

Transient Elastography is superior to the FIB 4 Index, Aspartate Platelet Ratio Index, Aspartate Alanine Aminotransferase Ratio, Age Platelet Index and Fibrosis Index in Diagnosing Fibrosis in Chronic Hepatitis B Patients

Kwok R*, Lee AU and Ngu MC

Department of Gastroenterology and Hepatology, Concord Repatriation General Hospital NSW, Australia

Submission: April 14, 2016; Published: May 09, 2016

*Corresponding author: Raymond Kwok, Department of Gastroenterology and Hepatology, Concord Repatriation General Hospital NSW, Khoo Teck Puat Hospital, 90 Yishun Central, Singapore, Tel: 6591772658; Email: Raymond.dr.kwok@gmail.com

Abstract

Introduction: Transient Elastography is a promising form of non-invasive assessment of fibrosis in chronic hepatitis B. The diagnostic accuracy and usefulness of TE is evaluated and compared with FIB 4 index, Aspartate Platelet Ratio Index (APRI), Aspartate Alanine aminotransferase Ratio (AAR), Age Platelet Index (API) and Fibrosis Index (FI).

Methods: Chronic hepatitis B patients who had a liver biopsy within the past 6 months were identified and invited to have TE. Clinical history, laboratory data and pathology were collected TE was performed as per the manufacturer's instructions. At least 10 successful measurements were required for a valid Liver Stiffness Measurement (LSM). An interquartile range to median ratio of < 30% was required when the LSM \geq 7.1 kPa for a reliable LSM. A second histology assessment was performed on liver biopsies slides that were available.

Results: 71 patients were recruited. LSM Area Under Receiver Operator Characteristic (AUROC) curves for F \geq 1, 2, 3 and 4 were 0.825 (95% CI 0.728-0.922, p<0.001), 0.792 (95% CI 0.689-0.895, p< 0.001), 0.874 (95% CI 0.775-0.973, p<0.001) and 0.945 (95% CI 0.867-1.000, p=0.001) respectively. Using ALT level specific LSM Cut-offs, F \geq 2 and F \geq 3 can be diagnosed or excluded with a very high degree of certainty (>90%) in 49.3% and 57.7% respectively. TE was the most superior non-invasive measure for every stage of fibrosis when compared with FIB-4I, APRI, API, AAR and FI.

Conclusions: TE has excellent accuracy for F4 and F \geq 3 and can reduce the need for liver biopsies in the majority of chronic hepatitis B patients.

Keywords: Chronic hepatitis B; Liver fibrosis; Noninvasive diagnosis; Transient elastography; Fibroscan; FIB 4 index; Aspartate platelet ratio index; Aspartate alanine aminotransferase ratio; Age platelet index; Fibrosis index

Abbreviations: CHB: Chronic Hepatitis B; HBV: Hepatitis B Virus; TE: Transient Elastography; FIB 4: FIB 4 index; APRI: Aspartate Platelet Ratio Index; AAR: Aspartate Alanine aminotransferase Ratio; API: Age Platelet Index; FI: Fibrosis Index; LSM: Liver Stiffness Measurement; AUROC: Area Under the Receiver Operator Characteristic; WHO: World Health Organization; HCC: Hepatocellular Carcinoma; HbsAg: Hepatitis B surface Antigen; HbeAg: Hepatitis b e Antigen; kPa: kilopascals; INR: International Normalized Ratio; ALT: Alanine aminotransferases; AST: Aspartate aminotransferase; SD: Standard Deviation; ETOH: Alcohol; F stage: Fibrosis stage; CPA: Collagen Proportionate Area; IA: Imaging Analysis

Introduction

According to the latest estimates from the World Health Organization (WHO), 240 million people are chronically infected with hepatitis B virus (HBV) as defined by hepatitis B surface antigen positive for at least 6 months [1,2]. There is a varying prevalence geographically, with a 75% majority of

affected individuals residing in Asia and the Western Pacific [3]. The spectrum of disease and natural history of chronic Hepatitis B (CHB) Virus infection are diverse and variable. However, liver cirrhosis and hepatocellular carcinoma (HCC) may occur during the natural course of infection. The annual average incidence of cirrhosis is estimated to be 2.1%, and the annual average incidence of HCC in cirrhotics is estimated

to be 3-6% [4]. However, depending on risk factors, the cumulative lifetime risk of cirrhosis and HCC may range from 15.9 - 76.2% and 4.4 - 61.8% respectively [5]. CHB is the major cause of HCC, and is responsible for 60-80% of the world's total cases of liver cancer. The WHO estimates that HBV-related end stage liver disease and HCC are responsible for over 780 000 deaths per year [1,2] and comprise 5-10% of the cases of liver transplantation [6-8].

Limitations of current antiviral therapy

The goal of therapy is to improve quality of life and survival by preventing progression of the disease to cirrhosis, decompensated end-stage liver disease, HCC and death. Active HBV replication is the key driver of liver injury and disease progression. Antiviral therapy is able to provide viral suppression, but currently do not completely eradicate the virus. It is uncommon for antivirals achieve a complete virological response as defined by hepatitis B surface antigen (HbsAg) loss. There are 2 main treatment options available: pegylated interferon and nucleotide/nucleoside analogues (NAs). Pegylated interferon has a finite treatment duration, but has significant side effects and only achieves a durable response (defined as 6 months post therapy) for hepatitis b e antigen (HbeAg) seroconversion in 32%; undetectable viral load in only 14-19% (defined as < 60-80 IU/ml); and ALT normalization in 41-59% [9]. The current first line NAs, entecavir and tenofovir, are highly potent, have a high genetic barrier to drug resistance and generally have few side effects. Entecavir is able to achieve viral suppression with undetectable DNA in over 90% of patients with a 0.4% rate of resistance in 4 years [10]. Tenofovir treatment after 5yrs was not found to have any resistance and was able to suppress viral DNA to < 400 IU/ml in 99% of patients [11]. However, the disadvantage of NAs is that treatment is commonly given for many years and is usually indefinite. In patients whom HbeAg positive, NAs have not demonstrated a durable effect on viral suppression after cessation even if there is HbeAg seroconversion. Detectable levels of viraemia occur in more than 90% within 4 years [12,13]. For those who are HbeAg negative, viral relapse occurs in 91.4% within 1 year of treatment cessation [14]. Guidelines generally recommend treatment until loss of HbsAg which is uncommon. Thus, treatment becomes indefinite [4,10,15,16]. An additional disadvantage is the cost of long term NAs therapy. This is particularly challenging as countries which have the highest CHB infection prevalence generally are developing nations which have low to intermediate gross national income per capita [17].

The importance of assessing liver fibrosis

Given the limitations of currently available antiviral drugs, the approach to initiating treatment involves selecting patients who will benefit. Liver fibrosis is a well-recognized prognostic factor. All guidelines recommend evaluation of the degree of liver fibrosis as part of the workup in the management of CHB patients

[4,10,15,16,18]. Firstly, the presence of moderate (Metavir stage F2) or advanced fibrosis (Metavir stage F3) is an indication for antiviral treatment. The goal of initiating therapy at this level of disease is to prevent the progression to F4 cirrhosis and its related complications. Secondly, compensated cirrhosis patients with no overt clinical, biochemical or radiological signs can be identified. The timing of surveillance for cirrhotic related complications such as HCC and varices can then be optimized. Appropriate interventions may then be applied to prevent liver related morbidity and mortality.

The need for non-invasive markers

Liver biopsy is considered the reference standard for assessing liver fibrosis and is universally recommended by guidelines [4,10,15,16]. However there are several reasons which make liver biopsy not so ideal for assessing liver fibrosis. Although it is the reference standard, liver biopsy is subject to sampling error, especially when the length of biopsy is less than 2cm and the specimen contains less than 15 portal tracts [19]. The volume of liver sampled in a biopsy is only 1:50000 of the entire liver volume. It has also been shown that 1/3rd liver biopsies had a difference of at least 1 stage of fibrosis when samples from the left and right lobe were compared after taken laparoscopically [20]. There can also be significant inter-observer variability reported to be between 15%-33% [20-22]. There is a risk of severe complications. The rate of serious bleeding requiring blood transfusion is reported to occur in 1.7%. Puncture of viscous, inadvertent biopsy of other organs and arterial-venous fistula formation are rare [23]. The risk of death is estimated to be 0.1-0.01% of cases [24,25]. Despite these serious adverse effects, these are uncommon and liver biopsy is generally considered to be a safe procedure when guided by imaging and performed by experienced hands. Pain is common and occurs in 87% of cases, and persists in 20% beyond the day of the procedure [26]. Most protocols for post biopsy care require monitoring for several hours after the procedure to ensure there has been no serious side effects. This usually requires a one whole day commitment from the patient. Therefore, there is low patient acceptability and tolerance for liver biopsy. In addition, for liver biopsy to be performed safely, it requires skilled operators, imaging equipment and a facility for monitoring post procedure. The resource heavy requirement of liver biopsy makes it impractical to be used as routine tool on a large scale. Thus there is a need for non-invasive methods to be developed in order to cope with the estimated 240 million CHB patients worldwide.

Non-invasive markers of fibrosis in chronic hepatitis B

At the time of this research, very few studies existed that examined non-invasive markers for liver fibrosis in specifically in CHB. Most studies had focused on chronic hepatitis C. Transient elastography (TE) was an emerging tool for the noninvasive

assessment of liver fibrosis. Most of the TE data at the time was from Europe on Caucasian hepatitis C patients. There was scant data available for hepatitis B patients. A literature review at the time found only 2 studies which focused on the CHB population [27,28]. The goal of this study was to determine the diagnostic performance of Fibroscan and develop LSM cut-offs for each stage of fibrosis in CHB patients. Since there was little data for non-invasive markers of liver fibrosis in hepatitis B we performed head to head comparisons of TE with other non-invasive measures. Markers that were compared include the FIB 4 index (FIB4), Aspartate Platelet Ratio Index (APRI), Aspartate Alanine aminotransferase Ratio (AAR), Age Platelet Index (API) and Fibrosis Index (FI).

Methods

Patient selection and recruitment

From June 2008 to September 2009, chronic hepatitis B patients (as defined by presence of Hepatitis B surface antigen positive > 6 months) who had a valid liver biopsy (At least 6 portal tracts and 15mm in length) at Concord Repatriation Hospital, Concord Sydney Australia were identified. Patients who already had TE performed as part of their clinical management within the last 6 months of their liver biopsy were included retrospectively in the study. Those who did not have a TE were invited to do so, and included in the study if it was within 6 months of their liver biopsy. TE was performed during routine clinic visits at Concord Repatriation Hospital from February 2009 to September 2009. Patients who were recruited prospectively gave written informed consent and were older than the age of 18yrs. 73 patients were recruited.

Transient elastography assessment

TE was performed according to the training and instructions provided by the manufacturer. Scans were taken on the right lobe of the liver. The probe is placed in the intercostal space along the axillary line with the subject lying supine and the right arm at maximum abduction. A minimum of ten successful measurements was required, with the median score taken as the LSM. The success rate is the percentage of successful scans out of total number of attempts. The LSM is expressed in kilopascals (kPa). The LSM was considered reliable if the interquartile range/median ratio (IQR/M ratio) was less than 30% when the result was greater than 7.1 kpa [29]. Two officially trained operators (R.K and V.G) were responsible for carrying out the LSM. 2/73 cases were unable to achieve the minimum of 10 valid scans and so 71 cases were included in the final analysis.

