Clinical Utilities of Transient Elastography

Dr Raymond Kwok

MBBS BSc FRACP

A Thesis Submitted for the Degree of

Doctor of Philosophy

Faculty of Medicine

The University of Sydney

2017

Declaration

I hereby declare that all work presented in this thesis describes original research undertaken at:

1. Gastroenterology Department of Concord Repatriation Hospital, Sydney Medical School, The University of Sydney under the supervision of Associate Professor Meng Ngu and Associate Professor Alice Lee.

AND

2. Institute of Digestive Diseases Department of Medicine and Therapeutics, Prince of Wales Hospital Shatin Hong Kong Special Administrative Region, The Chinese University of Hong Kong under the supervision of Professor Vincent Wong, Professor Grace Wong and Professor Henry Chan.

This thesis has not been submitted for any other degree or other purposes. The author performed all the studies within this thesis.

Raymond Kwok

MBBS BSc FRACP

September 30th, 2016

Dedication

This is by far the greatest work I have ever achieved. I dedicate this to my dear wife Esther. I owe her several lifetimes' worth of gratitude. She has stuck by me and sacrificed the most to support me to complete my academic goals. Love you always.

To my kids Emilene and Nicholas, I hope you can look at my work when you are older and understand that accomplishments are not about being the smartest or the strongest, but about being the most persistent.

Acknowledgements

Reaching the end of this long journey required much work and sacrifice. But most of all, it required help. I was only able to get here because I had great support from so many people along the way.

First of all, acknowledgement must go to Professor Henry Lik-Yuen Chan, Vincent Wai-Sun Wong and Grace Lai-Hung Wong. You provided the infrastructure, guidance, and expert perspectives that brought my PhD to another level, enabling publication in prominent journals. It was truly an honour to work alongside you. My time spent in Hong Kong was some of the most fun I ever had. I have learnt invaluable lessons in what makes a successful leadership and team culture.

Thank you to those many who have provided me help in the form of statistical assistance, data collection, patient recruitment and insightful opinions. This includes Yee Kit Tse, Yeon Jung Ha, Yuying Zhang, Andrea On-Yan Luk, Sally She-Ting Shu, Anthony Wing-Hung Chan, Ming-Wai Yeung, Juliana Chung-Ngor Chan, Alice Pik-Shan Kong, Veronica Gomez, Ken Liu, Bain Shenstone, Anna Cunningham, Christian and Christina Guirgis, Roger Chen and Gordon Park.

And finally, my biggest thanks go to Associate Professor Meng Chong Ngu and Associate Professor Alice Unah Lee. I appreciate the many years of support and I am indebted to both of you for the opportunity and endorsement. I am grateful for the insightful feedback and opinions. I thank you for the flexibility and the autonomy you have given me, without which I could not have completed this thesis.

Ethics approval

All studies have been approved by the corresponding local ethics boards.

The Sydney South West Area Health Service Human Research Ethics Committee approved of the studies presented in chapter 2, 5, 6 and 7 of this thesis.

The Joint Chinese University of Hong Kong – Northern Territories East Cluster Research Ethics Committee approved of the study presented in chapter 4 of this thesis.

General Abstract

Introduction: Chapter 1

Chronic liver disease causes 1.75 million deaths globally and is within the top 10 leading causes of death in middle income countries. Chronic liver injury occurs via a process of inflammation and fibrosis formation. Patients often do not present to healthcare until advanced stages of disease and when there is already decompensated cirrhosis. Liver biopsy has been used to identify earlier stages of fibrosis, but it is poorly accepted by patients and has limitations. Transient Elastography (TE) using Fibroscan ® is a non-invasive tool for the diagnosis liver fibrosis. The clinical application of Fibroscan in non-alcoholic fatty disease (NAFLD), chronic hepatitis B (CHB), and methotrexate induced liver fibrosis were examined.

Clinical Utility of Transient Elastography in non-alcoholic fatty liver disease: Chapters 2, 3 and 4

Non-alcoholic fatty liver disease affects 20-35% of the global population, but only a small subset develop the histological subtype of non-alcoholic steatohepatitis (NASH), which can lead to progressive liver disease by causing fibrosis and eventually cirrhosis. Fibroscan can potentially identify those patients who have fibrosis and are at increased risk of further progression. Patients with type 2 diabetes, who are at high risk of NASH, were assessed. A liver stiffness measurement (LSM) ≥9.8 kPa, used as a cut-off for advanced fibrosis (1), was found in 12% (10/77) of subjects. Higher LSM readings correlated with higher BMI and the use of insulin therapy. Patients on insulin had LSM ≥9.8 kPa with likelihood ratio (LR): 12.3, p=0.002 **(Chapter 2)**. The study was limited by a small sample size, and a high failure rate as the medium (M) probe was only available.

A systemic review evaluating all non-invasive methods for diagnosing NASH and NAFLD fibrosis was undertaken. This included a meta-analysis that focused on what was found to be the most widely studied markers of NASH and NAFLD fibrosis: cytokeratin-18 (CK-18) fragments and transient elastography respectively **(Chapter 3).** Not only was TE found to be the most extensively studied, it had also one of the highest diagnostic accuracies with pooled sensitivities and specificities to diagnose F≥2, 3 and 4 to be: 79% and 75%, 85% and 85%, and 92% and 92% respectively.

We then proceeded to perform a much larger study in diabetic subjects using the latest generation of Fibroscan ® 502 touch model **(Chapter 4)**. This included the extra-large (XL) probe for obese subjects and also featured the novel Controlled Attenuation Parameter (CAP), which assesses liver steatosis. A total of 1918 diabetes patients at Prince of Wales Hospital, Hong Kong were recruited. Each had a TE and CAP to assess liver stiffness and steatosis. Reliable scans were achieved in 98.2% of patients using the M or XL probes. The proportion of patients with increased CAP (suggestive of steatosis) and increased LSM (suggestive of advanced fibrosis) were 72.8% and 17.7% respectively. By multivariate analysis, female gender, higher body mass index, triglycerides, fasting plasma glucose and alanine aminotransferase, and non-insulin use were associated with increased CAP. Longer duration of diabetes, higher body mass index, alanine aminotransferase, spot urine albumin-creatinine ratio, and lower high-density lipoprotein-cholesterol were associated with increased LSM. The17.7% prevalence of advanced fibrosis suggests type 2 diabetic patients would benefit from routine screening for liver disease**.**

Clinical Utility of Transient Elastography in chronic hepatitis B: Chapters 5 and 6

Transient elastography was initially applied for staging patients with chronic hepatitis C (CHC) with data rapidly growing on its utility for the assessment in patients with CHB infection. Our study contributes to this by further evaluating the diagnostic accuracy and usefulness of TE, and also comparing its performance against the FIB4 index, Aspartate Platelet Ratio Index (APRI), Aspartate Alanine Aminotransferase Ratio (AAR), Age Platelet Index (API), Fibrosis Index (FI) and Caffeine Breath Test (CBT) **(Chapter 5)**.

In 71 CHB patients, the diagnostic performance of the LSM for Metavir fibrosis stage F≥1, 2, 3 and 4 were: Area under Receiver Operator Characteristic (AUROC) = 0.825, 0.792, 0.874 and 0.945 respectively. Patients with high ALT required higher LSM cut-offs. Dual cut-offs are needed to "rule in" and to "rule out" stage of fibrosis with a high level of certainty. Using normal vs high ALT specific cut-offs, F≥2 and F≥3 can be "ruled in" or "ruled out" with certainty in 49.3% and 57.7% of CHB patients respectively. TE was the superior non-invasive test when compared with FIB4-I, APRI, API, AAR and FI. Caffeine breath test compared well against TE in a small cohort, but is not as practical. Liver histology is limited by interobserver variability, with 44% of liver biopsies being classified a different stage on second evaluation, and the intraclass correlation coefficient showing moderate agreement (K =0.457). Although routinely compared, this highlights the limitations of assessing the accuracy of TE and other non-invasive tests against a reference standard that has such a degree of variation.

The use of TE in the longitudinal monitoring of fibrosis is important in the follow up of patients with CHB **(Chapter 6)**. Current literature was conflicting and seemed to suggest that decline in LSM was influenced more by the fall in ALT with decline in necroinflammatory activity, rather than fibrosis regression. We sought to evaluate the factors that affected LSM change and assess which clinical subgroups experienced an LSM decline.

In 124 CHB patients who were followed for 31.2 months (SD 13.1), LSM decline was greatest in those who had active disease and were subsequently treated with antivirals. This is associated with ALT normalization, HBeAg seroconversion and viral suppression. In CHB patients with quiescent disease - ie did not require antiviral treatment, or who had persistently normal ALT irrespective of treatment - only a small or non-significant decline in LSM was observed. The change in LSM was strongly correlated with length of time and may suggest fibrosis regression. Further studies are required, as our findings are limited by a lack of correlation with liver biopsy, and the low baseline levels of liver stiffness in those with inactive CHB.

Clinical Utility of Transient Elastography in methotrexate induced liver fibrosis: Chapter 7

Long term use of methotrexate has been associated with risk of liver fibrosis and the role of TE in this cohort was evaluated. The relationship between liver fibrosis and methotrexate dose, and other factors associated with moderate fibrosis (F2) using an LSM cut-off of ≥7.1 kPa were examined.

In 39 patients with a mean intake dose of 5.3g of methotrexate, no correlation was found between the LSM and the cumulative dose or duration of treatment. Of the 7/39 cases of LSM≥7.1 kPa (17.9%), BMI≥30 was the only risk factor with a likelihood ratio (LR) of 4.442, p=0.029. One patient had cirrhosis (2.6%). This is much lower than rates reported from early studies [26% (2, 3)], and more in line with recent data [around 2% (4)], and lends support to the suggestion that early studies overestimated the risk of methotrexate induced fibrosis due to lack of controls for pre-existing liver

disease (5). There was also no difference in the LSM of methotrexate subjects and matched population controls.

Conclusion

Our studies lend further support to the utility of LSM on identifying those at increased risk of liver fibrosis progression, which will continue to remain a significant clinical challenge in both individuals and as a public health burden. In particular we feel that major contributions have been made on the subject of screening for advanced fibrosis in a high-risk population of type II diabetic patients. Our longitudinal studies on the role of using TE in follow up and comparing its performance in CHB patients are also significant. Despite the small cohort of methotrexate users, this further supports the utility of TE in a wide range of liver diseases that manifest with progressive fibrosis. The next area of further development in the clinical use of TE is as a stand-alone marker that has prognostic significance.

Publications

1. **Screening diabetic patients for non-alcoholic fatty liver disease with controlled attenuation parameter and liver stiffness measurements: a prospective cohort study.**

   Kwok R, Choi KC, Wong GL, Zhang Y, Chan HL, Luk AO, Shu SS, Chan AW, Yeung MW, Chan JC, Kong AP, Wong VW

   Gut 2016 Aug; 65(8):1359-68

   This publication forms chapter 4 of this thesis. Raymond Kwok is the corresponding author and was primarily responsible for recruiting subjects, performing the study, collecting data, interpreting the analysis and writing the drafts. The study was co-designed the study with the co-authors.

2. **Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease - the role of transient elastography and plasma cytokeratin-18 fragments**.

   Kwok R, Tse YK, Wong GL, Ha Y, Lee AU, Ngu MC, Chan HL, Wong VW

   Alimentary pharmacology & therapeutics 2014; 39(3):254-69.

   This publication forms chapter 3 of this thesis. Raymond Kwok is the corresponding author and was primarily responsible for the literature review, data extraction, and writing the drafts. The study was co-designed and analysis co-interpreted with the co-authors.

Oral Presentations and Conference Posters

3. **The Role of Fibroscan and Controlled Attenuation Parameter (CAP) in Fatty Liver Disease**

   GiHep Singapore 2016 Annual Scientific Meeting

   Liver symposium oral presentation

4. **Clinical Utilities of Transient Elastrography**

   Kwok R, Gonzalez-Arce V, Ngu MC, Lee AU.

   Australian Liver Association 2009 Conference (Yarra Valley, Victoria)

   Oral presentation

5. **Should we screen diabetic patients for fatty liver and advanced fibrosis? A prospective study with 2080 controlled attenuation parameter and liver stiffness measurements.**

Kwok R, Wong GLH, Chan HLY, Chan JC, Kong AP, Wong VW.

Hepatology. 2014;60:611A-2A.

Annual Meeting of the American Association for the Study of Liver Disease 2014 (Boston, United States) - Poster

6. **A longitudinal study of fibroscan in chronic hepatitis B patients**

Kwok R, Lee AU, Ngu MC.

Hepatology International. 2013;7:S220.

Asia-Pacific Association for Study in the Liver 2013 conference (Singapore)-Poster

7. **A pilot study of Fibroscan®, caffeine breath test, APRI and AAR: Comparison of non-invasive markers of liver fibrosis in chronic hepatitis B patients**

Kwok R, Park GJH, Ngu MC, Lee AU.

Hepatology International. 2010;4(1):251-2

Asia-Pacific Association for Study in the Liver 2010 conference (Beijing, China) - Poster

8. **Diabetes mellitus is associated with high liver stiffness measurements by Fibroscan®**

Kwok R, Ngu MC, Kim A, Girgis C, Chen RC, Lee AU.

Hepatology International. 2010;4(1):251.

Asia-Pacific Association for Study in the Liver 2010 conference (Beijing, China) - Poster

9. **Transient elastography as a potential screening method for methotrexate induced hepatic fibrosis**

Cunningham A, **Kwok R**, Lee A, Shenstone B.

Internal Medicine Journal. 2010;40:10.

Australian Rheumatology Association/Rheumatology Health Professional Association 51st annual Scientific Meeting (Melbourne, Australia) – Oral Presentation

10. **Evaluation of hepatic fibrosis in chronic hepatitis B using transient elastography**

Kwok R, Gonzalez-Arce V, Kim A, Ngu MC, Lee AU.

Journal of Gastroenterology and Hepatology. 2009;24:A283.

Australian Gastroenterology Week 2009 (Sydney, Australia) - Poster

11. **Evaluation of liver fibrosis in diabetes using transient elastography.**

Kwok R, Girgis C, Chen R, Ngu MC, Lee AU.

Journal of Gastroenterology and Hepatology. 2009;24:A283.

Australian Gastroenterology Week 2009 (Sydney, Australia) - Poster

12. **Transient elastography (Fibroscan) in chronic hepatitis B infection: Reduce number of biopsies and save money**.

Guirgis M, Manoharan S, Scott DR, **Kwok R**, Lee AU, Connor SJ, Levy M.

Hepatology. 2009;50:543A.

