY
This study should inform other UK National Health Service centres that prior to using Fibroscan in clinical practice, novices should be trained to understand the clinical implications of the LSE validity criteria, and should undertake a minimum of 10 observed scans on patients prior to using the Fibroscan independently.

Main messages

▶ Fibroscan requires minimal operator training (≥10 observed scans on patients) prior to independent clinical use.
▶ The liver stiffness evaluation (LSE) validity criteria should be adhered to in clinical practice to ensure diagnostic accuracy.
▶ A valid LSE is an accurate, non-invasive tool for excluding advanced liver fibrosis

Current research questions

▶ What is the impact of Fibroscan on the clinical decision-making process? (ie, does it determine the requirement for liver biopsy, the choice of treatment and the decision to discharge from follow-up?)
▶ How accurate is Fibroscan (and liver stiffness evaluation validity criteria) in predicting clinical events (ie, liver failure, cancer, death)?
▶ What is the diagnostic accuracy and feasibility of Fibroscan in primary care?

Key references

▶ Tsochatzis EA, Gurusamy KS, Ntaoula S, et al. Elastography for the diagnosis of severity of fibrosis in chronic liver disease: a meta-analysis of diagnostic accuracy. *J Hepatol* 2011;54:650–9.
▶ Castéra L, Foucher J, Bernard P-H, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. *Hepatology* 2010;51:828–35.
▶ McCorry RB, Palaniyappan N, Chivinge A, et al. Development and evaluation of a nurse-led transient elastography service for the staging of hepatic fibrosis in patients with suspected chronic liver disease. *QJM* 2012;105:749–54.

Author affiliations
[1]Centre for Liver Research and NIHR Liver Biomedical Research Unit, University of Birmingham, Birmingham, UK

[2]Liver and Hepatobiliary Unit, Queen Elizabeth Hospital, Birmingham, UK
[3]Department of Statistics, Wolfson Computer Laboratory, University of Birmingham, Birmingham, UK
[4]India Department of Pathology, Post Graduate Institute of Medical Sciences, Rohtak, Haryana, India
[5]Centre for Endocrinology, Diabetes and Metabolism, University of Birmingham, Birmingham, UK
[6]Department of Cellular Pathology, Queen Elizabeth Hospital, Birmingham, UK
[7]School of Cancer Sciences, University of Birmingham, Birmingham, UK

REFERENCES
1 Bravo AA, Sheth SG, Chopra S. Liver biopsy. *N Engl J Med* 2001;344:495–500.
2 Rockey DC, Bissell DM. Noninvasive measures of liver fibrosis. *Hepatology* 2006;43:S113–20.
3 Tsochatzis EA, Gurusamy KS, Ntaoula S, et al. Elastography for the diagnosis of severity of fibrosis in chronic liver disease: a meta-analysis of diagnostic accuracy. *J Hepatol* 2011;54:650–9.
4 Friedrich-Rust M, Ong M-F, Martens S, et al. Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. *Gastroenterology* 2008;134:960–74.
5 Talwalkar JA, Kurtz DM, Schoenleber SJ, et al. Ultrasound-based transient elastography for the detection of hepatic fibrosis: systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2007;5:1214–20.
6 Castéra L, Foucher J, Bernard P-H, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. *Hepatology* 2010;51:828–35.
7 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.
8 Myers RP, Crotty P, Pomier-Layrargues G,, et al Prevalence, risk factors and causes of discordance in fibrosis staging by transient elastography and liver biopsy. *Liver Int* 2010;30:1471–80.
9 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.
10 Ishak K, Baptista A, Bianchi L, et al. Histological grading and staging of chronic hepatitis. *J Hepatol* 1995;22:696–9.
11 Rosenberg WM, Voelker M, Thiel R, et al. Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology* 2004;127:1704–13.
12 Roulot D, Czernichow S, Le Clésiau H, et al. Liver stiffness values in apparently healthy subjects: influence of gender and metabolic syndrome. *J Hepatol* 2008;48:606–13.
13 McCorry RB, Palaniyappan N, Chivinge A, et al. Development and evaluation of a nurse-led transient elastography service for the staging of hepatic fibrosis in patients with suspected chronic liver disease. *QJM* 2012;105:749–54.
14 Kettaneh A, Marcellin P, Douvin C, et al. Features associated with success rate and performance of FibroScan measurements for the diagnosis of cirrhosis in HCV patients: a prospective study of 935 patients. *J Hepatol* 2007;46:628–34.
15 Boursier J, Konate A, Guilluy M, et al. Learning curve and interobserver reproducibility evaluation of liver stiffness measurement by transient elastography. *Eur J Gastroenterol Hepatol* 2008;20:693–701.
16 Wong VW-S, Vergniol J, Wong GL-H, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. *Hepatology* 2010;51:454–62.
17 Castéra 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.
18 Foucher J, Chanteloup E, Vergniol J, et al. Diagnosis of cirrhosis by transient elastography (FibroScan): a prospective study. *Gut* 2006;55:403–8.
19 Boursier J, Calès P. Clinical interpretation of Fibroscan® results: a real challenge. *Liver Int* 2010;30:1400–2.
20 Boursier J, Zarski J-P, de Ledinghen V, et al. Determination of reliability criteria of liver stiffness evaluation by transient elastography. *Hepatology* 2013;57:1182–91.

Operator training requirements and diagnostic accuracy of Fibroscan in routine clinical practice

M J Armstrong, C Corbett, J Hodson, et al.

Updated information and services can be found at:

http://pmj.bmj.com/content/89/1058/685.full.html

These include:

Data Supplement	*"Supplementary Data"* http://pmj.bmj.com/content/suppl/2013/08/07/postgradmedj-2012-131640.DC1.html
References	This article cites 20 articles, 2 of which can be accessed free at: http://pmj.bmj.com/content/89/1058/685.full.html#ref-list-1
Open Access	This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/
Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections	Articles on similar topics can be found in the following collections Open access (15 articles) General surgery (141 articles) Radiology (345 articles) Surgical diagnostic tests (137 articles) Liver disease (67 articles)

To request permissions go to:

http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:

http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:

http://group.bmj.com/subscribe/

Notes