Data collection

Clinical data for these patients were obtained from existing medical records that were available from public hospital and private specialists' rooms. Where available, data recorded include

age, gender, alcohol intake, any other documented chronic liver disease, imaging results, INR, liver function tests, hepatitis B surface antigen status, hepatitis B E antigen status, hepatitis B DNA viral load, details of antiviral therapy, platelet count, alpha fetoprotein and any other liver disease related morbidity. The value that was recorded for laboratory data would be the one that was the closest to the date of when the Fibroscan was performed and not exceeding 1 month.

Histological assessment

Where available, a histological assessment was performed on liver biopsy slides for specifically for the purposes of the study. Two experienced histopathologists (B.P.C.L and J.T), evaluated the specimens and determined the Metavir fibrosis stage. Any differences in assessment were deliberated. The final interpretation for stage of fibrosis was agreed upon unanimously after discussion. If the histological specimen was no longer available, then the original histology report that accompanied the liver biopsy was used and the Metavir Fibrosis score was taken as stated from the report.

Non-invasive markers

The formulas used to calculate the non-invasive measures analysed in the study are follows:

- $FIB-4 = [age \text{ (yrs)} \times AST \text{ (U/L)}] / [platelet \text{ count } (x10^9/L) \times \text{square root}(ALT(U/L))] \text{ [30]}$
- $APRI = 100 \times (AST \text{ (U/L)} / \text{upper level of normal}) / \text{platelet count } (x10^9/L). \text{ [31]}$
- $API \text{ score} = \text{The sum of age score and platelet count score}$
[age (years): <30 = 0, 30–39 = 1, 40–49 = 2, 50–59 = 3, 60–69 = 4, >70 = 5; platelet count (109/l): >225 = 0, 200–224 = 1, 175–199 = 2, 150–174 = 3, 125–149 = 4, <125 = 5] [32]
- $AAR = AST \text{ (U/L)} / ALT \text{ (U/L)}. \text{ [33]}$
- $FI \text{ score (fibrosis index)} = 8 - 0.01 \times \text{number of platelets } (10^9/L) - \text{albumin (g/dl)} \text{ [34]}$

Data analysis

All statistical analyses were performed using SPSS version 21.0 (IBM inc). Continuous variables were analysed using linear regression and independent samples T-test. Paired-related continuous variables were analysed using the paired T-test. Chi-squared test was used for categorical variables and Fisher's exact test when appropriate. Multivariate analysis was performed using multiple stepwise logistic regressions on variables found to be significant on univariate analysis. The overall accuracy of LSM in diagnosing histological bridging fibrosis and cirrhosis was calculated using the receiver operating characteristics (ROC) curve and its 95% CI. The accuracy of APRI, AAR and FIB-4, FI AND API were also calculated using the receiver operating

characteristics curve. A P-value of <0.05 was considered statistically significant. All statistical tests were two-sided.

Results

Liver stiffness measurement characteristics

71/73 subjects had a valid TE and were included in the analysis. The median LSM was 6.9 kPa (IQR 5.3-10.7 kPa). The mean IQR/M ratio was 0.20 (standard deviation (SD) 0.24). 66/71 (93.0%) scans were reliable, as defined as having an IQR/M ratio of greater of equal to 0.30, when the LSM ≥ 7.1kpa [29]. The findings are showing in Table 1.

Table 1: Liver Stiffness Measurement characteristics.

Liver Stiffness Measurement Characteristics	
Fibroscan parameter	Subjects
LSM (kpa)	6.9 (5.3-10.7) ^a
IQR/M ratio	0.20 (0.24) ^b
Valid Scans	71/73 (97.3%) ^c
Reliable scans	66/71 (93.0%) ^c
Success rate	90.4 (14.7) ^b

- a. Median LSM is reported with interquartile range
- b. Mean (standard deviation)
- c. Proportion (percentage)

Clinical characteristics of patients

The clinical characteristics of the study population are shown in Table 2. The mean age of the study population was 46.1 yrs (SD 11.9) and were predominantly male (69%). The mean log10HBVDNA viral load was 5.0 IU/ml (SD 2.5), mean ALT 121 U/L (SD 284) and mean AST 79 U/L (SD 161). These were all elevated. 33/71 (46.5%) were HBeAg positive. The mean bilirubin, albumin and platelet count were not elevated. The mean ETOH consumed per week was 4.8g (SD 10), which is much lower than the recommended level for safe alcohol consumption of 2 standard drinks (20g) per day. Data for INR, alphafetoprotein, imaging, height and weight were incomplete and was not included in the analysis.

A second histological assessment was able to be performed in 54/71 cases. The final Metavir fibrosis score that was attributed each liver biopsy and used for statistical analysis was based on the results of the second assessment except for 17/71 cases in which the original slides were no longer available. For these cases, the fibrosis stage that was reported in the original assessment was used. The frequency of each fibrosis stage were as follows: F0=14/71 (19.7%), F1=12/71 (16.9%), F2=26/71 (36.6%), F3=14/71 (19.7%) and F4=5/71 (7.0%). Details of the liver biopsy length and portal tracts were not routinely reported and could not be analysed. Treatment with antivirals occurred in 27/71 (38.0%) subjects at the time of the Fibroscan. Another 18/71 (25.4%) were planned to have treatment initiated, while 25/71 (35.2%) were to be monitored and managed without

antivirals.

Table 2: Clinical Characteristics of Patients.

Clinical Characteristics of Patients	
Characteristic	Subjects in Analysis
Male ^a	49/71 (69%)
Age (yrs) ^b	46.1 (11.9)
Log10HBVDNA (IU) ^b	5.0 (2.5)
HBe Antigen positive ^a	33/71 (46.5%)
Bili (umol/l) ^b	13 (8)
ALT (u/l) ^b	121 (287)
ALB (g/L) ^b	44 (5)
AST (u/l) ^b	79 (161)
Platelets (x10 ⁹) ^b	217 (50)
ETOH (g/week) ^b	4.8 (10)
Metavir Stage (after 2 nd assessment histology assessment)	
F0	14/71 (19.7%)
F1	12/71 (16.9%)
F2	26/71 (36.6%)
F3	14/71 (19.7%)
F4	5/71 (7.0%)
Treatment	27/71 (38.0%)
Entecavir ^a	20/27
Pegylated interferon ^a	3/27
Lamivudine and Adefovir ^a	2/27
Other (clev trial, adefovir only) ^a	2/27
Treatment planned, yet to be initiated ^a	18/71 (25.4%)
Treatment Previously ^a	1/71
No antivirals – monitoring only ^a	25/71 (35.2%)

- a. Subjects/total subjects (percentage)
- b. Mean (SD)

Factors associated with LSM

Statistical analysis was performed for the various clinical factors assessing for associations with the LSM. For scale variables, univariate analysis using linear regression revealed several statistically significant correlations with the LSM. These included the Metavir F score (p<0.001), bilirubin (p<0.001), ALT (p<0.001), high ALT (p=0.027), AST (p<0.001) and albumin (p=0.006). Gender, E antigen positivity, log HBV DNA, platelet count, alcohol intake and current treatment with antivirals were not associated. Age was also not statistically significant but showed a trend in univariate analysis (r = 0.165, p=0.085). After converting categorical variables into dummy scale variables, multivariate analysis using multiple linear regression showed that the only independently associated variables were Metavir F score (p<0.001) and AST (p<0.001). These findings are shown in the Table 3&4.

Table 3: Factors associated with LSM on Linear Regression Analysis.

Factors associated with LSM on Linear Regression Analysis		
	Correlation (r)	P value
Metavir F score	0.533	<0.001
Age (yrs)	0.165	0.085
logHBVDNA	0.026	0.414
Bili	0.387	<0.001
ALT	0.631	<0.001
ALB	-0.295	0.006
AST	0.742	<0.001
Platelets	-127	0.146
ETOH g/week	-0.030	0.403

Table 4: Factors associated with LSM on independent samples T-test.

Factors Associated with LSM on Independent Samples T-test			
	N	Mean LSM kpa (SD)	P value
Male	49/71 (69%)	M 9.3 (6.4) F 9.9 (9.6)	0.744
ALT high	37/71 (52.1%)	High ALT: 11.4 (9.7) Normal ALT: 7.4(3.0)	0.027
HBe Ag positive	33/71 (46.5%)	Pos 10.8 (9.5) Neg 8.4 (5.0)	0.181
Current treatment	27/71 (38.0%)	Yes 11.0 (10.4) No 8.5 (4.9)	0.181

Table 5: LSM AUROCs for diagnosing F≥1,2,3,4 and Optimal LSM cut-offs in all subjects.

LSM AUROCs for diagnosing F≥1,2,3,4 and Optimal LSM cut-offs with corresponding sensitivity and specificity in all subjects						
F≥	AUROC	95% ConfidenceInterval	P value	Best Overall LSM kpa (Sn,Sp%)	Best LSM kpa for > 90% sensitivity, (Sn,Sp%)	Best LSM kpa for > 90% specificity: (Sn,Sp%)
1	0.825	0.728-0.922	<0.001	6.5 (68.4, 92.9)	4.7 (93.0, 33.6)	6.5 (68.4, 92.9)
2	0.792	0.689-0.895	<0.001	7.5 (71.4, 77.8)	5.2 (91.4, 36.1)	9.7 (51.4, 94.4)
3	0.874	0.775-0.973	<0.001	9.7 (84.2, 92.3)	6.0 (94.7, 44.2%)	9.7 (84.2, 92.3)
4	0.945	0.867-1.000	0.001	11.9 (100, 84.8)	11.9 (100, 84.8)	15.9 (80.0, 97.0)

The area under the receiver operator characteristic (AUROC) curves for LSM diagnosing F0 vs. F≥1, F01 vs. F≥2, F012 vs. F≥3 and F0123 vs. F4 was calculated Figure 2. All patients in the study were included in this analysis. Hence, both normal and elevated ALT level subjects were included in this first analysis.

For F≥1, the AUROC was 0.825 (95% CI 0.728-0.922, p<0.001).The cut-off with best diagnostic accuracy overall was determined by choosing the value which corresponded to the greatest sum of the sensitivity and specificity. For diagnosing F≥1, this value was LSM ≥6.5 kpa corresponding to a sensitivity of 68.4% and specificity of 92.9%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was LSM≥4.7 kpa (sensitivity 93%, specificity 33.6%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was also the overall LSM≥ 6.5 kpa cut-off that had the best diagnostic accuracy.

Diagnostic performance of liver stiffness measurement for fibrosis stage for all subjects

A box plot for each liver biopsy according to fibrosis stage on the x-axis and the corresponding LSM on the y – axis is shown in Figure 1. Table 5 summarises the findings.