Annual Meeting of the American Association for the Study of Liver Disease 2009 (Boston, United States) - Poster

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AAR	aspartate aminotransferase/alanine aminotransferase ratio
ALB	Albumin
ALP	alkaline phosphatase
ALT	alanine aminotransferase
API	age platelet index
APRI	AST platelet index
ARFI	acoustic radiation force impulse
AST	aspartate aminotransferase
AUROC	area under the receiver operator characteristics
BMI	body mass index
BR	Bilirubin
CAP	controlled attenuation parameter
CBT	caffeine breath test
cccDNA	covalently closed circular DNA
CCL2	chemokine ligand-2
CHB	chronic hepatitis B
CHC	chronic hepatitis C
CI	confidence interval
CK-18	cytokeratin-18
CRP	c-reactive protein
CT	computed tomography
DM	diabetes mellitus
eGFR	estimated glomerular filtration rate
ELF	European liver fibrosis score
ETOH	Ethanol
F stage	fibrosis stage
FI	fibrosis index
FIB4-I	Fibrosis 4 index
GGT	gamma-glutamyl transferase
HAIR	hypertension, increased ALT and insulin resistance
HbA1c	glycated haemoglobin
HBeAb	hepatitis B e antibody
HBeAg	hepatitis B e antigen
HBsAb	hepatitis B surface antibody
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HCC	hepatocellular carcinoma
HDL	high density lipoprotein
HIV	human immunodeficiency virus
HOMA-IR	homeostasis model assessment of insulin resistance
HR	hazards ratio
HSROC	Hierarchical summary receiver-operating characteristic
IBD	inflammatory bowel disease

IHTC	intrahepatic triglyceride content
IL	interleukin
INR	international normalised ratio
IQR	inter-quartile range
IQR/M	inter-quartile range/median
kPa	kilopascals
LDL	low density lipoprotein
LR	likelihood ratio
LRE	liver related events
LSM	liver stiffness measurement
M probe	medium probe
MCP	monocyte chemoattractant protein
MRE	magnetic resonance elastography
MRI	magnetic resonance imaging
MRS	magnetic resonance spectroscopy
NAFLD	non-alcoholic fatty liver disease
NA	nucleotide/nucleoside analogue
NASH	non-alcoholic steatohepatitis
NPV	negative predictive value
NS	non-significant
OR	odds ratio
PLT	platelet count
PNALT	persistently normal alanine aminotransferase
PPV	positive predictive value
QUADAS	quality assessment of diagnostic accuracy studies
SCD	skin capsule distance
SD	standard deviation
SMR	standardised mortality ratio
SPEA	serum prolidase enzyme activity
sRAGE	soluble receptor for advanced glycation endproducts
SS	simple steatosis
SVR	sustained viral response
Tc99-MIBI	Technetium-99 m-2-methoxy-isobutyl-isonitrile
TE	transient elastography
TNF	tumour necrosis factor
US	ultrasound
XL probe	extra-large probe
αFP	alpha fetoprotein

This page has been intentionally left blank

CHAPTER 1: GENERAL INTRODUCTION

1.1 The global burden of chronic liver disease

Chronic liver disease is a common disorder that causes significant health and economic burdens. Patients are susceptible to a variety of complications and their life expectancy can be markedly reduced. Irrespective of the aetiology, liver damage is insidious, with most patients experiencing no symptoms. Approximately 40% of patients with cirrhosis are asymptomatic and are identified only through incidental abnormalities on laboratory or radiographic studies (6). Often, the first symptoms experienced are that of decompensated cirrhosis, whereby it is generally too late for the condition to be reversed and life expectancy is limited.

Chronic liver disease is one of the major causes of death worldwide. The WHO Global Burden of Disease Study estimated there were 1.75 million deaths - 752 000 from liver cancer and 1.03 million from cirrhosis - in 2010 (7). Liver cancer and cirrhosis each ranked within the top 10 leading causes of death in upper and lower middle income countries in 2012 ***(Figure 1 and Figure 2)*** (8).

The top 3 aetiologies of chronic liver disease that cause liver related mortality and morbidity are chronic hepatitis B (CHB), chronic hepatitis C (CHC) and alcohol abuse. On a global scale, CHB is estimated to be responsible for 45% of liver cancer deaths and 30% of cirrhosis related deaths. Chronic hepatitis C causes 26% and 28% of liver cancer and cirrhosis related deaths, while alcohol abuse accounts for 25% of liver cancer and cirrhosis deaths. Variations across different regions are seen. CHC was the predominant cause of liver cancer and cirrhosis related deaths in the USA (40 and 41% respectively) and Western Europe (36 and 40% respectively). Chronic hepatitis B is the predominant cause of liver cancer and cirrhosis related deaths in China (54 and 46% respectively) and India (48 and 35% respectively). In Australia, alcohol is the leading cause of cirrhosis related death (33%), but CHB is the leading cause of liver cancer deaths (41%)(7).

Non-alcoholic fatty liver disease (NAFLD) is imminently the next major cause of liver disease burden. Increasing worldwide obesity rates (9) has led to doubling of NAFLD over the past 20 years, with prevalence ranging from 6.3 to 33% with a median of 20% (10). This increase has been observed not just in Western countries, but in the Middle East, Far East, Africa, the Caribbean and Latin America. The World Gastroenterology Organisation (WGO) has identified NAFLD as a worldwide pandemic (11).

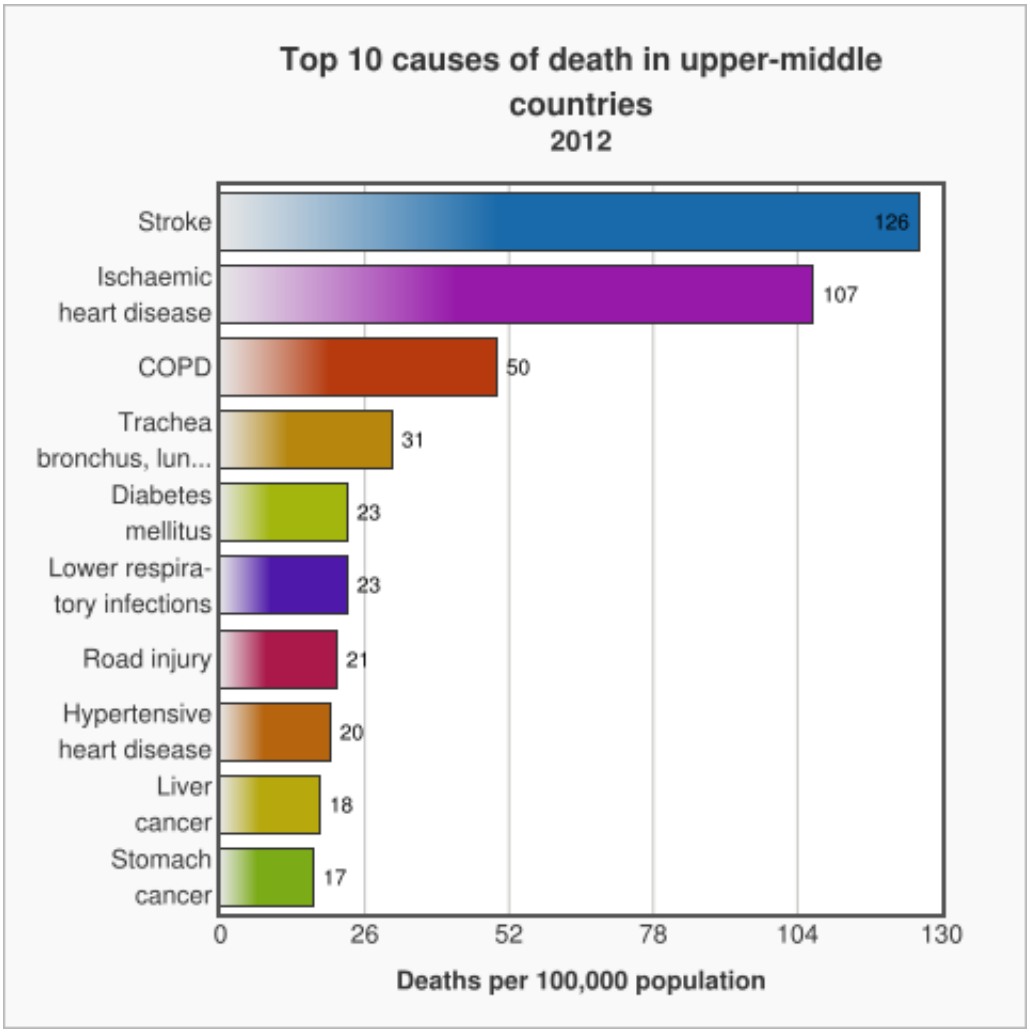


Figure 1: Top 10 causes of death in upper-middle income countries in 2012

Figure 1 taken from "WHO fact sheet #310"
http://www.who.int/mediacentre/factsheets/fs310/en/index1.html (8)

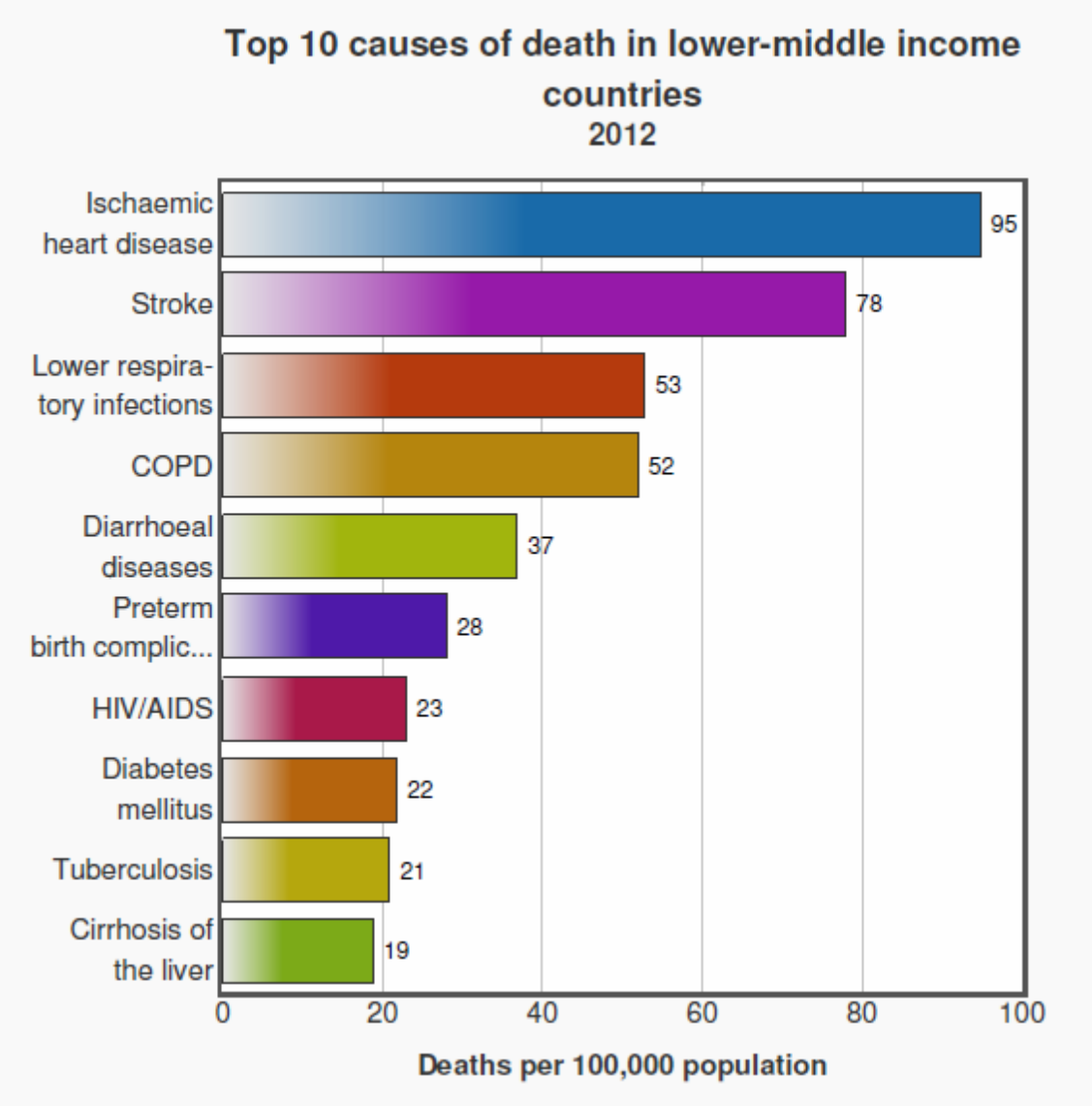


Figure 2: Top 10 causes of death in lower-middle income countries in 2012

Figure 2 taken from "WHO fact sheet #310"
http://www.who.int/mediacentre/factsheets/fs310/en/index1.html (8)

1.2 Liver related mortality and economic burden in Australia

According to the Australian Bureau of Statistics data, 7266 liver related deaths occurred in 2012, ranking it as the 19th leading cause. Other diseases such as cardiovascular, respiratory and cancer contribute a much greater proportion of mortality, but the median age for death is much younger in liver disease. For instance, in alcoholic cirrhosis, the median age at death was 58.3 years and 56.3 years for males and females respectively. The overall impact to society is under-represented than what mortality statistics suggests because there are greater impacts due to premature loss of life and associated social and economic costs (12).

To demonstrate the total economic burden of liver disease, the Gastroenterological Society of Australia (GESA) and Australian Liver Association (ALA) commissioned a Deloitte Access Economics report (13). Several findings were of note. Firstly, a staggering 6,179,285 persons had liver disease (approximately 28% of the population). NAFLD accounted for an estimated 90% ***(Table 1)***. Secondly, the total financial impact (including health system costs, loss of productivity due to inability to be employed, absenteeism, premature death, care giver costs, program and welfare payments, funeral costs and taxation foregone) was a massive $5.4 billion. Burden of disease quantifies the impact in terms of disability adjusted life years, and is a better reflection of the greater loss in function seen in liver disease patients who are generally younger and within working age. This was calculated to an estimated $50.7 billion ***(Table 2)***.

The total economic burden of chronic liver disease is 40% greater than that from diabetes and chronic kidney disease combined, and it is two fifths of the cost of cardiovascular disease, which has the greatest economic burden. Alarmingly, the prevalence of liver disease is projected to affect 8,092,339 persons by 2030 ***(Table 1)*** (13).