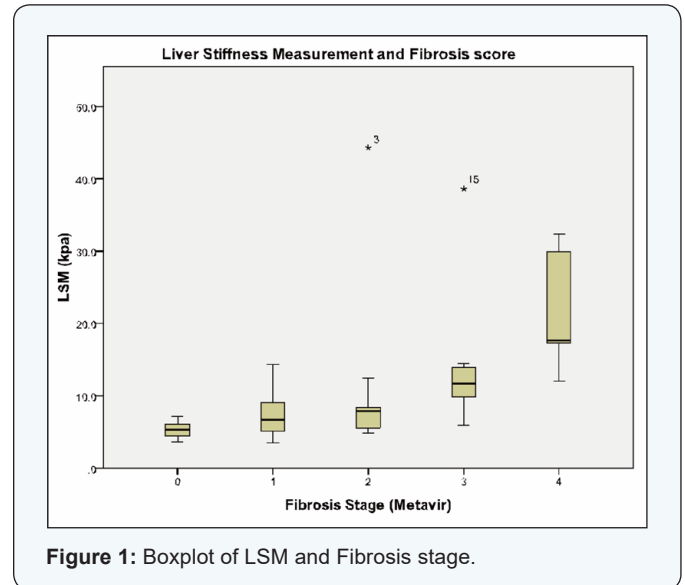


Figure 1: Boxplot of LSM and Fibrosis stage.

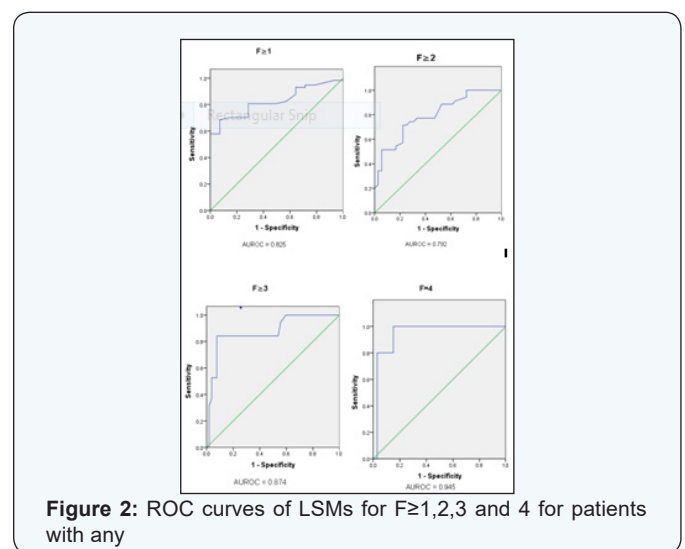


Figure 2: ROC curves of LSMs for F≥1,2,3 and 4 for patients with any

For $F \geq 2$, the AUROC was 0.792 (95% CI 0.689-0.895, $p < 0.001$). The cut-off with best diagnostic accuracy was $LSM \geq 7.5$ kpa corresponding to a sensitivity of 71.4% and specificity of 77.8%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was $LSM \geq 5.2$ kpa (sensitivity 91.4%, specificity 36.1%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was $LSM \geq 9.7$ kpa (sensitivity 51.4%, specificity 94.4%).

For $F \geq 3$, the AUROC was 0.874 (95% CI 0.775-0.973, $p < 0.001$). The cut-off with best diagnostic accuracy was $LSM \geq 9.7$ kpa corresponding to a sensitivity of 84.2% and specificity of 92.3%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was $LSM \geq 6.0$ kpa (sensitivity 94.7%, specificity 44.2%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was also the overall best diagnostic cut-off $LSM \geq 9.7$ kpa.

For $F = 4$, the AUROC was 0.945 (95% CI 0.867-1.000, $p = 0.001$). The cut-off with best diagnostic accuracy was $LSM \geq 11.9$ kpa corresponding to a sensitivity of 100% and specificity of 84.8%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was also $LSM \geq 11.9$ kpa (sensitivity 100%, specificity 84.8%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was $LSM \geq 15.9$ kpa (sensitivity 80.0%, specificity

97.0%).

Diagnostic performance of liver stiffness measurement for fibrosis stage for normal ALT patients

The area under the receiver operator characteristic (AUROC) curves for LSM diagnosing F_0 vs. $F \geq 1$, F_{01} vs. $F \geq 2$, F_{012} vs. $F \geq 3$ and F_{0123} vs. F_4 was calculated (Figure 3). Only normal ALT level subjects were included. Table 6 summarises the findings.

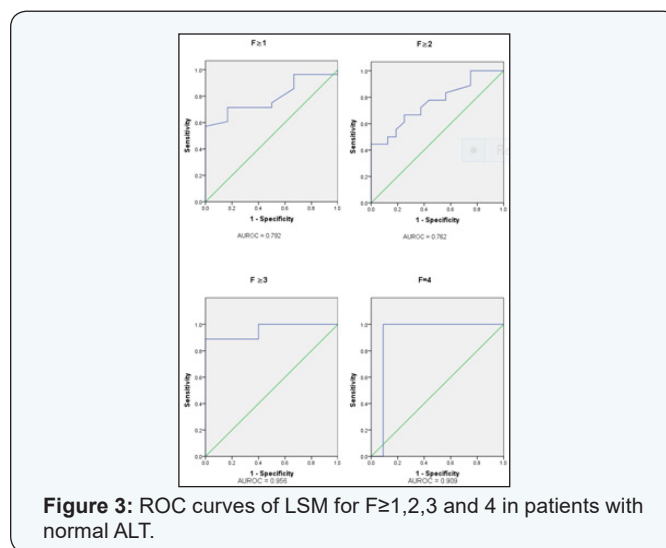


Figure 3: ROC curves of LSM for $F \geq 1, 2, 3$ and 4 in patients with normal ALT.

Table 6: LSM AUROCs for diagnosing $F \geq 1, 2, 3, 4$ and Optimal LSM cut-offs in Normal ALT subjects.

LSM AUROCs for diagnosing $F \geq 1, 2, 3, 4$ and Optimal LSM cut-offs with corresponding sensitivity and specificity in high ALT subjects						
$F \geq$	AUROC	95% Confidence Interval	P value	Best Overall LSM kpa (Sn,Sp%)	Best LSM kpa for > 90% sensitivity, (Sn,Sp%)	Best LSM kpa for > 90% specificity: (Sn,Sp%)
1	0.886	0.751-0.982	0.002	6.5 (79.3, 87.5)	5.8 (90.0, 62.5)	7.6 (69.0, 100)
2	0.847	0.723-0.971	<0.001	7.6 (88.2, 75)	5.8 (100, 40.0)	12.3 (52.9, 95.0)
3	0.817	0.639-0.994	0.003	10.5 (80.0, 85.2)	5.8 (100, 29.6)	12.5 (70.0, 92.6)
4	0.939	0.858-1.000	0.005	15.9 (100, 93.9)	15.9 (100, 93.9)	15.9 (100, 93.9)

For $F \geq 1$, the AUROC was 0.792 (95% CI 0.632-0.952, $p = 0.027$). Examining the coordinates of the curve, the cut-off with best diagnostic accuracy overall was determined by choosing the value which corresponded to the greatest sum of the sensitivity and specificity. For diagnosing $F \geq 1$, this value was $LSM \geq 6.6$ kpa corresponding to a sensitivity of 57.1% and specificity of 100%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was $LSM \geq 4.5$ kpa (sensitivity 96.4%, specificity 33.3%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was also the overall $LSM \geq 6.6$ kpa cut-off that had the best diagnostic accuracy.

For $F \geq 2$, the AUROC was 0.762 (95% CI 0.603-0.921, $p = 0.009$). The cut-off with best diagnostic accuracy was $LSM \geq 9.7$ kpa corresponding to a sensitivity of 44.4% and specificity of 100%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was $LSM \geq 4.7$ kpa (sensitivity 100%, specificity 25%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was also the overall $LSM \geq 9.7$ kpa cut-off that had the best diagnostic accuracy.

For $F \geq 3$, the AUROC was 0.956 (95% CI 0.868-1.000, $p < 0.001$). The cut-off with best diagnostic accuracy was $LSM \geq 9.7$ kpa corresponding to a sensitivity of 88.9% and specificity

of 100%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was $LSM \geq 6.0$ kpa (sensitivity 100%, specificity 60%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was also the overall $LSM \geq 9.7$ kpa cut-off that had the best diagnostic accuracy.

For $F=4$, the AUROC was 0.909, but was not significant (95% CI 0.801-1.000 95%, $p=0.169$).

Diagnostic performance of liver stiffness measurement for fibrosis stage for high ALT patients

The area under the receiver operator characteristic (AUROC) curves for LSM diagnosing F_0 vs. $F \geq 1$, F_{01} vs. $F \geq 2$, F_{012} vs. $F \geq 3$ and F_{0123} vs. F_4 was calculated (Figure 4) (Table 7). Only elevated ALT level subjects were included.

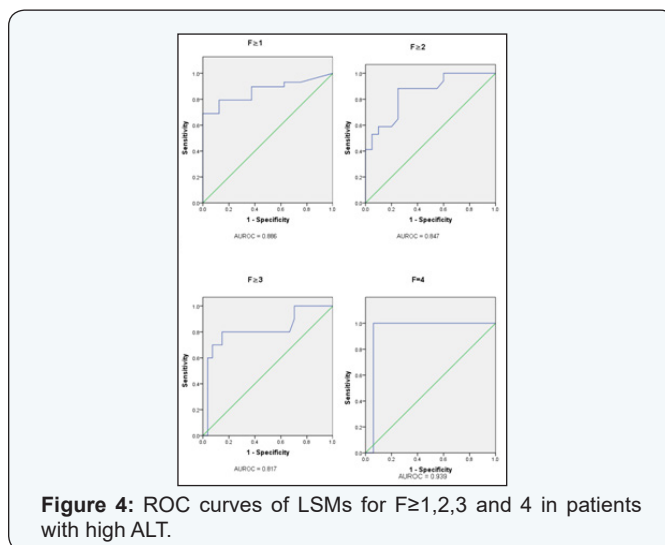


Figure 4: ROC curves of LSMs for $F \geq 1, 2, 3$ and 4 in patients with high ALT.

Table 7: LSM AUROCs for diagnosing $F \geq 1, 2, 3, 4$ and Optimal LSM cut-offs in high ALT subjects.

LSM AUROCs for diagnosing $F \geq 1, 2, 3, 4$ and Optimal LSM cut-offs with corresponding sensitivity and specificity in high ALT subjects						
$F \geq$	AUROC	95% Confidence Interval	P value	Best Overall LSM kpa (Sn,Sp%)	Best LSM kpa for > 90% sensitivity, (Sn,Sp%)	Best LSM kpa for > 90% specificity: (Sn,Sp%)
1	0.886	0.751-0.982	0.002	6.5 (79.3, 87.5)	5.8 (90.0, 62.5)	7.6 (69.0, 100)
2	0.847	0.723-0.971	<0.001	7.6 (88.2, 75)	5.8 (100, 40.0)	12.3 (52.9, 95.0)
3	0.817	0.639-0.994	0.003	10.5 (80.0, 85.2)	5.8 (100, 29.6)	12.5 (70.0, 92.6)
4	0.939	0.858-1.000	0.005	15.9 (100, 93.9)	15.9 (100, 93.9)	15.9 (100, 93.9)

For $F \geq 1$, the AUROC was 0.886 (95% CI 0.751-0.982, $p=0.002$). Examining the coordinates of the curve, the cut-off with best diagnostic accuracy overall was determined by choosing the value which corresponded to the greatest sum of the sensitivity and specificity. For diagnosing $F \geq 1$, this value was $LSM \geq 6.5$ kpa corresponding to a sensitivity of 79.3% and specificity of 87.5%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was $LSM \geq 5.8$ kpa (sensitivity 90.0%, specificity 62.5%). While the cut-off that corresponded to a specificity of at least 90% with best possible sensitivity was $LSM \geq 7.6$ kpa (Sensitivity 69.0%, specificity 100%).