Table 1: Prevalence of liver disease Australia 2012

Disease	2012 Males	2012 Females	2012 Persons	2030 Males	2030 Females	2030 Persons
Hepatitis A	148	135	284	180	164	344
Hepatitis B	105,555	105,535	211,089	131,782	131,061	262,842
Hepatitis C	185,468	121,572	307,040	251,391	162,887	414,278
NAFLD	2,713,372	2,825,305	5,538,677	3,566,969	3,693,619	7,260,588
Primary liver cancer	1,064	387	1,451	1,652	601	2,253
Alcoholic liver disease	4,605	1,598	6,203	5,816	2,008	7,824
PBC	43	389	433	63	553	616
PSC	554	318	872	748	425	1,174
Haemochromatosis	56,343	56,894	113,237	70,999	71,421	142,421
Total	**3,067,152**	**3,112,133**	**6,179,285**	**4,029,600**	**4,062,739**	**8,092,339**

Prevalence of liver disease in Australia by gender in 2012 and 2030 – taken from "The economic cost and health burden of liver disease in Australia" (13)

Table 2: Cost of Liver Disease Australia 2012

	Federal government	States and territories	Individuals	Other parties	Total
Burden of Disease			45,256.0		45,256.0
Health system costs					
Health expenditure	164.9	102.0	70.7	48.7	386.2
National immunisation program	14.5	9.0	6.2	4.3	34.0
Research funding	5.0	3.1	2.1	1.5	11.7
Productivity costs					
Employment	680.3		1,386.9		2,067.2
Absenteeism	68.2			139.0	207.1
Premature death	630.1		1,284.5		1,914.6
Carer costs	85.1		173.5		258.7
Program payments					
National respite for carers	0.3				0.3
Palliative care	1.6				1.6
Funeral costs			33.9		33.9
Welfare payments	53.7			-53.7	-
Transfer DWLs				526.8	526.8
Total	**1,703.7**	**114.1**	**48,213.8**	**666.6**	**50,698.1**

Total costs of liver disease, by type and bearer in 2012 ($ million) – taken from "The economic cost and health burden of liver disease in Australia" (13)

1.3 Fibrosis in chronic liver injury

The response to liver injury is a ubiquitous process of inflammation and fibrosis regardless of the cause. Fibrosis develops in all patients with chronic liver injury at variable rates depending upon the aetiology and host factors (14). There is collapse of hepatic lobules, formation of fibrous septae, and hepatocyte regeneration with nodular formation. This diffuse process may ultimately progress to cirrhosis with its accompanying consequences of portal hypertension and impaired hepatic function. Cirrhosis represents a late stage of progressive hepatic fibrosis characterized by distortion of the hepatic architecture and the formation of regenerative nodules.

Once thought of as irreversible, hepatic fibrosis is now recognized as a dynamic process with the potential for significant resolution. However, late stage cirrhosis with profuse fibrous nodules, severe portal hypertension and grossly impaired synthetic function is generally thought of as irreversible. The exact point of no return is not well defined. Evidence suggests that early cirrhosis is reversible when the offending agent is removed. In chronic hepatitis B, patients experience a regression of cirrhosis in 74% of cases after 5 years of viral suppression with tenofovir disoproxil fumerate (15). Chronic hepatitis C patients who achieve a sustained virological response (SVR) after antiviral therapy undergo cirrhosis regression in 61% after 61 months (16). NAFLD patients observed over 3 years had fibrosis regression in 25%, associated with corresponding reductions in waist circumference and low density lipoprotein cholesterol (17).

Regression of hepatic fibrosis is possible with specific disease targeted treatment. Further, antifibrotic therapy may be available as we gain new molecular insights into fibrogenesis. Hence accurate evaluation of liver fibrosis remains an important clinical tool.

1.4 Cellular and molecular mechanisms of fibrosis progression and regression

Activated hepatic stellate cell (HSC) and their transformation into myofibroblasts is the main driver of liver fibrogenesis (18). The transition of HSCs into myofibroblasts is regulated by their interaction with several other cell types and the activation of specific pathways (19). Injured hepatocytes, hepatic macrophages (Kupffer cells), endothelial cells, and lymphocytes drive HSC activation. The death of hepatocytes leads to the release of cellular contents and reactive oxygen species that activate Kupffer cells to release pro-inflammatory factors such as TNFα, IL-1b, IL-6, CCL2, TLR4 and pro-fibrogenic factors TGFβ (20, 21).

Kupffer cells drive fibrosis progression in chronic injury, but also coordinate the regenerative response in acute injury. They stimulate the influx of bone marrow derived immune cells via release of CCL2 and CCL5 recruitment of immature monocyte-derived Ly6Chi macrophages. In mouse models, the absence of Ly6Chi macrophages results in inhibition of the pro-fibrogenic response in carbon tetrachloride injury suggesting that Ly6Chi macrophages are central to the fibrosis mechanism and activation (22). Ly6Chi macrophages can differentiate into pro-resolution (restorative) Ly6Clo macrophages which secrete large quantities of fibrolytic matrix metalloproteinases such as MMP-9 and MMP-13, and the anti-inflammatory cytokine IL-10, which are all implicated in fibrosis resolution (23). This may involve the fractalkine receptor CX3CR1 (24), but further research is needed. These pathways may be potential targets for antifibrotic therapy and remain of intense interest (20).

Oxidative stress may also be an important aspect of fibrogenesis. Reactive oxygen species causing oxidative stress in chronic tissue damage can lead to the overexpression of critical genes related to extracellular matrix remodeling, inflammation and fibrogenesis, especially in alcoholic hepatitis and non-alcoholic steatohepatitis (NASH) (25, 26). Other factors that have been implicated include the intestinal microbiota (27), tissue hypoxia (28), epigenetic modification in conditioning the progression of fibrosis (29) and the mechanical properties of the underlying matrix on the progression of the fibrogenic process (30).

1.5 Liver biopsy and the histological staging of liver fibrosis

Liver biopsy is the reference standard for the assessment of hepatic fibrosis. There are several histological scoring systems used to stage liver fibrosis. The most common include the Metavir, Knodell, Ishak, Scheuer, and Laennac scores (31-35). All scoring systems provide a semi-quantitative score based on the distribution and extent of fibrosis. Each was developed originally for chronic viral hepatitis, and also includes an assessment of necroinflammatory activity. In general they describe increasing severity of each stage based on the location of fibrosis within the hepatic lobule. Fibrosis begins around the portal triad, then extend to septations which later connect and form bridges between the triads. Finally, the fibrous septations becomes so numerous and prominent that discrete nodules are formed which is the basis of histological cirrhosis. In studies of non-invasive assessment of liver fibrosis, Metavir score is the most commonly used system, and hence adopted in our research. The description and typical appearance of each Metavir stage is provided in ***Figure 3.***

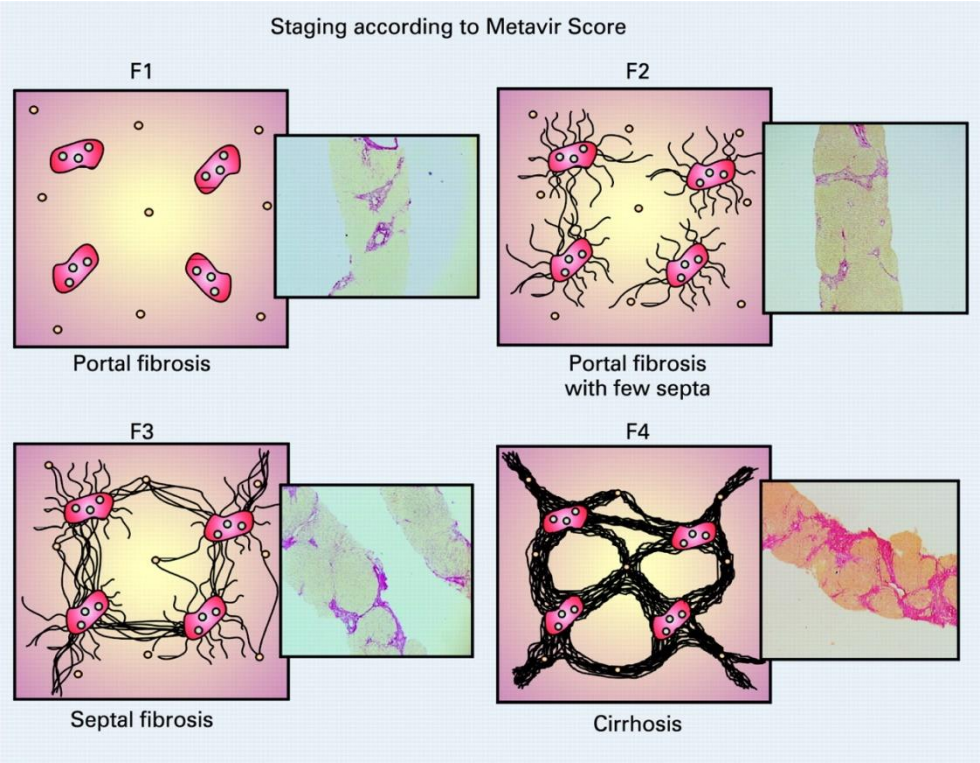


Figure 3: Metavir Fibrosis Score

Adopted from Asselah 2008 (36)

Although liver biopsy is considered the reference standard for fibrosis assessment, there are several limitations. The decision to proceed with a liver biopsy between the clinician and the patient is usually not a trivial one. It is a painful invasive procedure that is associated with uncommon, but serious complications (see 1.5.1). Patients are reluctant to submit to repeated biopsies which limit the ability to monitor disease progression and treatment response. Interpretation requires expertise, and remains subject to variability as well as sampling error. As such, liver biopsies cannot be used on a large scale.

1.5.1 Complications of liver biopsy

Complications from liver biopsies are overall uncommon, with large retrospective studies reporting a risk of bleeding of approximately 0.3-1.7% and a risk of death of 0.09 -0.33% (37-39).

In a British nationwide audit of 1500 cases across 189 health institutions (37), bleeding complicated 26 procedures (1.7%), and transfusion was required in 11 (0.73%). There were two definite and three possible procedure related deaths, giving an overall mortality of 0.13-0.33% (37).

In a retrospective study of 9212 liver biopsies over 21 years, there were 10 fatal and 22 non-fatal haemorrhages (0.11% and 0.24%, respectively) (38). Risk factors were malignancy, age, sex, and the number of biopsy needle passes. The rate of fatal and non-fatal haemorrhage were 0.4% and 0.57% compared with 0.04% and 0.16% respectively (38).

The largest review comprised of 68,276 biopsies over 10 years across 36 Italian institutions (39) . A total of 147 (2.2%) complications were found. Complications related to bleeding occurred in 30 (0.44%), inadvertent puncture into the chest cavity in 51 (0.75%), puncture of other viscous in 13 (0.19%) and others (such as sepsis, shock, biliary peritonitis, reaction to anaesthetic) in the remaining 53 (0.78%). Fatal complications occurred in 6 (0.09%), all due to bleeding and in which 3 patients had cirrhosis (39).

Pain is a very common side effect. A study showed that pain is experienced in 87% of cases and extend beyond the day of the procedure in 20% (40). Observation for at least 6 hours post procedure is required by most protocols to monitor for complications and adds inconvenience to patients. For all the aforementioned reasons, liver biopsy has a poor patient acceptance and tolerance.

1.5.2 Sampling error

A biopsy represents approximately 1/50000th volume of the liver, so that sampling error can occur where disease distribution is uneven or when the size of the biopsy inadequate. In a study of 124 hepatitis C patients, 33.1% had a difference of at least 1 stage of fibrosis in biopsy samples taken laporascopically simultaneously from the left and right lobe, leading to an under-diagnosis of cirrhosis in 14.5% (41).

Length of biopsy is important, with inadequate samples being associated with greater variability. One study of 17 liver samples demonstrated 65% of biopsies 15 mm in length were categorized correctly

according to the reference value. However this increased to 75% for a 25-mm liver biopsy specimen (42).

In a study of 161 liver biopsies, assessors were blinded and specimens were repeatedly examined after shortening the visible length from ≥3cm to 1.5cm and then to 1cm. Shorter length was associated with reporting of mild fibrosis. They also reported that 11-15 portal tracts were required for accurate staging. The authors concluded that to achieve 11-15 portal tracts, a biopsy at least 2cm in length would be adequate in 94% of specimens (43).

This standard of a 2cm biopsy length however, may be unrealistic. A meta-analysis of 32 studies with 10027 liver biopsies showed that the mean length was 17.7mm and mean number of portal tracts was 7.5. Only 8 studies had a mean length of at least 20mm (44) . These biopsies were performed in research studies, where strict study criteria for minimum biopsy length were required. Having adequate length liver biopsy (>2cm) is unlikely to be achievable in the real world, and would be difficult to enforce because of concern with the additional risks of bleeding and other complications.

1.5.3 Variability of histology interpretation

Interobserver variability is a well described issue in histological fibrosis staging. It is generally reproducible amongst pathologists in academic centres or who are specialised in hepatic histopathology. A high level of concordance in 30 liver biopsies specimens examined by 10 specialist liver pathologist in the METAVIR study group was found for portal fibrosis (K=0.80) and cirrhosis (K=0.91)(32). Another study of 95 liver biopsies showed an 84% agreement across 3 observers in an academic centre (45).

Outside of academic settings, the variability of interpretation is significantly higher. In 391 biopsies, there was complete agreement between specialist liver pathologists and community pathologists in only 49.9% of cases. Across all stages of fibrosis, the correlation (K) coefficient was 0.41 which is considered only to be a poor to fair level of agreement, with 73% cases understaged by community pathologists (46).

1.5.4 Barriers for use in large numbers of patients

Most protocols for post liver biopsy care requires monitoring for several hours after the procedure to ensure there has been no serious side effects. This usually requires a full day commitment from the patient. For liver biopsy to be performed safely, skilled operators, and a facility for monitoring post procedure are necessary. Use of imaging to safely guide the location of biopsy is now common practice and necessitates ultrasound imaging equipment. Furthermore, for liver histology to be interpreted with a high level of accuracy, it is best examined by specialist liver pathologists. Liver biopsy and assessment is thus is a resource intense procedure. It is impractical to be used as routine tool on a large numbers of patients.

Given the limitations of liver biopsy and combined with the increasing burden of liver disease, the need for alternative methods of assessment is accentuated.

1.5.5 Transient Elastography as an alternative to liver biopsy

There has been much research in non-invasive measures of hepatic fibrosis. At the time of this thesis' conception in 2009, an emerging method for liver fibrosis evaluation was Transient Elastography (TE).

Transient elastography assesses the elasticity of tissue in the liver. Fibroscan ® is a non-invasive device developed by Echosens ® (Paris, France) that applies TE to measure liver fibrosis. Fibroscan overcomes the drawbacks of liver biopsy with regards to complications, pain, poor patient acceptance, and high resource demand. It also assesses a much greater volume of liver compared to liver biopsy.

Fibroscan is rapidly performed with each case requiring around 10 or fewer minutes to complete. Results are displayed immediately after each scan. The patient feels a soft tap on the skin surface from the probe. There is no pain or complications. There is negligible operator dependence after a short period of training and supervision. The volume of liver assessed by Fibroscan is approximately $4cm^3$. This is 100 to 200 times greater than that of a liver biopsy, and potentially reduces sample variation. Fibroscan has been widely accepted by patients and physicians because it is safe, painless, rapid, and produces results that are instantly available.