For $F \geq 2$, the AUROC was 0.847 (95% CI 0.723-0.971, $p < 0.001$). The cut-off with best diagnostic accuracy was $LSM \geq 7.6$ kpa corresponding to a sensitivity of 88.2% and specificity of 75%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was $LSM \geq 5.8$ kpa (sensitivity 100%, specificity 40.0%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was $LSM \geq 12.3$ kpa (Sensitivity 52.9%, specificity 95.0%).

For $F \geq 3$, the AUROC was 0.817 (95% CI 0.639-0.994, $p=0.003$). The cut-off with best diagnostic accuracy was $LSM \geq 10.5$ kpa corresponding to a sensitivity of 80.0% and specificity of 85.2%. The cut-off that corresponded to a sensitivity of at least 90%

with the best possible specificity was $LSM \geq 6.0$ kpa (sensitivity 90.0%, specificity 29.6%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was $LSM \geq 12.5$ kpa (Sensitivity 70.0%, specificity 92.6%).

For $F=4$, the AUROC was 0.939 (95% CI 0.858-1.000, $p=0.005$). The cut-off with best diagnostic accuracy was $LSM \geq 15.9$ kpa corresponding to a sensitivity of 100.0% and specificity of 93.9%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was also the best overall $LSM \geq 15.9$ kpa (sensitivity 100.0%, specificity 93.9%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was yet again the best overall $LSM \geq 15.9$ kpa (sensitivity 100.0%, specificity 93.9%).

Optimal LSM cut offs for moderate and advanced fibrosis

Optimal cut offs were chosen to “rule in” and “rule out” $F \geq 2$ and $F \geq 3$ were derived from the analyses made from the previous section. The cut off corresponding to at least 90% sensitivity with the best possible specificity was used to “rule out” disease. While the cut off that corresponded to at least 90% specificity with the best possible sensitivity was used to “rule in” disease. The group of cut-offs selected was also specific to whether the ALT was normal or abnormal. This is summarized in Table 8.

Table 8: Optimal LSM cut-offs for F \geq 2 and F \geq 3 according to normal or elevated ALT.

Optimal LSM cut-offs for F \geq 2 and F \geq 3 according to normal or elevated ALT				
	F \geq 2		F \geq 3	
	Rule out	Rule in	Rule out	Rule in
Normal ALT	4.7 (100, 25.0)	9.7 (44.4, 100)	6.0 (100, 60)	9.7 (88.9, 100)
High ALT	5.8 (90.0, 76.0)	12.3 (52.9, 95.0)	5.8 (100, 29.6)	12.5 (70.0, 92.6)

Values listed are the LSM (Kpa) with corresponding sensitivity and specificity %

These cut-offs were corresponded to the LSMs of the patients in the study to determine the number of subjects which could have moderate or advanced fibrosis ruled out or ruled in with a high degree (greater than 90%) of accuracy. Thus:

For F \geq 2:

- In normal ALT subjects, it is ruled out for 8/34 and ruled in for 8/34 subjects
- In high ALT subjects, it is ruled out for 9/37 and ruled in for 11/37 subjects
- Overall 36/71 subjects (50.7%) were ruled out or ruled in for F \geq 2
- 35/71 subjects (49.3%) could not be determined with at least 90% accuracy and are in the ‘grey zone’
- For F \geq 3
- Normal ALT subjects, ruled out for 15/34 and ruled in for 8/34 subjects
- High ALT subjects, ruled out for 9/37 and ruled in for 9/37 subjects.
- Overall 41/71 (57.7%) of subjects were ruled out or

ruled in for F \geq 3

- 30/71 subjects (42.3%) could not be determined with at least 90% accuracy and are in the “grey zone”.

Clinical, biochemical, imaging features of cirrhotic subjects compared with LSM

5 patients in this study had histological proven cirrhosis. The results of their LSM ranged from 12.0 kpa to 32.4 kpa. All subjects’ LSM scores were compared to the cut-offs that were established previously shown in Tables 5-7. All 5/5 subjects had a derived F score using the LSM as F4. None had any overt clinical features of decompensated cirrhosis. None had any abnormalities of the bilirubin, albumin, INR, or platelet count. 2 subjects had recent ultrasound imaging available that showed no features of associated with cirrhosis, such as modularity, portal vein dilatation, hypersplenism and hepatofugal. The alpha-fetoprotein was mildly elevated in subjects #1 and #2 to a level of 15 IU/ml and 17 IU/ml, which is consistent with cirrhosis. However this mild elevation is not specific for cirrhosis and can also be consistent with liver regeneration in chronic hepatitis. It may also indicate hepatocellular carcinoma with or without cirrhosis. The results are shown in Table 9.

Table 9: Comparison of markers in histologically proven cirrhosis patients.

Comparison of clinical, biochemical, ultrasound imaging and LSM scores of 5 histologically proven cirrhosis patients											
Subject	Clinical features of cirrhosis	Br (u/mol)	Alb (g/L)	INR	ALT (u/l)	AST (u/l)	Afp (IU/ml)	Platelets (x10 ⁹)	US	LSM (kpa)	LSM derived F score
1	None	12	42	1.1	45	48	15	173	No cirrhotic features	12.0	4
2	None	12	43	1.1	75	54	17	196	n/a	32.4	4
3	None	16	40	1.0	85	54	6	164	n/a	17.6	4
4	None	4	44	1.0	177	110	3	250	n/a	29.9	4
5	None	13	40	1.0	232	175	7	178	No cirrhotic features	17.3	4

Diagnostic performance of FIB-4 index for Fibrosis Stage

The FIB-4 AUROC for Fibrosis was calculated for F \geq 1, 2 3 and F=4.

The summary of the AUROC results and cut-offs are shown in Table 10 and the corresponding AUROC curves are shown in Figure 5.

Table 10: FIB4-Index AUROCs for diagnosing F≥1,2,3,4 and Optimal LSM cut-offs.

FIB4-Index AUROCs for diagnosing F≥1,2,3,4 and Optimal LSM cut-offs with corresponding sensitivity and specificity						
F≥	AUROC	95% Confidence Interval	P value	Best Overall FIB 4 (Sn,Sp%)	Best FIB 4 for > 90% sensitivity, (Sn,Sp%)	Best FIB 4 for > 90% specificity: (Sn,Sp%)
1	0.677	0.536-0.817	0.042	1.4733 (43.9, 92.9)	0.7298 (90.0, 28.6)	1.4733 (43.9, 92.9)
2	0.711	0.660-0.883	<0.001	1.3427 (71.4, 77.8) & 1.3786 (68.6, 80.6)	0.7761 (91.4, 25)	1.6128 (48.6, 91.7)
3	0.635	0.475-0.795	0.084	n/a	n/a	n/a
4	0.912	0.843-0.981	0.02	1.8342 (100, 86.6)	1.8342 (100, 86.6)	2.174 (60, 91.9)

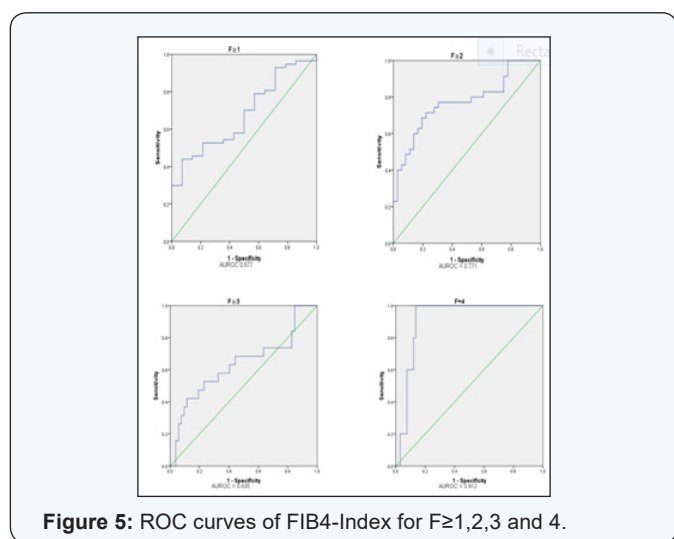


Figure 5: ROC curves of FIB4-Index for F≥1,2,3 and 4.

For F≥1, the AUROC was 0.677 (95% CI 0.536-0.817, p=0.042). The cut-off with best diagnostic accuracy overall for F≥1, was FIB-4≥1.4733 corresponding to a sensitivity of 43.9% and specificity of 92.9%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was FIB-4≥0.7298 (sensitivity 90.0%, specificity 28.8%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was also the score which have the best diagnostic accuracy: FIB-4≥ 1.4733 (sensitivity 43.9%, specificity 92.9%).

For F≥2, the AUROC was 0.711 (95% CI 0.660-0.883, p<0.001). The cut-off with best diagnostic accuracy overall for F≥2 there

was held equally by 2 cut-offs, and these values were: FIB-4≥1.3427 corresponding to a sensitivity of 71.4% and specificity of 77.8%; and FIB-4≥1.3786 corresponding to a sensitivity of 68.6% and specificity of 80.6%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was FIB-4≥0.7761 (sensitivity 91.4%, specificity 25.0%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was also the score which have the best diagnostic accuracy: FIB-4≥ 1.6128 (sensitivity 48.6%, specificity 91.7%).

The AUROC for diagnosing F3 using FIB-4 was not found to be significant (p=0.084).

For F=4, the AUROC was 0.912 (95% CI 0.843-0.981, p=0.002). The cut-off with best diagnostic accuracy for F=4, was FIB-4≥1.8342 corresponding to a sensitivity of 100% and specificity of 86.6%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was also the score with the overall best diagnostic accuracy: FIB-4≥1.8342 (sensitivity 100.0%, specificity 86.6%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was FIB-4≥ 2.174 (sensitivity 60.0%, specificity 91.9%).

Diagnostic performance of Aspartate Platelet ratio index (APRI) for Fibrosis Stage

The APRI AUROC for Fibrosis was calculated for F≥1, 2 3 and F=4.

The ROC curves for APRI for each stage of fibrosis is shown in Figure 6. The summary of the AUROC results and cut-offs are shown in Table 11.

Table 11: APRI AUROCs for diagnosing F≥1,2,3,4 and Optimal LSM cut-offs.

APRI AUROCs for diagnosing F≥1,2,3,4 and Optimal LSM cut-offs with corresponding sensitivity and specificity						
F≥	AUROC	95% Confidence Interval	P value	Best Overall APRI (Sn,Sp%)	Best APRI for > 90% sensitivity, (Sn,Sp%)	Best APRI for > 90% specificity: (Sn,Sp%)
1	0.672	0.534-0.809	0.048	0.5185 (49.1, 85.7)	0.2329 (90.0, 14.3)	0.559 (43.9, 92.9)
2	0.698	0.576-0.821	0.004	0.559 (57.1, 83.3)	0.286 (91.4, 33.3)	1.174 (17.1, 91.7)
3	0.720	0.596-0.844	0.005	0.559 (68.4, 75.0)	0.3493 (90.0, 42.3)	1.2929 (15.8, 90.4)
4	0.821	0.712-0.930	0.017	0.6106 (100, 74.2)	0.6106 (100, 74.2)	1.3429 (20.0, 90.9)

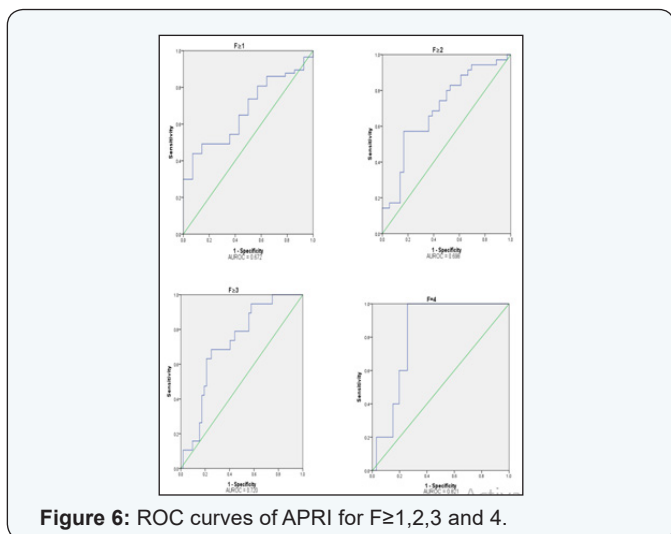


Figure 6: ROC curves of APRI for F≥1,2,3 and 4.