1.6 Transient Elastography: basic principle and the concept of the Liver Stiffness Measurement

1.6.1 Concept of Transient Elastography

Transient Elastography characterizes the elasticity of soft tissues. Using this method, an imaging system follows in real time, the propagation of a low frequency shear wave. The displacement of the propagating shear wave is measured as a function of time and space. The aim of tissue elastography is to create high-resolution shear stiffness images of human tissue for diagnostic purposes (47).

The shear stiffness of any material refers to its propensity to deform under mechanical stress. It is known as the Young's Modulus or Elastic Modulus and is measured in units of Pascals (Pa). The concept is commonly applied in engineering and was first described by 18th century scientists Leonhard Euler and Giordano Riccati, and was further developed by Thomas Young in the 19th century (48).

The stiffness of human tissue has long been utilized in medicine as a method of distinguishing different pathological states. For instance, the 'hardness' of a lump is traditionally taught in clinical examination to medical students to help determine the likelihood of a lesion being benign or malignant. In the case of liver fibrosis, the progressive deposition of collagen leads to a "stiffer" liver. The goal of transient elastography is to utilise this property to generate high-resolution images. The expectation is that shear stiffness images will identify abnormal tissue not identified by standard ultrasound techniques.

1.6.2 What is Fibroscan® and Liver Stiffness Measurement?

Fibroscan® is a non-invasive medical device developed by Echosens® (Paris, France) that applies the principles of transient elastography to diagnose the magnitude of fibrosis in the liver. It was first introduced in 2003 in Europe when it received European Medical Association approval. In Australia, Fibroscan® was registered with the Therapeutic Goods Administration on 28th April 2008.

Fibroscan consists of a specialised ultrasound probe and an integrated computer ***(Figure 4 and Figure 5).*** The specialised probe has an ultrasound transducer fitted on the axis of an electrodynamic transducer. The electrodynamic transducer creates a low-amplitude mechanical pulse and generates a low frequency (50 Hz) elastic wave, also known as a shear wave, which propagates throughout the liver tissue.

The ultrasound transducer transmits and senses radiofrequency signals during the shear wave. Comparison of consecutive signals allows for the mapping of the local strain of the medium. A strain rate image is then generated by the integrated computer. The strain rate image reflects rates of deformation generated in the liver by the propagation of the elastic wave as a function of time (horizontal axis in milliseconds) and of depth (vertical axis in millimeters) ***(Figure 6)***. The colour scale indicates the sign of the deformations (compression or dilatation) and their amplitude. The speed of the propagation of this elastic wave is proportional to the slope. The greater the slope, the greater the propagation speed,

known formally as the shear wave velocity. The square of the shear velocity (V_s) is proportional to the elastic modulus (E): $\mathbf{E \propto V_s^2}$**.** Consequently the elastic modulus of the liver can then be derived.

The elastic modulus of the liver has been coined as the Liver Stiffness Measurement (LSM) and thus from here on will be referred to as such.

Figure 4: The Fibroscan probe

Figure obtained and used with permission from UITC, distributor of Fibroscan.

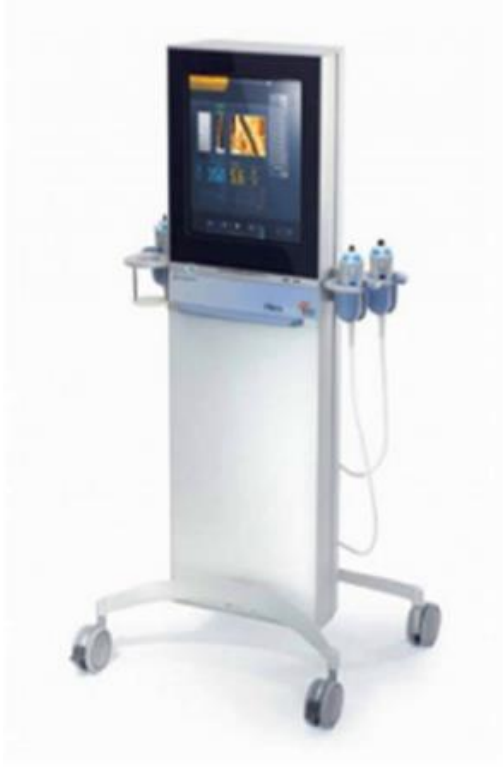


Figure 5: The Fibroscan Chassis

Figure obtained and used with permission from UITC, distributor of Fibroscan.

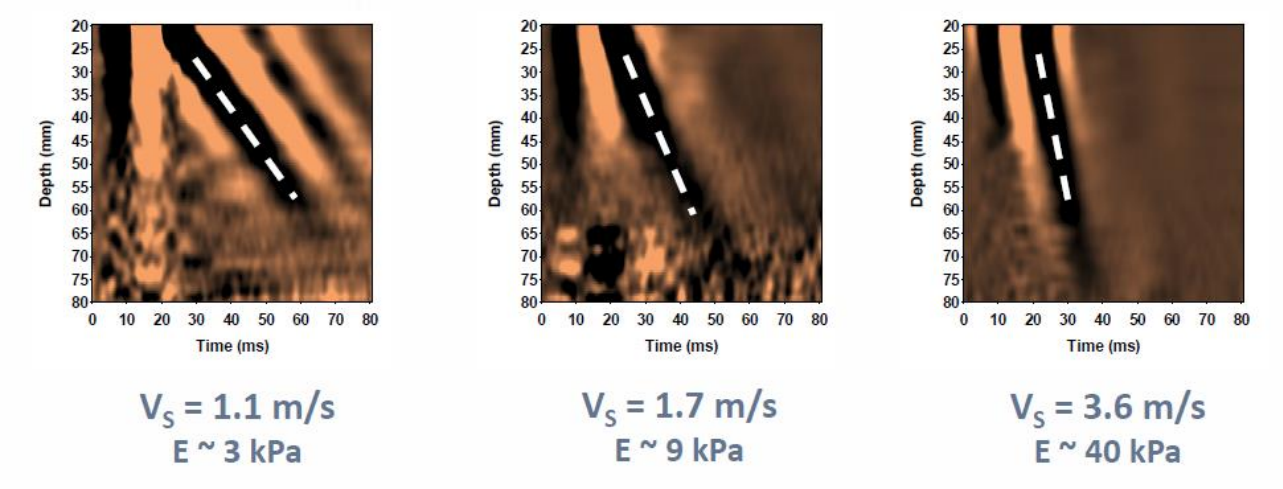


Figure 6: Strain Rate Images

The vertical axis represents the depth of measurement (millimetres), while the horizontal axis represents time (milliseconds). The deformation caused by the shear wave is the prominent diagonal dark stripe. The slope of this stripe, as indicated by the dotted white line, represents the shear wave velocity. Below each strain rate image is the corresponding shear wave velocity (V_s) in metres per second, and the derived elastic modulus (E) in kilopascals. Figure obtained and used with permission from UITC, distributor of Fibroscan.

1.7 Performing and Interpreting a Fibroscan

1.7.1 Performing a Fibroscan

Patient preparation and positioning

Portal blood flow increases after a meal and this has been found to increase the LSM (49-51). 3 hours fast prior to performing the Fibroscan is recommended.

The measurement of the stiffness of the liver is carried out on the right lobe of the liver at the intercostal spaces in the mid-axillary line. The patient is instructed to be lying down on the examination table in supine position with the right arm in maximum abduction.

Taking measurements

Ultrasound gel is applied on the tip of the probe sensor and placed in contact with the skin of the patient. As soon as the probe is in contact with the patient's skin, a pressure variation is detected which

initiates the ultrasonic sensor functions in ultrasound mode. The ultrasound signal is represented by the display in mode A and the display as function of time in M Mode. These two modes allow verification that the measurement zone includes the liver only. The manufacturer recommends choosing a liver zone at least 7cm thick, away from the liver edges and large vascular structures. The measurement can only be triggered within a specific range of pressure applied as indicated by a green zone. Once the measurement has been chosen, and the probe is kept perpendicular to the patient's skin with an appropriate pressure, the measurement of the stiffness can be triggered by pressing on one of the probe buttons. The vibrator generates a low frequency (50Hz) elastic wave, and the acquisition lasts less than a tenth of a second. A progression bar appears in the information window while the acquisition data are transferred and processed. The results are then displayed. The images generated for the signals in A mode, M mode and the strain rate image are known as the elastogram ***(Figure 7).***

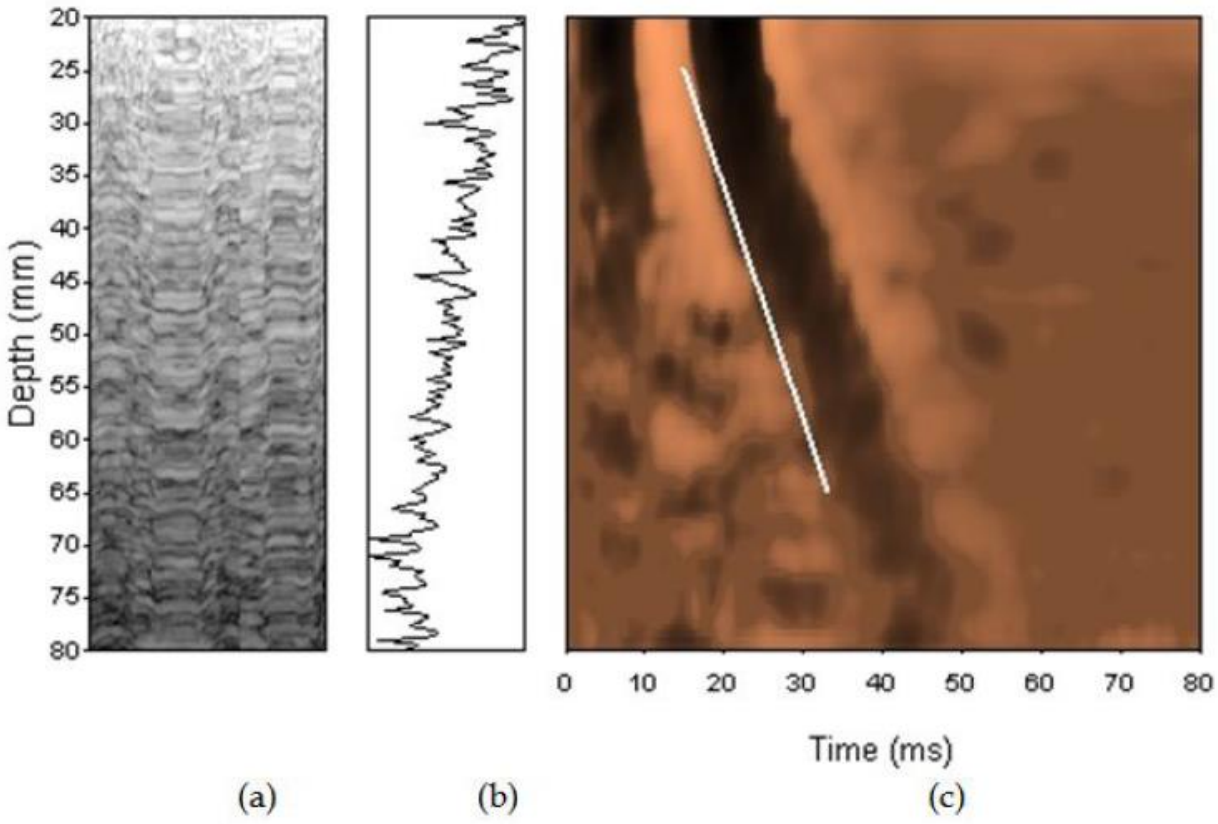


Figure 7: The Elastogram

(A) Represents the M mode. (B) Represents the A mode and (C) is the strain rate image. Figure obtained and used with permission from UITC, distributor of Fibroscan.

1.7.2 Correct evaluation of liver stiffness by interpreting elastograms

A correct evaluation of the liver requires the elastogram to meet three quality criteria.

The A mode signal needs to be linear. This means that the region of measurement is uniform and suggests there is only one type of tissue being evaluated.

The M mode image needs to show the characteristic layer structure of the liver at all the depths. The image pattern should be homogenous, suggesting that only liver parenchyma is being measured. Heterogeneous patterns that feature white bands may represent the liver border with the lung and vascular structures. Patterns with dark bands represent ribs.

The third criterion is that the elastogram is free of defects. Fibroscan's internal software automatically rejects and cancels elastograms when the signal detected and strain rate image produced are not satisfactory. However not all defective elastograms are correctly filtered by the software and liver stiffness results for these inaccurate elastograms are still calculated and displayed. These defects need to be correctly recognized by the operator and the corresponding LSM results should be disregarded. The known defects that can occur are "A" waves, "E" waves and angled waves (52).

"A" waves are dual shear waves. They form due to shear waves emerging from the same point and then diverge, which then creates the appearance of the capital letter "A". The steeper shear wave being detected by the software instead of the shallower wave causes an over-estimation ***(Figure 8)***.

"E" waves are enlarged waves. These are similar to "A" waves, where dual shear waves emerge from the same point and diverge, but are so close together that they coalesce forming the shape of a wedge. If the steeper slope is detected by the software as being the liver stiffness, this results in an over-estimation ***(Figure 9)***.

A waves and E waves are due to shear waves arising from ribs. Since the hardness of bone is much greater than that of the liver, this leads to an overestimation. "A" waves and "E" waves can be avoided by finding a larger intercostal space.

Angled waves describe the shear wave seen as having a point of inflection at the proximal depths of measurement, causing an angled appearance ***(Figure 10)***. The corresponding A mode will show a non-linear signal, and TM mode non-homogenous saturation. Angled waves occur when the distance between the liver capsule and skin surface, called the skin capsule distance (SCD) is within the probe's region of measurement. Structures between the liver surface and the skin, such as the fibrous liver capsule may be included in the interpretation of liver stiffness, and can cause an over estimation. Angled waves can be avoided by finding a measurement spot where the liver is closer to the skin surface.

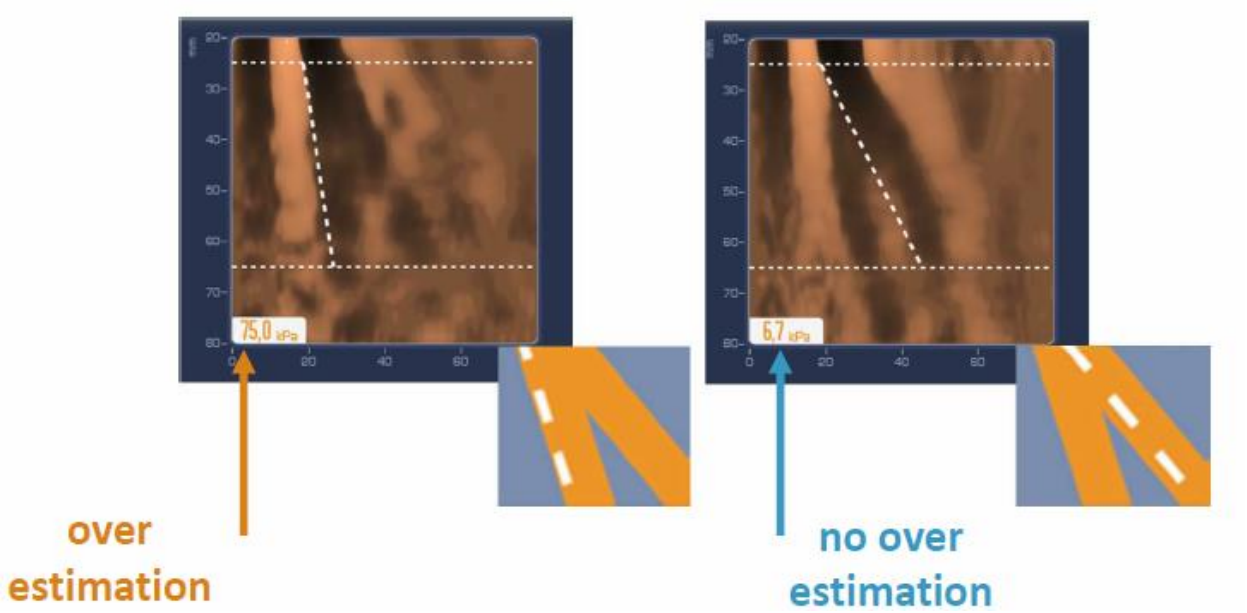


Figure 8: "A" wave

The left image shows the steeper shear wave being interpreted by Fibroscan software as the shear velocity (as indicated by the dotted white line), which leads to an overestimation of the liver stiffness. The right image shows the shallower shear wave being interpreted as the shear velocity, which is the correct estimation of liver stiffness. Figure obtained and used with permission from UITC, distributor of Fibroscan.

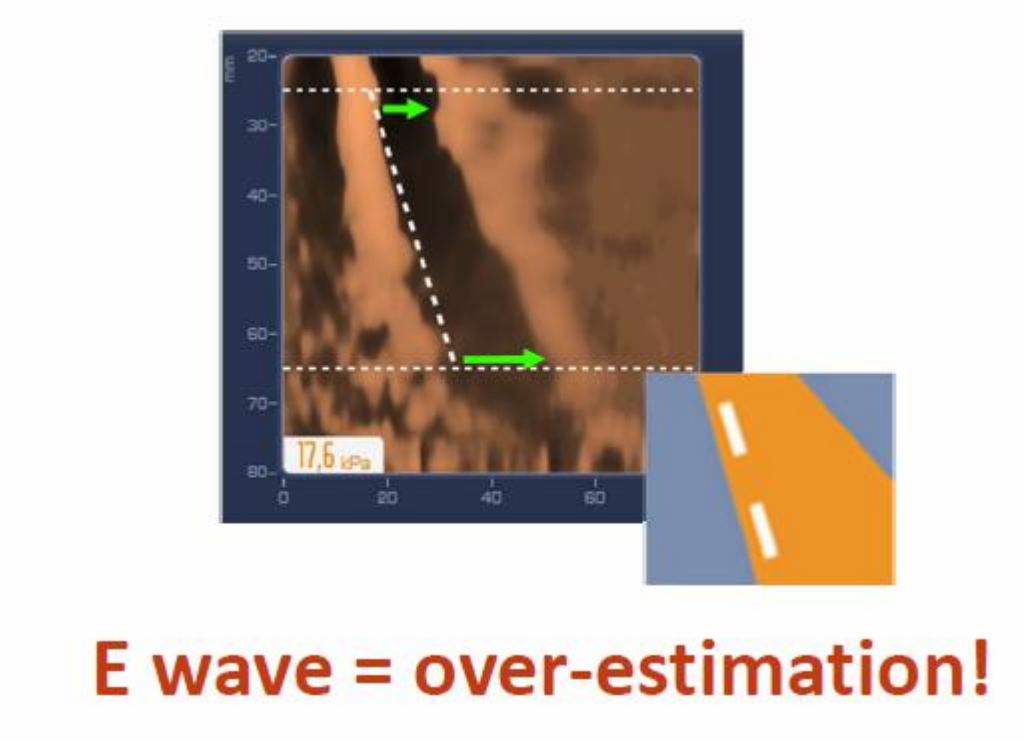


Figure 9: "E" wave

The shear wave becomes gradually larger at deeper levels from the coalescence of 2 shear waves. The slope of the steeper shear wave is interpreted as the shear wave velocity leading to an overestimation. Figure obtained and used with permission from UITC, distributor of Fibroscan.

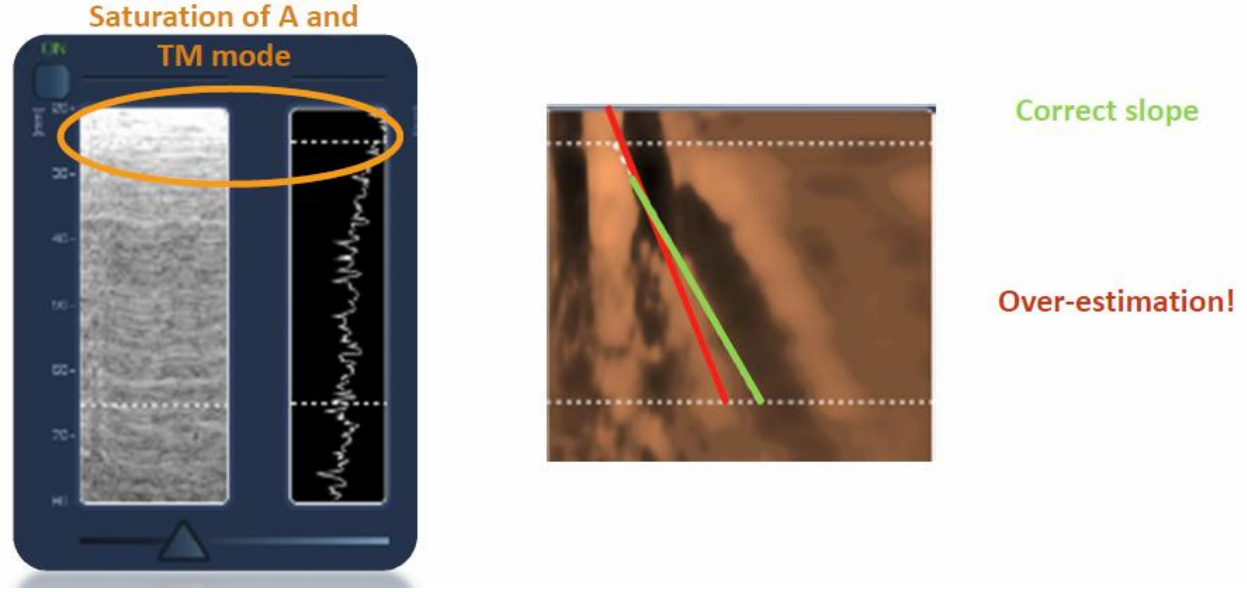


Figure 10: Angled wave elastograms

The shear wave appears to have an inflection point at approximately 35mm depth. This leads to an overestimation if the slope of the shear wave prior to the inflection point (represented in red) in measured, as opposed to the slope of the shear wave beyond the inflection point (represented in green). At depths between 20-30mm, the corresponding A mode and M mode shows a non-linear signal and a much brighter saturation respectively. Figure obtained and used with permission from UITC, distributor of Fibroscan.

1.7.3 Calculation of the Liver Stiffness Measurement

The final Liver Stiffness Measurement (in kPa) is represented by the median value of all the valid stiffness measurements. The interquartile range (IQR, in kPa) represents the interval around the median (m) containing the measurements that are between the 25^{th} to 75^{th} percentiles. If the measurement is not valid, the current stiffness is not defined. The software calculates a "success rate" in % to express the number of valid measurements in relation to the number of attempted measurements and shows this information in the display.

1.7.4 Requirements for a reliable stiffness examination: changing definitions

When this study began in 2009, a “reliable” stiffness examination was based on 3 criteria (52):

1. At least 10 valid measurements performed at the same spot in the right lobe of the liver

2. A Success Rate (SR) ≥60%

3. An IQR/Median Ratio of liver stiffness ≤ 0.30

Excellent interobserver agreement is observed when the IQR to median ratio (IQR/M) is ≤ 30%. A study of 800 TE examinations performed on 200 patients by 4 operators revealed the intraclass correlation coefficient to be excellent (K= 0.98) (53). From this study, the term “reliable” LSM became synonymous with reproducibility.

A lower IQR/M ratio may even reduce discordance further. In a study of 254 patients, the most discriminant cutoff value for discordance was IQR/M = 0.21. When IQR/M ≤ 0.21, discordance was observed in in 10/135 cases (7.4%) as compared to 18/119 (15.1%), p≤0.05 when IQR/M > 0.21 (54).

Subsequently, a study of 1165 patients in 2013 challenged the usual definitions for LSM reliability which led to the criteria being refined (55). In this study, TE findings were compared to liver biopsy. Using the usual definition of a “reliable” LSM, cirrhosis was not diagnosed any more accurately than in those with an “unreliable” LSM. When the LSM was < 7.1kPa, there was no difference in the accuracy of fibrosis stage classification regardless of the IQR/M ratio. When the LSM was 7.1-12.5 kPa, the accuracy was significantly superior if the IQR/M ≤ 0.30. When the LSM was ≥12.5 kPa, accuracy was superior if IQR/M< 0.10. Thus the investigators proposed new system of definitions. Instead of “reliable” vs “unreliable”, the new classification included ‘‘very reliable’’ (IQR/M <0.10); ‘‘reliable’’ (IQR/M 0.10-0.30, or IQR/M >0.30 with LSM median <7.1 kPa), and ‘‘poorly reliable’’ (IQR/M >0.30 with LSE median ≥7.1 kPa). Using this new classification, 74.3% of scans were now considered “reliable” and 16.6% of scans were “very reliable”. Scans that were “unreliable” accounted for 9.1% of total cases. In contrast, the previous definition would classify 24.3% of scans as “unreliable” (55).

In addition, the LSM success rate in the study was not found to have any influence on diagnostic accuracy and the authors concluded that success rate is irrelevant. The same conclusion was made in another study of 251 patients which also found that success rate has no bearing on the accuracy of TE (56).

The following are revised definition of reliable LSM based on the following criteria (57):

1. A minimum of ten valid measurements
2. An IQR/median ratio of ≤ 30%. This is not required if final stiffness result is < 7.1kpa

These were adopted for all analyses in this thesis.

1.8 Histological scoring systems for the diagnosis of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis

A diagnosis of Non-alcoholic fatty liver disease (NAFLD) can be made simply based on the presence of >5% steatosis on liver biopsy, and after secondary causes (e.g. alcohol) have been clinically excluded. However, there is a wide spectrum of other histological features that may be present, including varying degrees of inflammation, liver injury, fibrosis and cirrhosis. Early studies suggest that Non-alcoholic steatohepatitis (NASH), is much more likely to progress to cirrhosis and hepatocellular carcinoma compared to bland steatosis (17, 58, 59). Since liver biopsy is the gold standard for the diagnosis of NASH, it is therefore important that histological interpretation is accurate and reproducible. Histology based scoring systems have been proposed make the diagnosis of NASH more objective and less prone to interobserver variance. The first large collaboration to address this was by the Pathology Committee of the NASH Clinical Research Network (NASH CRN), which designed and validated a histological scoring system to address the full spectrum of lesions of NAFLD, and to propose a NAFLD activity score (NAS) for use in clinical trials. The NAS comprised of a total score calculated from grade of steatosis (0 pts: < 5%, 1 pt: 5-33%, 2 pt: 33-66%, 3 pt: >66%), lobular inflammation (0 pts: 0, 1 pt: 1-2. 2 pt: 2-4, 3pt: > 4 foci/20x power field), and ballooning (0 pts: none, 1 pt: mild/few, 2 pts: moderate/many). The NAS score would thus range 0-8 points. A score of 1-2 would be considered no NASH, 3-4 borderline NASH and 5-8 definite NASH. Fibrosis staging was based on the Brunt's or Kleiner's system: F0 = no fibrosis; F1 = perisinusoidal or portal; F2 = perisinusoidal and portal/periportal; F3 = septal or bridging fibrosis; and F4 = cirrhosis (60, 61).

However, the NAS score was not an ideal a diagnostic tool. It was designed more to provide a continuous scale for activity assessment, which would be useful in clinical trials. A notable issue was that a significant proportion of patients that would be classified as borderline NASH under NAS, would be considered definite NASH or definite non-NASH by consensus expert opinion (62).

To address some of these issues, a new system known as the Steatosis Activity Fibrosis (SAF) score was developed by Bedossa et al. (63) and validated by the Fatty Liver Inhibition of Progression (FLIP) consortium. Like the NAS, the SAF comprises a score for each of the main components of NAFLD: steatosis, ballooning, lobular inflammation, and fibrosis. However, disease activity is determined by the sum of ballooning (0-2) and lobular inflammation (0-2) only. Steatosis is excluded as a factor in determining disease severity in the SAF score, of which its inclusion was among the chief criticisms of the NAS score. In addition, the SAF proposes an algorithm for determining NAFLD vs NASH. Essentially, the presence of any steatosis (grade 1,2 or 3) in association with both ballooning (grade 1 or 2) and lobular inflammation (grade 1 or 2) would be classified as NASH. NAFLD is the diagnosis instead of NASH when there is absence of ballooning or lobular inflammation (either being grade 0). The absence of steatosis (grade 0) would preclude a diagnosis NAFLD altogether.

Agreement between expert liver pathologists and general pathologists on the diagnosis of NASH was found to improve after applying the SAF score algorithm (64). Despite this, the NAS and SAF scores are both tools that are mostly used in research. Greater validation in the hands of non-expert pathologists

are required. Furthermore, while it is important to differentiate NASH from simple steatosis, recent studies have revealed that ultimately, the degree of fibrosis has the greatest prognostic influence (65).

1.9 Rationale for examining the clinical utility of transient elastography in nonalcoholic fatty liver disease, chronic hepatitis B and long term methotrexate patient populations

At the time of this study's conception, few publications existed for TE. Most research focused on hepatitis C, with scarce data for other patient populations. We identified nonalcoholic fatty liver disease, hepatitis B and methotrexate induced liver fibrosis as areas requiring further study.

A systematic web-based literature search of all publications in MEDLINE (via OvidSP), PUBMED (NLM) and EMBASE was conducted on 10th January 2009 from the date of inception for each of the databases. Our primary search strategy for identifying studies used the following free-text words: "transient elastography", "Fibroscan" and "liver stiffness measurement". Search limits included abstracts and publication in peer-reviewed journals. This list of articles was corroborated with the latest official summary of "Publications and Communications for Fibroscan" provided by Echosens (66). A secondary search of the reference lists of the articles that were identified on the primary search was performed to locate any studies missed by electronic search strategies.