For F≥1, the AUROC was 0.672 (95% CI 0.534-0.809, p=0.048). The cut-off with best diagnostic accuracy overall for diagnosing F≥1 was APRI≥0.5185 corresponding to a sensitivity of 49.1% and specificity of 85.7%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was APRI≥0.2329 (sensitivity 90.0%, specificity 14.3%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was APRI≥ 0.559 (sensitivity 43.9%, specificity 92.9%).

For F≥2, the AUROC was 0.698 (95% CI 0.576-0.821, p=0.004). The cut-off with best diagnostic accuracy overall for diagnosing F≥2 was APRI≥0.559 corresponding to a sensitivity of 57.1% and specificity of 83.3%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was APRI≥0.286 (sensitivity 91.4%, specificity 33.3%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was APRI≥ 1.174 (sensitivity 17.1%, specificity 91.7%).

For F≥3, the AUROC was 0.720 (95% CI 0.596-0.844, p=0.005). The cut-off with best diagnostic accuracy overall for diagnosing F≥3 was APRI≥0.559 corresponding to a sensitivity of 68.4% and specificity of 75.0%. The cut-off that corresponded

to a sensitivity of at least 90% with the best possible specificity was APRI≥0.3493 (sensitivity 90.0%, specificity 42.3%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was APRI≥ 1.2929 (sensitivity 15.8%, specificity 90.4%).

For F=4, the AUROC was 0.821 (95% CI 0.712-0.930, p=0.017). The cut-off with best diagnostic accuracy overall for diagnosing F=4 was APRI≥0.6106 corresponding to a sensitivity of 100% and specificity of 74.2%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was also APRI≥0.6106 with the highest overall diagnostic accuracy (sensitivity 100%, specificity 74.2%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was APRI≥ 1.3429 (sensitivity 20.0%, specificity 90.9%).

Diagnostic performance of Age Platelet Index (API) for Fibrosis Stage

The API AUROC for Fibrosis was calculated for F≥1, 2, 3 and F=4.

The ROC curves for API for each stage of fibrosis is shown in Figure 7. The summary of the AUROC results and cut-offs are shown in Table 12.

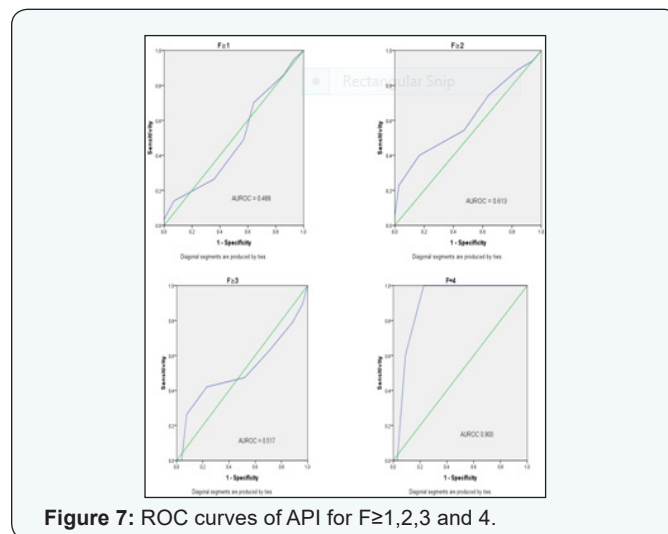


Figure 7: ROC curves of API for F≥1,2,3 and 4.

Table 12: API AUROCs for diagnosing F≥1,2,3,4 and Optimal LSM cut-offs.

API AUROCs for diagnosing F≥1,2,3,4 and Optimal LSM cut-offs with corresponding sensitivity and specificity						
F≥	AUROC	95% Confidence Interval	P value	Best Overall API (Sn,Sp%)	Best API for > 90% sensitivity, (Sn,Sp%)	Best API for > 90% specificity: (Sn,Sp%)
1	0.489	0.315-0.662	0.897	n/a	n/a	n/a
2	0.613	0.480-0.745	0.102	n/a	n/a	n/a
3	0.517	0.343-0.691	0.825	n/a	n/a	n/a
4	0.900	0.821-0.979	0.003	5 (100, 77.3)	5 (100, 77.3)	6 (60.0, 90.9)

API was only able to reliably diagnose F4, with AUROCs for F1, 2 and 3 not being statistically significant. The AUROC for F4 was 0.900 (95% CI 0.821-0.979, p=0.003). The cut-off with best diagnostic accuracy overall was determined by choosing the value which corresponded to the greatest sum of the sensitivity and specificity. For diagnosing F=4, this value was API≥5 corresponding to a sensitivity of 100.0% and specificity of 77.3%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was also the score with the overall best diagnostic accuracy: API≥5 (sensitivity 100.0%, specificity 77.3%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was API≥ 6 (sensitivity 60.0%, specificity 90.9%).

Diagnostic performance of Fibrosis Index (FI) for Fibrosis Stage

The FI AUROC for Fibrosis was calculated for F≥1,2,3 and F=4.

The ROC curves for FI for each stage of fibrosis is shown in Figure 8. The summary of the AUROC results and cut-offs are shown in Table 13.

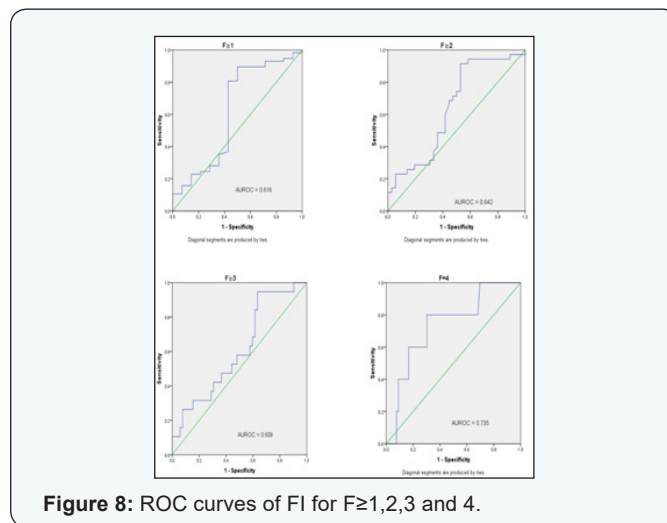


Figure 8: ROC curves of FI for F≥1,2,3 and 4.

Table 13: FI AUROCs for diagnosing F≥1,2,3,4 and Optimal LSM cut-offs.

FI AUROCs for diagnosing F≥1,2,3,4 and Optimal LSM cut-offs with corresponding sensitivity and specificity						
F≥	AUROC	95% Confidence Interval	P value	Best Overall FI (Sn,Sp%)	Best FI for > 90% sensitivity, (Sn,Sp%)	Best FI for > 90% specificity: (Sn,Sp%)
1	0.618	0.428-0.809	0.172	n/a	n/a	n/a
2	0.642	0.511-0.774	0.039	1.08 (91.4, 47.2)	1.08 (91.4, 47.2)	2.105 (22.9, 91.7)
3	0.608	0.464-0.753	0.165	n/a	n/a	n/a
4	0.735	0.523-0.947	0.082	n/a	n/a	n/a

Only F≥2 had a significant AUROC, which was 0.642 (95% CI 0.511-0.774, p=0.039). By examining the coordinates of the curve, the cut-off with best diagnostic accuracy overall was determined by choosing the value which corresponded to the greatest sum of the sensitivity and specificity. For diagnosing F≥2, this value was FI≥1.08 corresponding to a sensitivity of 91.4% and specificity of 47.2%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was also FI≥1.08 (sensitivity 91.4%, specificity 47.2%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was FI≥ 2.105 (sensitivity 22.9%, specificity 91.7%).

Diagnostic performance of Aspartate aminotransferase and alanine aminotransferase ratio (AAR) for Fibrosis Stage

The AAR AUROC for Fibrosis was calculated for F≥1, 2 3 and F=4.

The AUROCS for AAR diagnosing F≥1,3 and F=4 were not

significant. Table 14 summarises the findings and Figure 9 show the ROC curves.

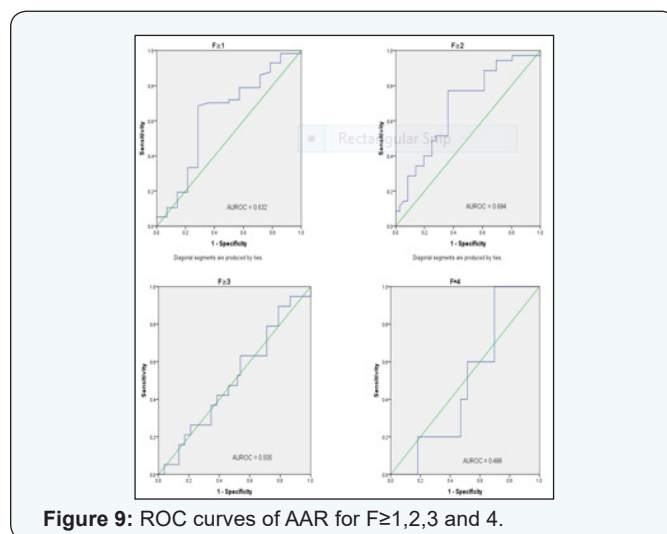


Figure 9: ROC curves of AAR for F≥1,2,3 and 4.

Table 14: AAR AUROCs for diagnosing F≥1,2,3,4 and Optimal LSM cut-offs.

AAR AUROCs for diagnosing F≥1,2,3,4 and Optimal LSM cut-offs with corresponding sensitivity and specificity						
F≥	AUROC	95% Confidence Interval	P value	Best Overall AAR (Sn,Sp%)	Best AAR for > 90% sensitivity, (Sn,Sp%)	Best AAR for > 90% specificity: (Sn,Sp%)
1	0.632	0.455-0.808	0.129	n/a	n/a	n/a
2	0.684	0.559-0.808	0.008	0.6954 (77.1, 63.9)	0.5353 (91.4, 30.6)	1.0571 (71.4, 91.7)
3	0.505	0.356-0.655	0.948	n/a	n/a	n/a
4	0.488	0.296-0.680	0.928	n/a	n/a	n/a

Only F≥2 had a significant AUROC, which was 0.684 (95% CI 0.559-0.808, p=0.008). By examining the coordinates of the curve, the cut-off with best diagnostic accuracy overall was determined by choosing the value which corresponded to the greatest sum of the sensitivity and specificity. For diagnosing F≥2, this value was AAR≥0.6954 corresponding to a sensitivity of 77.1% and specificity of 63.9%. The cut-off that corresponded to a sensitivity of at least 90% with the best possible specificity was AAR≥0.5353 (sensitivity 91.4%, specificity 30.6%). While the cut-off that corresponded to a specificity of at least 90% with best possible specificity was AAR≥ 1.0571 (sensitivity 71.4%, specificity 91.7%).