A total of 88 publications were found. The vast majority were from Europe where Fibroscan originated. 23 articles were in a non-English language and 36 articles were identified to be reviews, comments, letters or editorials. The remaining 29 articles were original research studies. The main aetiology of liver disease studied in each publication is broken down as follows:

- 13 publications focusing mainly on the hepatitis C population. 4 of these studies examined HIV/HCV co-infection and one reported the use of TE in liver transplantation for HCV patients (67-79).
- 6 publications examining the use of TE in chronic liver disease of multiple aetiologies. However, in every study, the majority of the patients had hepatitis C (80-84).
- 3 studies focused on the reliability and reproducibility of TE (53, 85, 86).
- 2 studies that examined TE for use in chronic hepatitis B or predominantly chronic hepatitis B (87, 88)
- 2 studies of liver fibrosis in patients on methotrexate (89, 90)
- 1 study of TE in HIV infected patients (91)
- 1 study of TE in primary biliary cirrhosis and primary sclerosing cholangitis (92)
- 1 study of TE in nonalcoholic fatty disease (1).

The studies which established TE as a promising non-invasive device and gave it prominence were 3 validation studies that compared the LSM with liver biopsy (77, 78, 84).

Two of these were studies in HCV patients and were published in 2005. The other assessed chronic liver disease patients of multiple aetiologies and was published in 2006. These landmark studies are summarized as follows:

Ziol et al. correlated TE with liver biopsy in a multicenter French study of 327 chronic hepatitis C patients. LSM correlated well with fibrosis stage. Areas under the receiver operator characteristics (AUROC) curves were 0.79 for F≥2, 0.91 for F≥3, and 0.97 for F= 4. Optimal stiffness cutoff values of 8.7 and 14.5 kPa were found for F≥2 and F= 4 respectively (77).

Castera et al. assessed the performance of Fibroscan in patients with chronic hepatitis C, in comparison against and in combination with Fibrotest and the aspartate transaminase to platelets ratio index [APRI]. One hundred and eighty three consecutive patients with chronic hepatitis C were recruited at a single centre. Optimal LSM Cut-off values were 7.1 kPa for F≥2, 9.5 kPa for F≥3, and 12.5 kPa for F=4. The diagnostic performance of Fibroscan and Fibrotest were similar, but both superior to APRI. The AUROCS for Fibroscan, Fibrotest and APRI were: 0.83, 0.85, and 0.78, respectively, for F≥ 2; 0.90, 0.90, and 0.84, respectively, for F≥ 3; and 0.95, 0.87, and 0.83, respectively, for F=4. The best performance was obtained by combining the Fibroscan and Fibrotest, with an AUROC of 0.88 for F≥ 2, 0.95 for F≥ 3, and 0.95 for F= 4. When the Fibroscan and Fibrotest results agreed, liver biopsy examination confirmed them in 84% of cases for F≥2, 95% for F≥3, and in 94% for F=4 (78).

Foucher et al. examined 711 patients with chronic liver disease of multiple aetiologies at a single centre. The accuracy of TE compared to liver biopsy, and the ability of LSM to correlate with clinical outcomes were assessed. LSM significantly correlated with fibrosis stage. AUROC curves were 0.80 for F≥2, 0.90 for F≥3, and 0.96 for F=4. Using a cut off value of 17.6 kPa, patients with cirrhosis were detected with a positive predictive value (PPV) and a negative predictive value (NPV) of 90%. Liver stiffness significantly correlated with clinical, biological, and morphological parameters of liver disease. With an NPV >90%, the cut off values for the presence of oesophageal varices grade 2 or 3, cirrhosis Child-Pugh B or C, past history of ascites, hepatocellular carcinoma, and oesophageal bleeding were 27.5, 37.5, 49.1, 53.7, and 62.7 kPa, respectively (84).

Our literature review on TE at the time demonstrated that the majority of research had focused on the chronic hepatitis C population, with very little data on other populations with chronic liver disease.

However, on a worldwide scale, nonalcoholic fatty liver disease and chronic hepatitis B are both more prevalent (10, 93, 94). Long term methotrexate was thought for a very long time to cause liver fibrosis and required rigorous monitoring with repeated liver biopsies but data from small studies using TE demonstrated contrary evidence suggesting that methotrexate induced liver fibrosis is not as common as previously thought (89, 90). We sought to explore the clinical utilities of TE in the patient populations of NAFLD, CHB and long term methotrexate use.

This page has been left intentionally blank

CHAPTER 2: ASSESSING LIVER STIFFNESS USING TRANSIENT ELASTOGRAPHY IN PATIENTS WITH TYPE 2 DIABETES – A PILOT STUDY

2.0 CHAPTER SUMMARY

Introduction: The development of liver fibrosis from NAFLD is associated with an increased risk of mortality. Fibroscan ® measures the liver stiffness measurement (LSM) and has been demonstrated to assess liver fibrosis. We measured the LSM using Fibroscan in a cohort of patients with type II diabetes who are at risk of developing progressive liver disease and compared them with non-diabetic controls.

Methods: Subjects with type II diabetes, no prior documented liver disease and alcohol intake < 140g/week were recruited from outpatient specialist clinics. Demographic, clinical history and laboratory data were collected. Fibroscan ® was performed with the M probe. At least 10 successful measurements were required for a valid LSM. An interquartile range to median ratio of < 30% was required when the LSM≥7.1 kPa for a reliable LSM. The cutoff of ≥9.8 kPa was used to indicate advanced fibrosis.

Results: Valid scans were obtained in 88/97 (90.7%), while valid and reliable scans were obtained in 77/97 (79.4%) of subjects. A cut-off ≥9.8 kPa was present in 12% (10/77) diabetes subjects. The LSM was significantly correlated with BMI and the need for insulin therapy. Patients requiring insulin had LSM≥9.8 kPa with likelihood ratio (LR): 12.3, p=0.002. Obesity was associated with invalid scans and was also probably the cause of unreliable scans. The mechanism is thought to be due to greater subcutaneous adiposity and hence higher skin to liver capsule distance (SCD).

Conclusions: The 12% prevalence of high LSM in type II diabetes suggests that this is an at risk group for developing progressive liver disease. This is likely to be an under representation given that those who were obese were unable to have successful readings. Those who are obese and are on insulin therapy may represent a particularly high risk group. There is a high failure rate for obtaining valid and reliable scans which stems from using the M probe in obese patients. Further studies with larger sample size and using newer generations of Fibroscan with the XL probe are needed.

2.1 BACKGROUND

2.1.1 Non-alcoholic fatty liver disease and the spectrum of histological changes

Non-alcoholic fatty liver disease (NAFLD) is the most common chronic liver disease worldwide, with a prevalence of 20-35% in the general population (10, 95-97). NAFLD is defined as the presence of steatosis after alcohol and other secondary causes of fat accumulation in the liver have been excluded (98). NAFLD has a spectrum of distinct histological changes, which include simple steatosis (SS), non-alcoholic steatohepatitis (NASH) and fibrosis which may progress to cirrhosis.

Simple steatosis is described when there is greater than 5% of hepatocytes affected by steatosis, without any features of hepatocellular injury (98). The extent of steatosis can be mild (5 to 33% of hepatocytes are steatotic), moderate (34 to 66%), or severe (>66%). The pattern of steatosis typically is macrovesicular, though mixed micro and macrovesicles may also be seen (98). NASH features steatosis as well as hepatocyte injury (such as ballooning degeneration and fibrosis), and lobular inflammation (99-101). Liver fibrosis may eventually develop, but is not a required diagnostic feature of NASH (99-101). Non-NASH is used to describe SS, and "borderline NASH" - biopsies which feature mild inflammation but without enough features of hepatocyte injury to diagnose as "definite" NASH (99-101) .

2.1.2 Natural history and prognosis of non-NASH, NASH and NAFLD fibrosis

The understanding of the natural history and prognosis of NAFLD comes mostly from longitudinal case control and cohort studies. Liver biopsy data is scarce and no large series that have assessed paired biopsies.

Patients with NAFLD are at risk of higher mortality compared to the general population. A study of 420 NAFLD patients in Olmsted County, Minnesota reported a standardized mortality ratio (SMR) of 1.34 after a median follow up of 7.6 years (102). Liver related mortality was the 3rd leading cause of death in this group of patients behind malignancy and ischaemic heart disease (102).

NAFLD is associated with metabolic syndrome, and when cohorts are compared in studies that have controlled for metabolic risk factors, the difference in mortality is not so profound. The US National Health and Nutrition Examination Survey (NHANES) of 14000 patients reported that after metabolic syndrome and other confounders are matched, mortality in NAFLD has a Hazards Ratio (HR) of only 1.038 (103). However, liver related mortality remained much higher: HR=9.32 (103).

Several studies have demonstrated that within groups of patients who have NAFLD, those with the NASH subtype have greater mortality compared with Non-NASH. In a study 129 NAFLD subjects compared with matched controls, survival after a median 13.7 years follow up was reported to be lower in NASH (70% versus 80% p=0.01), but was not significantly different in non-NASH subjects (104). Cleveland clinic registry data reported NASH vs Non-NASH liver related mortality to be 11% compared to 2% over a median follow up of 8 years (105), and 18% compared to 3% over a median of 18.5 years

respectively (106). A Swedish study of 256 NAFLD patients reported an SMR of 1.55 for SS and an SMR of 1.86 for NASH over a median follow up of 28 years (107).

Recent data suggests liver fibrosis may be the most important histological feature when it comes to predicting poor liver related outcomes in NAFLD. An international study of 619 NAFLD patients followed over a median of 12.6 years reported that fibrosis was the only histological feature of liver biopsies significantly associated with death or liver transplantation (108). The hazard ratio calculated against those with no fibrosis, increased progressively with the Metavir fibrosis stage (see section 1.5 for description of the Metavir stages): F1 HR=1.88 (95% confidence interval [CI], 1.28–2.77), F2 HR= 2.89 (95% CI, 1.93–4.33), F3 HR=3.76 (95% CI, 2.40–5.89), and F4 HR=10.9 (95% CI, 6.06–19.62). The presence of fibrosis was associated with lower survival regardless of the presence of NASH and furthermore, NASH was not associated with lower survival compared with non-NASH unless fibrosis was present (108).

2.1.3 Non-invasive assessment of NAFLD and fibrosis

As NASH and NAFLD with fibrosis are associated with increased liver related mortality, patients who have developed either of these conditions should be the considered for monitoring and interventions. Liver biopsy is traditionally considered the reference standard for diagnosis and is recommended by international and regional guidelines (11, 97), but is not an ideal method of assessment for liver fibrosis. Pain and risk of complications make it poorly accepted by patients. In addition, liver biopsy has resource intensive requirements, and can produce variable results due to interobserver interpretation.

A pertinent need arises for non-invasive methods to diagnose NAFLD fibrosis and NASH. Transient Elastography with Fibroscan ® allows for rapid and non-invasive measurement of the tissue stiffness. Most data on its utility was in those with chronic hepatitis C. (77, 78, 84). In contrast, only 1 study was published in 2009 in its role in NAFLD (1). Thus we sought to further explore the use of Fibroscan in NAFLD.

Subsequent to this study, we performed a systematic review of all non-invasive measures of NASH and NAFLD fibrosis with a meta-analysis focusing on the 2 modalities that had been the most widely studied: TE and CK-18(109). This meta-analysis and review is presented in detail in chapter 3.

2.1.4 Epidemiology of fibrosis in NAFLD is unknown

The overall prevalence of NAFLD is reported to be 20-35%, but the epidemiology of the more severe disease phenotypes of NASH and NAFLD with fibrosis is incomplete due to the difficulty of obtaining liver biopsies in large series. No data exists for the prevalence of steatohepatitis and NAFLD fibrosis for the general population. A recent review estimated the prevalence of NASH to be 3-5%, but provided none for fibrosis in NAFLD (10).

Type 2 Diabetes is a major risk factor for NAFLD, with prevalence rates of 63 -70% based on ultrasound (110-112). A study of asymptomatic mostly male middle aged and overweight adults reported NAFLD and NASH respectively in 46% and 12.2% overall, but was higher in diabetics with 74% and 22.2% respectively (95). Fibrosis occurred in 9 subjects (2.7%) overall.

2.1.5 Study goals, hypothesis and specific objectives

Given the gap in data on liver fibrosis in NAFLD patients, we sought to explore the use of TE in assessing liver stiffness as a surrogate marker of liver fibrosis in a cohort of patients with type II diabetes who are at higher risk for NAFLD and NAFLD fibrosis. Based on previous studies, we estimate the rate of advanced fibrosis to be 5% in this cohort (95).

- Objective 1: To determine the rate of advanced fibrosis in type II diabetes by using the cutoff of LSM ≥9.8kpa established by Yoneda et al. (1).
- Objective 2: To compare the LSM in type II diabetes patients and matched healthy controls without diabetes or known liver disease.
- Objective 3: To identify factors associated with high LSM results
- Objective 4: To determine the success rate of obtaining valid and reliable Fibroscans and factors that influences this.

2.2 METHODS

2.2.1 Patient Selection and data collection

Consecutive type II diabetic patients who attended the Diabetes and Endocrinology outpatient clinics at Concord Repatriation Hospital (Concord Sydney, Australia) from June 2009 to November 2009 were recruited into the study. Patients who were younger than 18 years or were pregnant were excluded from the study. Patients with a history of alcohol intake of greater than 20g per day, or had prior diagnosed chronic liver disease other than NAFLD were excluded from the study. These included viral hepatitis B or C, recent viral hepatitis A or E (within 1 year of recruitment), haemachromatosis, cardiac hepatic congestion, autoimmune hepatitis, primary or secondary biliary cirrhosis, primary or secondary sclerosing cholangitis, Wilson's disease and alpha-antitrypsin deficiency. Patients with prior diagnosis of secondary causes of fatty liver (steroids, tamoxifen, amiodarone, thyroid disease) were also excluded. All patients gave informed written consent. A total of 101 subjects were enrolled.

Clinical assessment was performed using a standardized questionnaire. Age, gender, duration of diabetes, pharmacological treatment, alcohol intake, diagnosis of any pre-existing liver disease and other co-morbid medical history were recorded. Height, weight were measured and BMI calculated.

Biochemistry results within 1 month to the date of the Fibroscan were recorded including : bilirubin (BR), alkaline phosphatase (ALP), gamma-glutamyl transpeptidase (GGT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), albumin (ALB), platelet count, international normalized ratio (INR) alpha fetal protein (αFP) and HbA1c.

Non diabetic controls with no known liver disease were recruited from staff at the clinics and from persons accompanying patients who attended the outpatient clinic. After matching for age, sex, height, weight, body mass index (BMI) and weekly alcohol intake, 26 controls were included in the analysis.

Fibroscan was performed as previously described (section 1.7). The author performed all the TE scans.

2.2.2 Statistical analysis

Statistical analysis was performed using the Statistical Package for Social Sciences (SPSS) version 17.0. Continuous variables were expressed as the mean ± standard deviation where appropriate. LSM between groups were compared using Independent samples T-test. Categorical variables were compared using the chi-square test and Fisher's exact test when appropriate. Multivariate analysis was performed using multiple stepwise linear or logistic regression on variables where appropriate. The LSM cut-off values used for advanced fibrosis (F≥3) was 9.8 kPa, as reported from the validation study by Yoneda (1).

2.3 RESULTS

2.3.1 Characteristics of the study population

Table 3 shows the characteristics of the study population. The mean age was 60.8 years (yrs), Standard Deviation (SD) 11.4 yrs with 40.3% being female. The mean duration of diabetes was 9.5 (SD 7.8) yrs with a mean HbA1c of 8.2% (SD 1.8) with 32.5% of patients on insulin therapy. The mean BMI was elevated: 30.6 SD 7.1 kgm^{-2}. Mean alcohol intake was 15 (SD 42) g/week (range 0-140 g/week). Apart from mildly elevated GGT (64 SD 104 IU/L), The mean of all other biochemical parameters were all within the normal reference range.

Table 3: Clinical and biochemical characteristics of the Study Population

Characteristic	Mean (SD)[a]
Age (yrs)	60.8 (11.4)
Female	31/77 (40.3%)
Duration of Diabetes (yrs)	9.5 (7.8)
HbA1c (%)	8.2 (1.8)
Insulin therapy	25/77 (32.5%)
Cholesterol (mmol/L)	4.2 (1.0)
HDL (mmol/L)	1.2 (0.3)
LDL (mmol/L)	2.2 (0.8)
Triglycerides (mmol/L)	1.8 (1.0)
Height (m)	1.64 (0.11)
Weight (kg)	82.5 (22.1)
BMI (kgm^{-2})	30.6 (7.1)
Alcohol intake (g/wk)	15 (42)
BR (µmol/L)	9 (4)
ALB (g/L)	44 (4)
ALP (IU/L)	92 (73)
GGT (IU/L)	62 (104)
ALT (IU/L)	32 (21)
AST (IU/L)	28 (21)
PLT ($x10^9$/L)	272 (107)
INR	1.0 (0.2)
αFP (ug/L)	1.9 (1.0)

a. The mean and standard deviation is shown except for female and insulin therapy where the frequency is represented

2.3.2 Patient recruitment, invalid and reliable Liver Stiffness Measurements

One hundred and one consecutive patients were identified as potential candidates for the study. After screening, 4 were ineligible. Two patients had type I diabetes, one patient excessive alcohol and one had haemachromatosis. Ninety seven patients were recruited for the study and had fibroscan performed. There were 9 subjects who had invalid LSM, defined as less than 10 measurements acquired. Out of the 88/97 (90.7%) subjects remaining, 11 subjects had unreliable results, defined by an IQR/M ratio of greater than 30% when the LSM≥7.1 kPa. This left 77/97 (79.4%) subjects with valid and reliable scans ***(Figure 11)***

A comparison of the factors between those who had valid LSM versus invalid LSM was performed. The mean BMI was significantly greater in those with an invalid LSM: 36.7 (SD 6.3) kgm^{-2} vs 30.1 (SD 6.9) kgm^{-2}, p=0.044. No other factors were found to be significant and a summary of the comparison is shown in ***Table 4***.

A comparison of the factors between subjects who had a reliable LSM versus an unreliable LSM was performed. In univariate analysis, LDL, LSM, IQR/M ratio and success rate were statistically significant. In multivariate analysis using a validated binary logistic regression model (Hosmer and Lemeshow test p=0.492), the only significant difference was a greater mean LSM in subjects with unreliable LSMs: 13.7 (SD 8.6) kPa vs 6.8 (SD 2.8) kPa (B coefficient 0.518, Wald statistic p = 0.049). A summary of the comparison is shown in ***Table 5***.

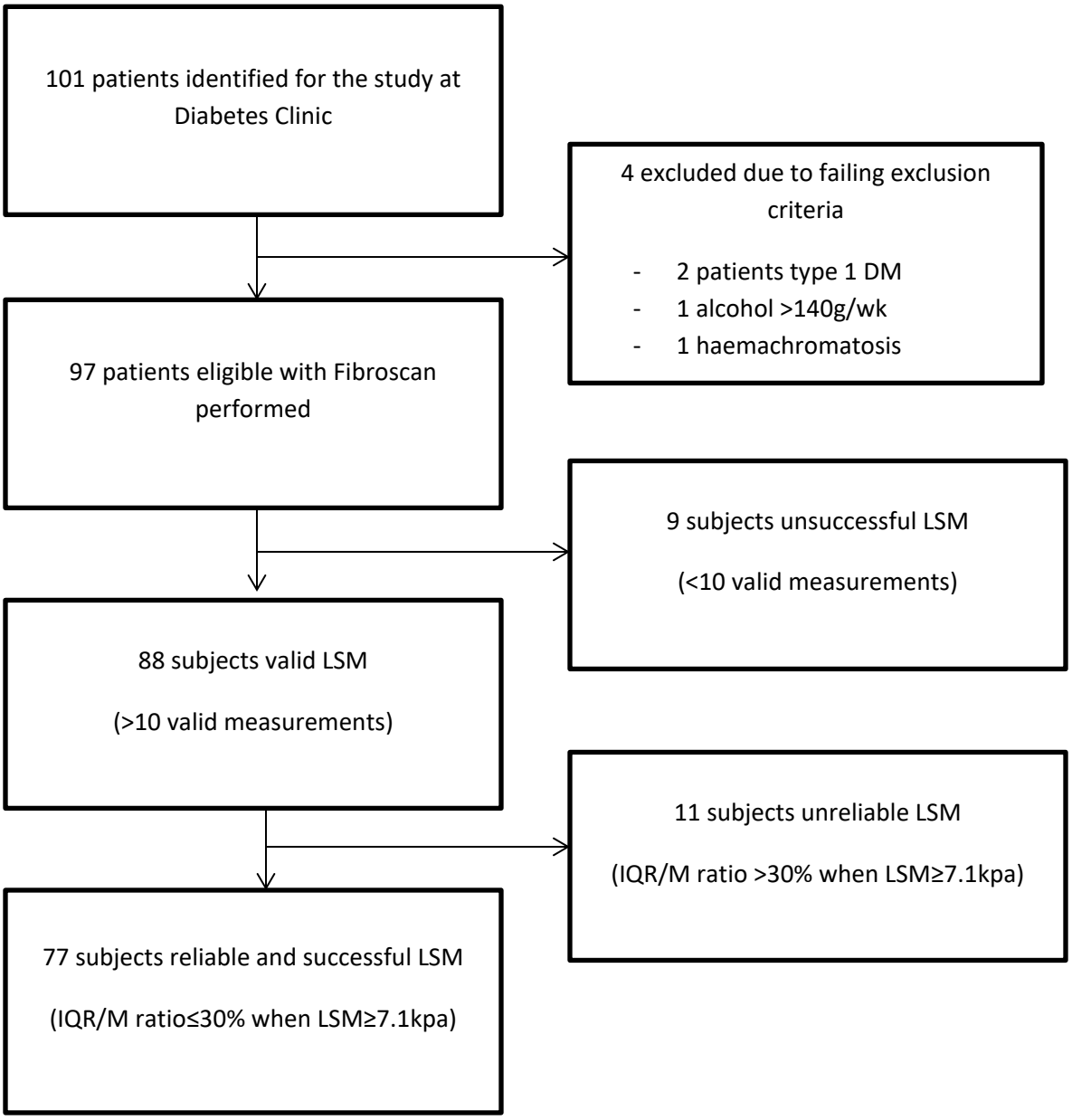


Figure 11: Patient recruitment flow chart

Table 4: Comparison of valid versus invalid Liver Stiffness Measurement subjects

Patient Characteristic	Valid LSM (n=88)	Invalid LSM (n=9)	P value
Age (yrs)	60.6 (11.6)	62.0 (9.9)	0.734
Gender Female[a]	36/88 (40.9%)	6/9 (66.6%)	0.137
Duration of Diabetes (yrs)	9.1 (7.3)	15.0 (12.3)	0.055
HbA1c (%)	8.2 (1.8)	8.5 (0.5)	0.754
Insulin therapy[a]	31/88 (35.2%)	6/9 (66.6%)	0.069
Cholesterol (mmol/L)	4.2 (1.0)	4.0 (0.6)	0.710
HDL (mmol/L)	1.2 (0.3)	1.1 (0.1)	0.627
LDL (mmol/L)	2.2 (0.8)	1.7 (0.2)	0.298
Triglycerides (mmol/L)	1.8 (1.0)	2.8 (2.0)	0.464
Height (m)	1.64 (0.11)	1.63 (0.08)	0.847
Weight (kg)	81.1 (21.6)	98.6 (23.7)	0.044
BMI (kgm^{-2})	30.1 (6.9)	36.7 (6.3)	0.016
Alcohol intake (g/wk)	15.5 (42.7)	0	0.471
BR (mmol/L)	9 (4)	13 (7)	0.108
ALB (g/L)	44 (4)	45 (2)	0.565
ALP (IU/L)	92 (74)	82 (12)	0.820
GGT (IU/L)	62 (105)	70 (71)	0.891
ALT (IU/L)	31 (21)	46 (16)	0.232
AST (IU/L)	27 (21)	42 (21)	0.242
PLT ($x10^9$/L)	270 (108)	327 (65)	0.371
INR	1.0 (0.2)	1.0 (0.1)	0.944
αFP (ug/L)	1.9 (1.0)	2.4 (0.5)	0.620

a. Chi square test applied to gender female and insulin therapy variables. All other variables are continuous and were compared using the independent samples t-test.

Table 5: Comparison of reliable versus unreliable Liver Stiffness Measurement subjects

Patient Characteristic	Reliable LSM (n=77)	Unreliable LSM (n=11)	P value
Age (yrs)	60.0 (11.5)	65.1 (11.6)	0.173
Female[a]	31/77(40.3%)	5/11(45.5%)	0.315
Duration of Diabetes (yrs)	8.6 (7.3)	12.2 (6.6)	0.125
HbA1c (%)	8.2 (1.8)	8.1 (1.8)	0.890
Insulin therapy[a]	24/77(31.2%)	4/11(36.4%)	1.000
Cholesterol (mmol/L)	4.1 (1.0)	4. (1.0)	0.089
HDL (mmol/L)	1.2 (0.3)	1.4 (0.4)	0.108
LDL (mmol/L)	2.2 (0.7)	2.7 (0.9)	0.072[b]
Triglycerides (mmol/L)	1.8 (1.0)	1.7 (0.7)	0.807
Height (m)	1.64 (0.11)	1.65 (0.12)	0.748
Weight (kg)	79.7 (20.7)	90.7 (25.6)	0.114
BMI (kgm^{-2})	29.6 (6.7)	33.1 (7.7)	0.121
Alcohol intake (g/wk)	14.9 (42.8)	19.5 (43.7)	0.741
BR (mmol/L)	9 (4)	9 (3)	0.680
ALB (g/L)	44 (4)	43 (3)	0.641
ALP (IU/L)	91 (76)	104 (55)	0.541
GGT (IU/L)	52 (92)	131 (163)	0.025
ALT (IU/L)	30 (19)	39 (31)	0.205
AST (IU/L)	26 (19)	40 (29)	0.034[b]
PLT ($x10^9$/L)	270 (112)	270 (79)	0.994
INR	1.0 (0.2)	1.0 (0.1)	0.682
αFP (ug/L)	1.8 (1.0)	2.1 (1.1)	0.409
LSM (kPa)	6.8 (2.8)	13.7 (8.6)	<0.001[b]
IQR/median	0.22 (0.14)	0.53 (0.25)	<0.001[b]
Scan Success Rate	79.1 (22.8)	62.5 (22.0)	0.029[b]

a. Chi square test applied to gender female and insulin therapy variables. All other variables are continuous and were compared using the independent samples t-test.

b. Only the LSM value was significant in multivariate analysis

2.3.3 Distribution of liver stiffness measurements in the diabetes subjects

Using the LSM cutoffs derived from Yoneda's study (1), 45 /77 (58%) of subjects had at least Metavir fibrosis stage 1(F≥1). The number of subjects with F≥2 was 36/77 (47%), F≥ 3 was 9/77 (12%), and F4 1/77 (1%). This is summarized in ***Table 6***. The distribution of the LSM scores and their frequencies are shown in ***Figure 12***.

Table 6: Distribution of LSM and in diabetes subjects

LSM (kpa)(1)	Derived F Stage	N (total = 77)
≥5.9	≥1	45 (58%)
≥6.7	≥2	36 (47%)
≥9.8	≥3	9 (12%)
≥17.5	4	1 (1%)

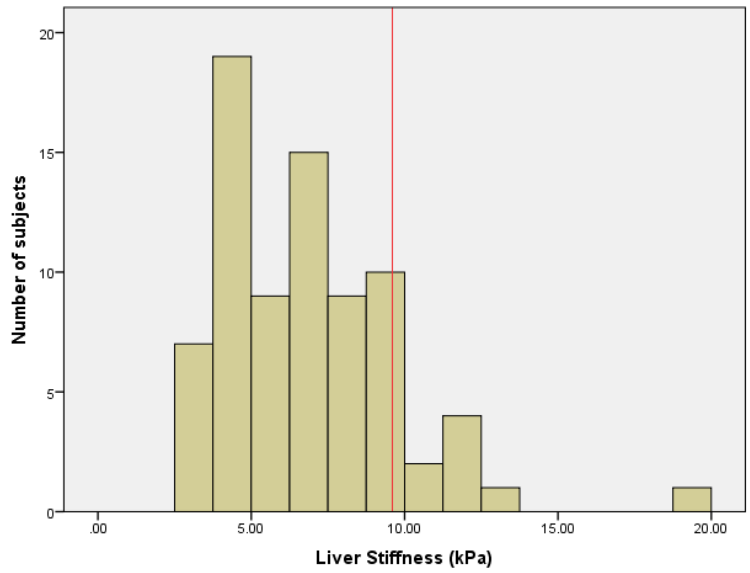


Figure 12: Distribution of Liver Stiffness in Diabetes subjects

Red line indicates LSM=9.8 kPa, the cutoff used for F≥3

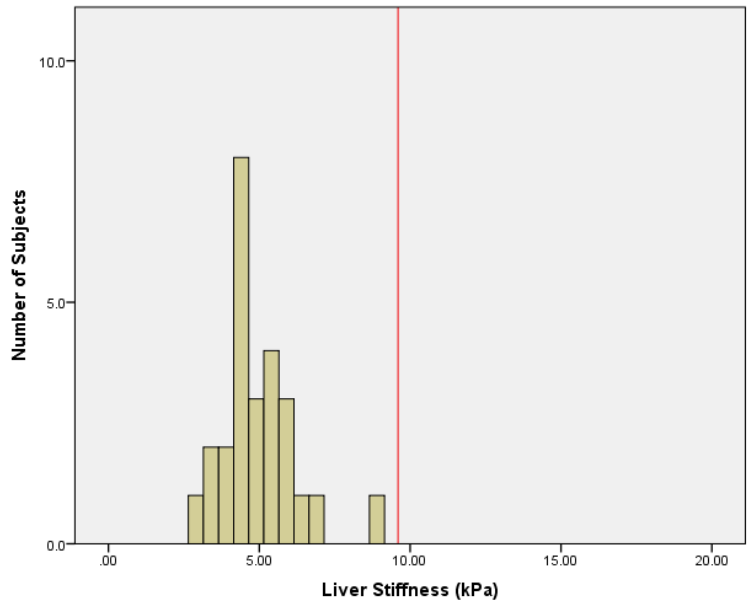


Figure 13: Distribution of Liver Stiffness in Controls:

Red line indicates LSM=9.8 kPa, the cutoff used for F≥3

2.3.4 Type II diabetes patients and non-diabetic controls

A comparison between type II diabetes patients in the cohort and non-diabetic controls with no known history of chronic liver disease was performed. Twenty six matched control subjects were included in the analysis. The LSM range for controls was 2.9-8.9 kPa ***(Figure 13)***. The mean LSM was found to be greater in type II diabetes cohort: 6.8 (SD 2.8) kPa versus 5.0 (SD 1.3) kPa, p= 0.002. No significant differences were found for age, gender, BMI, alcohol intake or IQ R/median ratio. The scan success rate was higher in type II diabetes 79.1% (SD 22.8%) versus controls 66.7% (SD 30.4%), p= 0.029. However, scan success is not a significant factor in the accuracy of readings (see 1.7.4). These findings are summarized in ***Table 7***.