Comparison of the diagnostic performance of non-invasive measures for liver fibrosis

Table 15 summarises the AUROCS for each stage of liver fibrosis for the non-invasive measures that have been analysed. Overall, Fibroscan had the most superior AUROCS for diagnosing each stage of fibrosis in this study. LSM AUROCS for F≥1 and 2 were good and fair, while for F≥3 and F=4 were excellent. FIB-4 AUROCS were poor for F≥12, failed for F≥3, but excellent for F=4. APRI AUROCS were poor for F≥123, and good for F=4. API fails in diagnosing F≥123, but us an excellent test for F=4. While AAR and FI overall can be considered as failed diagnostic tests.

Table 15: Comparison of Non invasive tests for fibrosis stage.

Comparison of Non invasive tests for fibrosis stage						
F Stage	LSM AUROC	FIB-4 AUROC	APRI AUROC	API AUROC	FI AUROC	AAR AUROC
≥1	0.825	0.677	0.672	NS	NS	NS
≥2	0.792	0.711	0.698	NS	0.642	0.684
≥3	0.874	NS	0.720	NS	NS	NS
=4	0.945	0.912	0.821	0.900	NS	NS

NS: Not Significant.

Discussion

Transient elastography in chronic hepatitis B patients was performed reliably in > 90% of cases

In our study a valid LSM was obtained using TE in 97.3% of cases. A reliable LSM was obtained in 93.7% of cases. Overall this demonstrates that TE can be feasibly performed with a high degree of success and accuracy in chronic hepatitis B patients. The operators of the Fibroscan for this study were also the dedicated operators performing Fibroscan for the clinical service, and toward the end of study recruitment had performed more than a combined 900 scans. This is consistent with other studies which have found that operator experience is important in obtaining a high rate of valid and reliable LSM's that are reproducible [35].

Liver Stiffness Area Under the Receiver Operator Characteristic curves for F≥1 was good (0.825), fair for F≥2 (0.792), good for F≥3 (0.874) and excellent for F=4 (0.945)

Our study shows slightly inferior findings compared with Marceline's study. In our study, F≥2 was fair AUROC = 0.792 (95% CI 0.689-0.895, p< 0.001). F≥3 was good: AUROC = 0.874 (95% CI 0.775-0.973, p<0.001). F=4 was excellent: AUROC = 0.945 (95% CI 0.867-1.000, p=0.001). Marcelin et al. [27] found the diagnostic performance of TE to be good for F≥2 (AUROC = 0.81), and excellent for F≥3 and F4 (both AUROC = 0.93). Since the completion of this research, many other studies have assessed the diagnostic performance of TE in hepatitis B patients. A 2013 review by Chen et al. found that in 12 studies, the range of AUROCS reported for F≥2 was 0.78-0.87; F≥3 0.87-

0.92 and F4 0.80-0.96 [36]. A separate 2012 review performed a pooled meta-analysis of 18 studies comprising of 2772 patients. The pooled AUROCS for $F \geq 2$, $F \geq 3$ and $F=4$ were 0.859 (95% CI 0.857–0.860), 0.887 (95% CI 0.886–0.887), and 0.929 (95% CI 0.928–0.929) respectively [37]. Our study's results are within the range observed in these reviews and meta-analysis.

Our results suggest that Fibroscan has superior diagnostic performance for the latter stages of fibrosis (ie $F \geq 3$ and F4) compared to the earlier stages (ie $F \geq 1$ and $F \geq 2$). The AUROC ranges from other studies that were summarized in the Chen's review also provide strong support for this observation. One explanation appears to be due to LSM having a stronger correlation with per cellular fibrosis, which occurs more in latter stages of fibrosis, compared with per portal and perivenular fibrosis as demonstrated by Wong et al. [28]. Another explanation perhaps is that LSM is a better reflection of the volume of fibrosis, rather than the stage of fibrosis. This is discussed later in detail

Higher ALT does not affect diagnostic performance, but increases optimal cut-offs by factor of 1.3

Wong's study had reported a reduction in the diagnostic performance of TE in high ALT patients compared to normal ALT patients. AUROC for F4 decreased from 0.93 to 0.86 [28]. However, in our study, the diagnostic performance of LSM was not observed to be reduced in higher ALT patients compared to normal ALT patients. The AUROCs were similar. For $F \geq 1$ (0.886 vs. 0.792) and $F \geq 2$ (0.847 vs. 0.766) were both higher in the high ALT cohort when compared to the normal ALT cohort. For $F \geq 3$ the AUROC was worse (0.817 vs. 0.956) for high ALT, but in F4 it was superior (0.939 vs. 0.909). The comparison was limited as F4 AUROC in normal ALT did not reach statistical significance due to low numbers. The overall comparison must also be interpreted with care due to the small size of the 2 groups.

Although the AUROCS were found to be fairly comparable between high and normal ALT subjects, this certainly does not mean that the exact same LSM cut-offs could be adopted for high and normal ALT patients. In fact, the diagnostic accuracy is reduced if cut-off adjustment is not made for high ALT cases. Specifically, in high ALT patients, $F \geq 3$ and F4 diagnostic accuracy is reduced, although is relatively unaffected for $F \geq 1$ and $F \geq 2$. The optimal cut-off $F \geq 3$ and F4 calculated for the entire cohort of patients were 9.7kpa (84.2% sin, 92.3% sp) and 11.9kpa (100% sin, 84.8% sp). If these same cutoffs were adopted for high ALT patients, the performance (determined by examining the ROC curve coordinates) for diagnosing $F \geq 3$ has 80% sensitivity and 80.1% specificity, compared to the preferred high ALT specific optimal cut-off of 10.5kpa, which has the same 80% sensitivity but greater specificity of 85.2%. For F4, the cut-off of 11.9kpa has 100% sensitivity, and 78.8% specificity compared to the preferred high ALT specific optimal cut-off of 15.9kpa which also has 100% sensitivity but greater specificity of 93.9%. Therefore, in order to remain optimal, the value of the cut-offs in high ALT

patients need to be increased by approximately a factor of 1.3x, which is identical to reports by Wong et al. [28]. A statistically significant AUROC for F4 for normal ALT subjects was unable to be determined in this study. Studies on a larger population will allow for cut-offs to be further refined.

Dual liver stiffness measurement cut-offs for diagnosing moderate and advanced fibrosis and grey zones

Due to the significance of liver fibrosis assessment in clinical practice, a non-invasive diagnostic tool should be able to rule out or rule in the fibrosis stage with a high degree of certainty. Accordingly, the most rational way to use noninvasive methods is first to assess whether patients can be diagnosed with high accuracy, and then use liver biopsy when the accuracy is not at an acceptable level. To reduce the need of liver biopsy by non-invasive methods, a high sensitivity cutoff for excluding fibrosis and a high specificity cutoff for confirming fibrosis should be determined. Therefore, the diagnostic cutoff for liver fibrosis assessment should use dual cutoffs rather than a single cutoff. Patients with test results in the grey zone between low and high cutoffs would be left undiagnosed and may require a liver biopsy.

In our study, the grey zones for in the normal ALT group for $F \geq 2$ was 4.8 - 9.6 kpa and $F \geq 3$ 6.0 - 9.7 kpa. In the abnormal ALT group, the grey zones for $F \geq 2$ was 5.8 - 12.3 kPa and for $F \geq 3$ 5.8 - 12.5 kpa. These LSM cutoffs corresponded with the ability to diagnose $F \geq 2$ and $F \geq 3$ at least a 90% degree of certainty. For the entire cohort, patients with LSM outside these grey zone ranges, and thus were able to be diagnosed with certainty, reached 49.3% for ruling in or ruling out moderate fibrosis and 57.7% for advanced fibrosis.

Chan et al. [38] proposed grey zones for advanced fibrosis that were slightly narrower. Normal ALT grey zone cutoffs: $F \geq 3$ 6.0 - 9.0kpa, and abnormal ALT grey zone cutoffs: $F \geq 3$ 7.5 - 12.0 kPa. In their study, TE was able to diagnose or exclude $F \geq 3$ in 62% and 58% of normal and abnormal ALT subjects respectively. Subsequently a liver biopsy was required only in 38% and 42% in normal and abnormal ALT CHB subjects [38]. Our study used cut-offs with a higher sensitivity for excluding advanced fibrosis compared with Chan et al: 100% vs. 93% in normal ALT patients; and 100% vs. 96% in abnormal ALT patients respectively. Though, our specificity was lower in abnormal ALT subjects: 92.6% vs 100%; it was the same in normal ALT subjects (100% vs. 100%). The method for choosing the optimal LSM has a rational basis, but also includes some clinical discretion. This author used the criteria of choosing the cut-off with a sensitivity/specificity of minimum 90%, while also optimizing the corresponding specificity/sensitivity. An alternate method would be to simply choose the highest/lowest possible LSM with at least 90% sensitivity/specificity, while ignoring the corresponding specificity/sensitivity. This method is not invalid since the focus is singularly on either optimizing the sensitivity

or specificity in turn. This approach would result in narrower grey zones at the cost of a lower degree of diagnostic accuracy. If this were to be adopted, the revised grey zone would be 6.0 – 9.0 kPa (for normal ALT subjects), which also happens to be identical to the Chan et al grey zone. For abnormal ALT, the revised grey zone would be 6.0 - 12.5kPa. Using these alternate values, the proportion of those that can diagnosed/excluded for F3 with at least 90% certainty increases from 57.7% to 67.6%. This would be at the expense of a slightly lower degree of certainty. Thus, apart from variation in the study population, the differences in how optimal LSM cutoffs are chosen can also account for the wider grey zone ranges.

Although 57.7% of patients can be diagnosed/excluded for F \geq 3 using these grey zone cutoffs, in clinical practice, the proportion of CHB patients avoiding a liver biopsy may be even greater. The decision to initiate antiviral therapy also rests upon whether patients have persistently abnormal ALT, HbeAg status and viral load. Potentially, many patients with high ALT and have grey zone LSM would be treated irrespective of the LSM. A grey zone LSM would cause some uncertainty regarding whether these patients truly F3, but performing a liver biopsy would not change the need for antiviral therapy. Arguably, the management is not changed with or without a liver biopsy, and many clinicians would choose not to perform one in these circumstances. This notion was demonstrated in an analysis of the LSM results of local CHB patients using Chan's algorithm along with serial ALT revealed that only 9/47 (19.1%) patients would require a liver biopsy. An implication for TE being able to reduce liver biopsies is that not only an invasive procedure may be avoided, but there would be associated cost-savings, which was estimated to be \$AUD 74 214 [39].

Liver stiffness is superior to fib-4 index, aspartate platelet ratio index, aspartate alanine aminotransferase ratio, age platelet index and fibrosis index

LSM had by far the most superior diagnostic accuracy in the estimation for F stages 1 through to 4 compared to the other noninvasive measures: FIB-4, APRI, API, AAR and FI. FIB-4 had the next best diagnostic performance, with an excellent AUROC for F4 (0.912) and fair accuracy for the F \geq 1, 2 and 3 (0.635-0.711). Third best was APRI which had good accuracy for F4 (0.821), but only fair accuracy for F \geq 1, 2 and 3 (0.698-0.720). Fourth was API which had an excellent AUROC for cirrhosis (0.900), but was completely unable to diagnose advanced fibrosis or significant fibrosis with AUROCS being too poor and non-significant. AAR and FI were poor measures overall and should not be used as diagnostic tests for fibrosis in CHB patients. Interpretation of these results may be limited by the small sample size.

FIB4-I was originally developed for a HCV-HIV co-infected cohort [30]. In 2010 Kim et al. performed a comparison study in 668 CHB patients with other simple non-invasive markers [40]. Their study found superior AUROCs for: F \geq 2 = 0.865, F \geq 3 = 0.910 and F4 = 0.926, which would make FIB4-I an excellent

overall non-invasive tool. Their study also found AUROC=0.928 for API which is excellent for F4. Furthermore, they were able to determine API was excellent for diagnosing advanced fibrosis (AUROC = 0.901). APRI and AAR were found to be fair tests to determine severe fibrosis and cirrhosis. APRI AUROC = 0.702 and 0.731 and AAR AUROC = 0.724 and 0.729 respectively for advanced fibrosis and cirrhosis. Their study has the advantage of a large sample size, and found generally better results for non-invasive markers other than TE that was analysed in our study. However there was no head to head comparison with TE.

API showed excellent performance for the diagnosis of F4 in another study by the same author Kim, who reported an AUROC = 0.93 [41]. A much weaker result by Chen et al was found with an AUROC of 0.77 [42] API has not been reported to be able to diagnose F \geq 2 except for one other study, which found a fair accuracy (AUROC = 0.77) [43].

In 2015 a meta-analysis examined 34 studies of APRI and FIB-4 with a total of 8855 patients [44]. The summary AUROC values for using APRI and FIB-4 for the diagnosis of F2, F3 and F4 are as follows:

APRI: F2 0.7407 (95% CI 0.7033-0.7781); F3 0.7844 (95% CI 0.7450-0.8238) and F4 0.7347 (95% CI 0.6790-0.7904).

FIB4: F2 0.8165 (95% CI 0.7707-0.8623); F3 0.7268 (95% CI 0.6578 - 0.7958) and F4 0.8448 (95% CI: 0.7742-0.9154).

The meta-analysis indicates that APRI and FIB-4 can identify hepatitis B related fibrosis and cirrhosis with only a fair to moderate degree of accuracy. The results of our study are more consistent with the overall body of literature compared to Kim et al. [40].

Transient elastography compared other non-invasive assessment of fibrosis in chronic hepatitis B patients – current literature

Since the conception of this study in 2009, there have been numerous developments on non-invasive markers for chronic hepatitis B. Of all these, the Fibro test (FT) has been the second most widely studied behind TE. A meta-analysis by Poynard [45] analysed 1842 CHB patients with liver biopsies across 8 studies and compared FT with LSM (5 studies, 618 patients). For the diagnosis of advanced fibrosis F3, AUROC was 0.84 (0.79–0.86) for FT and 0.89 (0.83–0.96) for LSM. Although TE had a numerically superior AUROC, there was no statistical difference. A later head to head comparison in 179 Australian and French CHB patients of FT with hepscore and other serum based markers found that FT was inferior and only had an AUROC of 0.72 compared with hepscore AUROC 0.83 [46]. On the other hand, a more recent head to head comparison of 194 Korean CHB patients between FT and LSM found AUROCs of FT were 0.903, 0.907, and 0.866, comparable to those of LSM: 0.873, 0.897, and 0.910 for F \geq 2, 3 and 4 respectively. This study found that by combining the 2 markers by multiplying the FT and LSM

showed the best AUROCs: 0.941, 0.931, and 0.929 for F \geq 2, 3, and 4, respectively [47].

Of the remaining potential non-invasive measures, a review by Chen [36] identified serum markers which had at least been independently validated with an AUROC of at least 0.85. Based on the reviewer's criteria, API and FIB4-I fulfill these requirements. API and FIB 4 have been discussed in the previous section. Overall API appears to be having a good diagnostic accuracy for F4 only. FIB 4 in a pooled meta-analysis show results indicating that it has fair to moderate diagnostic accuracy.

Other tests that fulfill these criteria were reported to include the Forms index, Hepascore, Fibro meter, Zing Index and Hue Index. In summary, these all report fair to good diagnostic performance for F \geq 2 (AUROC: 0.72-0.81), fair to good diagnostic performance for F \geq 3 (AUROC: 0.75-0.89) and good to excellent performance for F4 (0.89-0.93). Results are promising, but more studies are required for these markers. None are close to the level of repeated validation that TE has received. Further research into combining the best non-invasive markers for fibrosis in CHB patients may reveal potentially even more accurate combinations.

Controversies of non-invasive markers compared to Fibrosis stage and collagen proportionate area

TE may be better used as a standalone marker in the future because comparison of LSM with fibrosis stage is a flawed concept. This is because of several reasons.

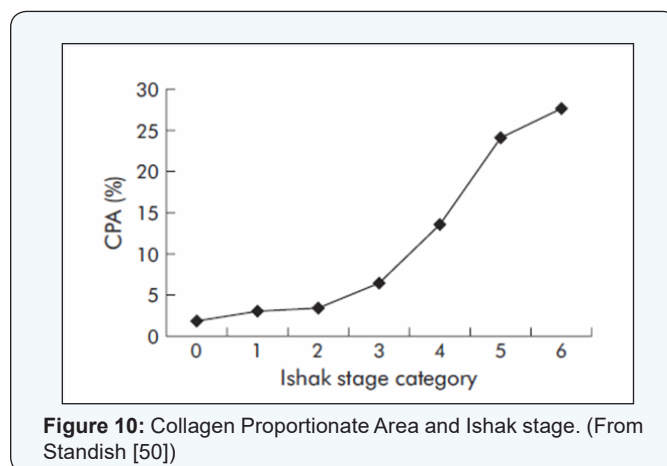
Firstly, the problems with liver histology assessment limit the accuracy of any comparisons with non-invasive measures. This study has highlighted the vast degree of inter-observer variability that may occur. Sampling error because of inadequate biopsy length [19] and patchy disease [20] are the other main causes of inaccurate assessment. Because of these limitations, some experts feel that an AUROC >0.90 cannot be truly achieved even for a perfect marker [48]. Many are of the opinion that non-invasive fibrosis tests with an AUROC of 0.85-0.90 are as good as liver biopsy for staging liver fibrosis [49].

Another reason for why LSM comparison to fibrosis stage is flawed relates to histological staging being a category that has ordinal features, but is strictly not a continuous variable. Essentially, LSM being a continuous variable is being compared to ordinal histological categories which render comparison to be awkward and some claim to be statistically flawed [50]. Whichever scoring system that is used, fibrosis staging is a qualitative morphological assessment of the distribution of fibrosis. There is no mix of features that includes the amount of fibrosis in the description of each stage. As the severity of liver disease progresses, the fibrosis distribution generally expands and hence fibrosis quantity increases with each stage. Each stage describes incrementally more extensive fibrosis distribution and so by corollary there is an increase in the quantity. But by no means do F stages scale linearly in a proportionate manner.

F4 certainly does not imply twice the fibrosis quantity of F2 for instance.

This concept is well illustrated by which shows the relationship between Ishak stage scores and measured quantity of fibrosis. The quantity of fibrosis, as denoted by the collagen proportionate area (CPA) and Ishak category clearly do not have a linear relationship. They are related, but in the end are different evaluations.

This author has described how our results and other studies consistently show LSM correlates better with latter stages of fibrosis rather than earlier stages of fibrosis. Wong et al. [28] study that shows LSM correlates better with per cellular fibrosis is only partly the answer. The over-arching reason for the poor correlation is that the relationship, as shown in Figure 10, is not linear or proportional between each stage. It appears to increase slowly in the earlier stages from Ishak 0 to 2, before the slope increases between Ishak 2 to 3. Then the slope becomes increasingly steeper between Ishak 3 to 4 and Ishak 4 to 5, before flattening out slightly between Ishak 5 to 6. This may explain why we observe that not only the LSM increases, but the rate of LSM increases with more advanced stages of fibrosis. It may also be the explanation why there are such tight LSM cut-offs between the early stages of fibrosis, compared to the latter stages of fibrosis. All studies, including our own, report a very wide range of LSM for F4, but narrow for F1. In our cohort, F4 had a wide range with any LSM from 11.9 – 75.0 kpa. In comparison F1 fell in the tight range of 6.5-7.4 kpa. Any non-invasive measure designed to be a marker of fibrosis quantity, but is then compared to fibrosis stage, will also be exposed to the inherent flaws.



Due to the limitations of histological staging, it is clear that proper measurement of liver collagen is unavoidable. Methods for histologically quantifying liver fibrosis are still in development. The most appropriate and practicable method appears to be using computer assisted imaging analysis (IA) of histologically stained sections. IA uses segmentation of digital images to measure the area of collagen and the area of tissue, producing a "fibrosis ratio" or collagen proportionate area (CPA).

To illustrate the point that non-invasive measures such as Fibroscan correlate better with a marker of fibrosis volume such as CPA rather than fibrosis stage, a study carried out by Isgro et al. showed that LSM was better predicted by CPA in CHB patients ($r^2=0.61$) compared to Ishak ($r^2=0.52$) [51].

And yet despite all these issues, non-invasive measures are still widely and routinely compared with the histological fibrosis stage and published in prominent research journals. The F stage is still the simplest, most established and common way of assessing fibrosis severity, although it may be misunderstood and misused as a quantifiable measure. Comparing all non-invasive measures with fibrosis stage is entrenched because it is meaningful. Fibrosis stage has been well correlated with prognosis in all forms of chronic liver disease. Collagen Proportionate area, although a more scientifically sound way of comparison, is an unknown quantity except to hepatologists and histopathologists with a special interest in this area. Few data exists that allow us to make prognostic assumptions based on the CPA. Thus despite being flawed, it is still useful to compare with fibrosis stage.

However, this begs the question upon whether LSM is better served as being considered as stand-alone outcome measure. As we have demonstrated, LSM being used to indirectly gauge disease severity through correlation with fibrosis stage is a flawed comparison. LSM comparison to CPA is a much more scientifically sound and biologically compatible. However, methods to even measure CPA are still in the early stage of research, let alone whether it can be used to reflect prognostic implications. Longitudinal studies which analyse the relationship between LSM and patient end points over long term follow up are required to determine the usefulness of the LSM being considered a stand-alone marker.

Conclusion

Fibroscan is a reliable and accurate non-invasive tool for diagnosing fibrosis stage. It has excellent accuracy for F4 and $F \geq 3$ and is superior to FIB-4, APRI, API, AAR and FI. It can reduce the need for liver biopsies in the majority of chronic hepatitis B patients. Fibroscan is the most widely studied non-invasive measure for fibrosis in chronic hepatitis B patients. Combining different non-invasive measures, such as the caffeine breath test, may improve the accuracy even further but further research is required. The comparison of Fibroscan with histological staging is a flawed concept, but one that has traction because of the well described prognosis of fibrosis stages.

Authorship

Guarantor of the article: Dr Raymond Kwok.

Author contributions: Dr Raymond Kwok performed the research, collected and analyzed the data and wrote the manuscript. Dr Alice Lee and Dr Meng Ngu gave critical comments to the manuscript. Dr Raymond Kwok and Dr Alice Lee conceived

and designed the study. All authors approved the final version of the manuscript.

References

- 1 Lavanchy D (2004) Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. *J_Viral_Hepat* 11(2): 97-107.
- 2 Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, et al. (2013) Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 381(9867): 628.
- 3 Lavanchy D (2004) Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. *J Viral Hepat* 11(2): 97-107.
- 4 Liaw YF, Kao JH, Piratvisuth T, Chan HLY, Chien RN, et al. (2012) Asian-Pacific consensus statement on the management of chronic hepatitis B: a 2012 update. *Hepatol Int* 6(3): 531-561.
- 5 Chen CJ, Yang HI (2011) Natural history of chronic hepatitis B REVEALed. *J Gastroenterol Hepatol* 26(4): 628-638.
- 6 Ganem D, Prince AM (2004) Hepatitis B virus infection--natural history and clinical consequences. *N Engl J Med* 350(11): 1118-1129.
- 7 Liaw YF (2005) Prevention and surveillance of hepatitis B virus-related hepatocellular carcinoma. *Semin Liver Dis* 25(Suppl 1): 40-47.
- 8 Hoofnagle JH, Doo E, Liang TJ, Fleischer R, Lok AS (2007) Management of hepatitis B: summary of a clinical research workshop. *Hepatology* 45(4): 1056-75.
- 9 Liaw YF, Jia JD, Chan HL, Han KH, Tanwandee T, et al. (2011) Shorter durations and lower doses of peginterferon alfa-2a are associated with inferior hepatitis B e antigen seroconversion rates in hepatitis B virus genotypes B or C. *Hepatology* 54(5): 1591-1599.
- 10 Ono A, Suzuki F, Kawamura Y, Sezaki H, Hosaka T, et al. (2012) Long-term continuous entecavir therapy in nucleos(t)ide-naive chronic hepatitis B patients. *J Hepatol* 57(3): 508-514.
- 11 Marcellin P, Gane E, Buti M, Afdhal N, Sievert W, et al. (2013) Regression of cirrhosis during treatment with tenofovir disoproxil fumarate for chronic hepatitis B: a 5-year open-label follow-up study. *The Lancet* 381(9865): 468-475.
- 12 Chung KT, Ha NB, Trinh HN, Garcia RT, Nguyen HA, et al. (2012) High frequency of recurrent viremia after hepatitis B e antigen seroconversion and consolidation therapy. *J Clin Gastroenterol* 46(10): 865-870.
- 13 Fung J, Lai CL, Tanaka Y, Mizokami M, Yuen J, et al. (2009) The duration of lamivudine therapy for chronic hepatitis B: cessation vs. continuation of treatment after HBeAg seroconversion. *Am J Gastroenterol* 104(8): 1940-6.
- 14 Seto WK, Hui AJ, Wong VW, Wong GL, Liu KS, et al. (2014) Treatment cessation of entecavir in Asian patients with hepatitis B e antigen negative chronic hepatitis B: a multicentre prospective study. *Gut* 64(4): 667-672.
- 15 Australia DHFotGSo (2009) Australian and New Zealand Chronic Hepatitis B (CHB) Recommendations. (2nd edn).
- 16 Terrault NA, Bzowej NH, Chang KM, Hwang JP, Jonas MM, et al. (2016) AASLD guidelines for treatment of chronic hepatitis B. *Hepatology* 63(1): 261-283.
- 17 Liaw YF (2009) Antiviral therapy of chronic hepatitis B: opportunities and challenges in Asia. *J Hepatol* 51(2): 403-410.
- 18 Lok AS, McMahon BJ (2009) Chronic hepatitis B: update 2009. *Hepatology* 50(3): 661-662.

- 19 Colloredo G, Guido M, Sonzogni A, Leandro G (2003) Impact of liver biopsy size on histological evaluation of chronic viral hepatitis: the smaller the sample, the milder the disease. *J Hepatol* 39(2): 239-244.
- 20 Regev A, Berho M, Jeffers LJ, Milikowski C, Molina EG, et al. (2002) Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. *Am J Gastroenterol* 97(10): 2614-2618.
- 21 Bedossa P, Dargere D, Paradis V (2003) Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 38(6): 1449-1457.
- 22 Westin J, Lagging LM, Wejstal R, Norkrans G, Dhillon AP (1999) Interobserver study of liver histopathology using the Ishak score in patients with chronic hepatitis C virus infection. *Liver* 19(3): 183-187.
- 23 Gilmore IT, Burroughs A, Murray-Lyon IM, Williams R, Jenkins D, et al. (1995) Indications, methods, and outcomes of percutaneous liver biopsy in England and Wales: an audit by the British Society of Gastroenterology and the Royal College of Physicians of London. *Gut* 36(3): 437-441.
- 24 McGill DB, Rakela J, Zinsmeister AR, Ott BJ (1990) A 21-year experience with major hemorrhage after percutaneous liver biopsy. *Gastroenterology* 99(5):1396-1400.
- 25 Piccinino F, Sagnelli E, Pasquale G, Giusti G (1986) Complications following percutaneous liver biopsy. A multicentre retrospective study on 68,276 biopsies. *J Hepatol* 2(2): 165-173.
- 26 Eisenberg E, Konopnik M, Veitsman E, Kramskay R, Gaitini D, et al. (2003) Prevalence and characteristics of pain induced by percutaneous liver biopsy. *Anesth Analg* 96(5): 1392-1396.
- 27 Marcellin P, Ziol M, Bedossa P, Douvin C, Poupon R, et al. (2009) Non-invasive assessment of liver fibrosis by stiffness measurement in patients with chronic hepatitis B. *Liver Int* 29(2): 242-247.
- 28 Wong GL, Wong VW, Choi PC, Chan AW, Chum RH, et al. (2008) Assessment of fibrosis by transient elastography compared with liver biopsy and morphometry in chronic liver diseases. *Clin Gastroenterol Hepatol* 6(9): 1027-1035.
- 29 Boursier J, Zarski JP, de Ledinghen V, Rousselet MC, Sturm N, et al. (2013) Determination of reliability criteria for liver stiffness evaluation by transient elastography. *Hepatology* 57(3): 1182-1191.
- 30 Vallet-Pichard A, Mallet V, Nalpas B, Verkarre V, Nalpas A, et al. (2007) FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. *Hepatology* 46(1): 32-36.
- 31 Wai C-T, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, et al. (2003) A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 38(2): 518-526.
- 32 Poynard T, Bedossa P (1997) Age and platelet count: a simple index for predicting the presence of histological lesions in patients with antibodies to hepatitis C virus. METAVIR and CLINIVIR Cooperative Study Groups. *J Viral Hepat* 4(3): 199-208.
- 33 Williams AL, Hoofnagle JH (1988) Ratio of serum aspartate to alanine aminotransferase in chronic hepatitis. Relationship to cirrhosis. *Gastroenterology* 95(3): 734-739.
- 34 Ohta T, Sakaguchi K, Fujiwara A, Fujioka S, Iwasaki Y, et al. (2006) Simple surrogate index of the fibrosis stage in chronic hepatitis C patients using platelet count and serum albumin level. *Acta Medica Okayama* 60(2): 77-84.
- 35 Fraquelli M, Rigamonti C, Casazza G, Conte D, Donato MF, et al. (2007) Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. *Gut* 56(7): 968-973.
- 36 Chen YP, Peng J, Hou JL (2013) Non-invasive assessment of liver fibrosis in patients with chronic hepatitis B. *Hepatol Int* 7(2): 356-368.
- 37 Chon YE, Choi EH, Song KJ, Park JY, Kim do Y, et al. (2012) Performance of transient elastography for the staging of liver fibrosis in patients with chronic hepatitis B: a meta-analysis. *PLoS One* 7(9): e44930.
- 38 Chan HL, Wong GL, Choi PC, Chan AW, Chim AM, et al. (2009) Alanine aminotransferase-based algorithms of liver stiffness measurement by transient elastography (Fibroscan) for liver fibrosis in chronic hepatitis B. *J Viral Hepat* 16(1): 36-44.
- 39 Guirgis M, Manoharan S, Kwok R, Scott DR, Lee AU, et al. (2009) Transient Elastography (FIBROSCAN) in chronic hepatitis B infection: Reduce number of biopsies and save money. *Hepatology* 50: 543A.
- 40 Kim BK, Kim do Y, Park JY, Ahn SH, Chon CY, et al. (2010) Validation of FIB-4 and comparison with other simple noninvasive indices for predicting liver fibrosis and cirrhosis in hepatitis B virus-infected patients. *Liver Int* 30(4): 546-553.
- 41 Kim BK, Kim SA, Park YN, Cheong JY, Kim HS, et al. (2007) Noninvasive models to predict liver cirrhosis in patients with chronic hepatitis B. *Liver International* 27(7): 969-976.
- 42 Chen YP, Liang XE, Dai L, Zhang Q, Peng J, et al. (2012) Improving transient elastography performance for detecting hepatitis B cirrhosis. *Dig Liver Dis* 44(1): 61-66.
- 43 Shin WG, Park SH, Jang MK, Hahn TH, Kim JB, et al. (2008) Aspartate aminotransferase to platelet ratio index (APRI) can predict liver fibrosis in chronic hepatitis B. *Dig Liver Dis* 40(4): 267-274.
- 44 Xiao G, Yang J, Yan L (2015) Comparison of diagnostic accuracy of aspartate aminotransferase to platelet ratio index and fibrosis-4 index for detecting liver fibrosis in adult patients with chronic hepatitis B virus infection: a systemic review and meta-analysis. *Hepatology* 61(1): 292-302
- 45 Poynard T, Ngo Y, Munteanu M, Thabut D, Ratziu V (2011) Noninvasive markers of hepatic fibrosis in chronic hepatitis B. *Curr Hepat Rep* 10(2): 87-97.
- 46 Raftopoulos SC, George J, Bourliere M, Rossi E, de Boer WB, et al. (2012) Comparison of noninvasive models of fibrosis in chronic hepatitis B. *Hepatol Int* 6(2): 457-467.
- 47 Kim BK, Kim SU, Kim HS, Park JY, Ahn SH, et al. (2012) Prospective validation of FibroTest in comparison with liver stiffness for predicting liver fibrosis in Asian subjects with chronic hepatitis B. *PLoS ONE* 7(4): e35825.
- 48 Mehta SH, Lau B, Afdhal NH, Thomas DL (2009) Exceeding the limits of liver histology markers. *J Hepatol* 50(1): 36-41.
- 49 Afdhal NH, Nunes D (2004) Evaluation of liver fibrosis: a concise review. *Am J Gastroenterol* 99(6): 1160-1174.
- 50 Standish RA, Cholongitas E, Dhillon A, Burroughs AK, Dhillon AP (2006) An appraisal of the histopathological assessment of liver fibrosis. *Gut* 55(4): 569-578.
- 51 Isgro G, Calvaruso V, Andreana L, Luong TV, Garcovich M, et al. (2013) The relationship between transient elastography and histological collagen proportionate area for assessing fibrosis in chronic viral hepatitis. *J Gastroenterol* 48(8): 921-929.