Table 7: Comparison of characteristics between type II diabetes patients and healthy controls

	Type II Diabetes (n=77)	Non-diabetic controls (n=26)	P value
Age (yrs)	60.0 (11.5)	55.5 (13.3)	0.101
Female[a]	31/77 (40%)	15/26 (58%)	0.122
BMI (kgm^{-2})	29.6 (6.7)	28.3 (6.4)	0.379
Alcohol intake (g/wk)	14.9 (42.8)	25.0 (61.0)	0.370
LSM (kPa)	6.8 (2.8)	5.0 (1.3)	0.002
LSM ≥ 9.8 kPa	9/77 (12%)	0/26 (0%)	0.068
IQR/median	0.21 (0.14)	0.25 (0.15)	0.371
Scan Success Rate	79.1 (22.8)	66.7 (30.4)	0.029

c. Chi square test applied to gender and LSM≥9.8 kPa with the proportions and percentages in parenthesis reported. All other variables are continuous in which the mean value and standard deviation in parenthesis were reported. Comparisons were performed using the independent samples t-test.

2.3.5 BMI and insulin use are associated with liver stiffness in diabetes

In the diabetes cohort, univariate and multivariate regression analysis was performed to determine associations with liver stiffness. The factors found to have a significant association with LSM in univariate analysis include insulin therapy (r=-0.190, p<0.001); weight (r=0. 393, p= 0.001); BMI (r= 0.349, p= 0.002); and ALT (r=0.281, p= 0.016). In multivariate analysis, the only remaining associations with LSM were BMI and insulin therapy. These findings are shown in ***Table 8***.

Table 8: Correlations of patient characteristics with liver stiffness in diabetes patients

Patient Characteristic	Correlation (r)	P value
Age (yrs)	0.168	0.145
Female[a]	-0.201	0.098
Duration of Diabetes (yrs)	0.105	0.376
HbA1c (%)	0.101	0.395
Insulin therapy[a]	-0.190	<0.001[b]
Cholesterol (mmol/L)	0.061	0.624
HDL (mmol/L)	-0.016	0.906
LDL (mmol/L)	0.080	0.543
Triglycerides (mmol/L)	0.090	0.477
Height (m)	0.178	0.133
Weight (kg)	0.393	0.001[b]
BMI (kgm^{-2})	0.349	0.002[b]
Alcohol intake (g/wk)	-0.036	0.763
BR (mmol/L)	-0.043	0.720
ALB (g/L)	0.026	0.826
ALP (IU/L)	0.039	0.745
GGT (IU/L)	0.095	0.423
ALT (IU/L)	0.281	0.016[b]
AST (IU/L)	0.216	0.067
PLT ($x10^9$/L)	-0.172	0.160
INR	-0.116	0.375
αFP (ug/L)	-0.202	0.131

a. Categorical variables were converted to continuous variables to enable multiple regression
b. Only BMI and use of insulin therapy were significant in the multivariate model

2.3.6 Comparison of factors in diabetes patients with LSM<9.8 kPa versus LSM≥ 9.8 kPa (cut-off indicative of F≥3)

A comparison of the clinical factors was performed in those with an LSM <9.8 kPa versus ≥9.8 kPa. The only significant difference was that the greater proportion of those who use insulin therapy in those with LSM≥ 9.8 kPa group: 77.8% versus 25%, likelihood ratio (LR) 12.3, p= 0.002.These findings are shown in ***Table 9***.

Table 9: Comparison of characteristics between diabetic patients with LSM < 9.8 kPa and subjects with ≥ 9.8 kPa

Patient Characteristic	LSM<9.8kPa (n=68)	LSM≥9.8 kPa (n=9)	P value
Age (yrs)	59.8 (11.8)	61.8 (8.8)	0.625
Female[a]	28/68 (41.2%)	3/9 (33.3%)	0.652
Duration of Diabetes (yrs)	8.6 (7.5)	8.8 (5.8)	0.947
HbA1c (%)	8.1 (1.9)	8.7 (1.3)	0.399
Insulin therapy[a]	17/68 (25.0%)	7/9 (77.8%)	0.002 (LR 12.3)[b]
Cholesterol (mmol/L)	4.2 (1.1)	4.0 (0.6)	0.601
HDL (mmol/L)	1.2 (0.3)	1.2 (0.4)	0.835
LDL (mmol/L)	2.2 (0.8)	2.1 (0.5)	0.629
Triglycerides (mmol/L)	1.8 (1.0)	1.7 (1.0)	0.609
Height (m)	1.63 (0.11)	1.67 (0.12)	0.383
Weight (kg)	78.4 (20.2)	90.3 (23.7)	0.125
BMI (kgm^{-2})	29.3 (6.7)	32.2 (7.0)	0.258
Alcohol intake (g/wk)	15.3 (44.7)	11.9 (23.9)	0.832
BR (mmol/L)	9 (4)	9 (2)	0.645
ALB (g/L)	44 (4)	44 (2)	0.905
ALP (IU/L)	91 (81)	85 (14)	0.804
GGT (IU/L)	53 (98)	44 (25)	0.785
ALT (IU/L)	28 (19)	42 (20)	0.056
AST (IU/L)	25 (19)	34 (18)	0.145
PLT ($x10^9$/L)	273 (120)	251 (36)	0.602
INR	1.0 (0.2)	1.0 (0.0)	0.677
αFP (ug/L)	1.9 (1.1)	1.4 (0.4)	0.181

a. Chi square test applied to categorical variables. All other variables were continuous and the independent samples t-test was applied
b. The likelihood ratio (LR) was calculated for insulin therapy use in those with LSM≥9.8 kPa vs < 9.8 kPa.

2.3.7 Summary of the results

In this cohort of type 2 diabetes patients:

- 12% had an LSM of ≥ 9.8 kPa indicative of at least advanced fibrosis (F≥3)
- The mean LSM was significantly higher in diabetics compared with non-diabetic matched controls: 6.8 (SD 2.8) kPa versus 5.0 (SD 1.3) kPa, p= 0.002
- The LSM was associated with insulin therapy (r=-0.190, p<0.001) and BMI (r= 0.349, p= 0.002).
- Diabetes patients with an LSM ≥ 9.8 kPa were significantly more likely to be on insulin therapy, with a likelihood ratio of 12.3 (p=0.002)
- Valid LSM was obtained in 90.7%, while a valid and reliable LSM were obtained in 79.4% of all subjects
- Subjects with invalid scans had a significantly higher BMI: 36.7 (SD 6.3) kgm^{-2} vs 30.1 (SD 6.9) kgm^{-2}, p=0.044
- Subjects with unreliable scans had a significantly higher LSM: 13.7 (SD 8.6) kPa vs 6.8 (SD 2.8) kPa (B coefficient 0.518, Wald statistic p = 0.049)

2.4 DISCUSSION

2.4.1 Rate of advanced fibrosis (as indicated by LSM≥9.8 kPa) in diabetes subjects

In our study cohort, 12% of type II diabetes patients had an LSM of ≥ 9.8 kPa consistent with at least F3 or advanced fibrosis. In contrast, the highest LSM reading in a non-diabetic control was 8.9 kPa. The LSM was significantly higher in diabetes patients: 6.8 (SD 2.8) kPa versus 5.0 (SD 1.3) kPa, p= 0.002. This result suggests that diabetes is associated with the development of progressive liver damage as a consequence of NAFLD, and is consistent with other studies (95, 110-112). The high rate of 12% with elevated LSM indicative of advanced fibrosis suggests that this group of patients should be routinely screened for liver disease.

A point of contention of the findings may be that they are somewhat limited by the accuracy of TE and that the LSM cutoffs adopted were derived from a single study of Japanese patients which may not be generalizable. At the time of the research, the only available data was from Yoneda's study, in which the LSM cutoff of ≥9.8 kPa for F≥3 was reported to have high sensitivity (85%) and specificity (81%), with an overall excellent AUROC of 0.904. In the later meta-analysis (chapter 3), pooled sensitivity and specificity for F≥3 to both be 85%, with the LSM cut-offs used in studies ranging from 8.0-10.4 kPa (109). Therefore, the chosen cutoff of 9.8 kPa for this study was reasonable and a reliable reflection of F≥3 advanced fibrosis.

Whether NAFLD is the only cause of fibrosis in this group of type 2 diabetic patients could be questioned. Due to the cross sectional nature of the study, we were only able to rule out other causes of chronic liver disease with clinical history and medical records. Not every patient had full serological or biochemical tests to exclude viral hepatitis, hemochromatosis, Wilson's disease and other causes of chronic liver disease. However the assumption that these patients with type 2 diabetes develop liver fibrosis because of NAFLD is still a sound one to make. It is established that 63-70% of type II diabetes patients have NAFLD, making it by far the most likely cause of liver fibrosis in this population (110-112). Meanwhile in our local setting, other causes of chronic liver disease are uncommon. The prevalence of chronic hepatitis B and C are each estimated to be 1% (113, 114). The next most common liver disorder is hemochromatosis in which the homozygous state has a prevalence of 0.5% (115). Therefore even if there were patients with undiagnosed chronic liver disease from these other etiologies, the background prevalence would only be around 2-3% compared to the 63-70% probability of having NAFLD. Patients with excessive alcohol intake were excluded as per study criteria. The contribution to fibrosis prevalence from other causes rather than NAFLD would therefore be minimal.

2.4.2 Liver Stiffness is associated with BMI and insulin therapy

Our results show BMI and need for insulin therapy are correlated with the LSM, and that patients on insulin therapy is associated with LSM ≥ 9.8, with a likelihood ratio of 12.3 (p=0.002)

Obesity and insulin resistance are key components of the metabolic syndrome. The need for insulin therapy probably implies a state of increase insulin resistance, in which endogenous insulin is insufficient for glycaemic control, thereby necessitating exogenous insulin administration. Our findings are unsurprising given that obesity, insulin resistance and metabolic syndrome are recognized as key features in the pathogenesis of NASH and fibrosis in NAFLD (116). Our observations are compatible with reports of LSM being associated with the HOMA-IR score (117, 118) and obese patients having very high rates (57-98%) of NAFLD (119-121). Furthermore, advanced liver fibrosis has been linked with the presence of diabetes and obesity in biopsy series (122, 123). BMI in the obese range and the need for insulin therapy are clinical features that could be used to further stratify diabetes patients who are at high risk of progressive liver disease for screening.

2.4.3 LSM has no association with ALT, AST and lipid levels, duration of diabetes or HbA1c

Our study found no correlation between LSM with ALT, AST and lipid levels.

LSM has been shown to be affected by elevated ALT levels in studies of TE in the viral hepatitis population (88) . While some biopsy series have reported an association between raised transaminases and advanced fibrosis (122, 123), the majority of the TE literature that feature liver biopsy (7 out of 9 studies) have found no association (1, 117, 118, 124-129). This suggests that ALT may have minimal effect on LSM reading in the setting of NAFLD. This could be explained by hepatic necroinflammation not being as severe as that are seen in those underlying viral hepatitis flares.

Hyperlipidaemia is associated with liver fibrosis (130) while statin use has been found to have a protective effect (108) in other studies. No correlations between LSM and lipid levels were found in our study, probably because patients were mostly on pharmacological treatment. The use of lipid lowering therapy was not closely examined in this cross sectional analysis, and may be addressed in future studies.

2.4.4 Overestimations of liver stiffness can occur when using the M probe in obese patients with excessive skin capsule distance

At the time of writing, 9 studies had compared liver histology with LSM in NAFLD patients while also examining factors which were associated with LSM (1, 117, 118, 124-129). These were also the studies that met QUADAS criteria and analysed in the meta-analysis to be presented in chapter 3.

There were reports by 3 of these studies that obese range BMI was a confounding variable for liver stiffness (117, 124, 126). In patients who were obese, the LSM was inaccurate for assessing fibrosis stage. Obese BMI patients often had a high LSM, but a lower degree of fibrosis severity found on biopsy compared to non-obese individuals. Mechanisms postulated include greater steatosis itself causing high liver stiffness, leading to overestimation of the liver fibrosis stage (117, 124, 126).

However, the 6 other studies reported no association between the LSM and obese range BMI or steatosis that was independent or separate to liver fibrosis (1, 118, 125, 127-129). Without comparison to liver biopsy, is not possible in our study to clarify whether high liver stiffness in obese BMI patients is due to fibrosis, or due to an alternate confounding mechanism associated with obesity. It would appear unlikely obesity is a confounder given the majority of the literature has not reproduced similar findings.

There may be a simpler reason why obese patients have less accurate Fibroscans. It is well established that inaccurate LSMs can occur in patients who have an excessive skin to liver capsule distance (SCD) that is greater than the depth of measurement that the Fibroscan probe allows (see 1.4.7) (52). The main reason for an excessive SCD is subcutaneous fat, and hence obese individuals are most likely to have inaccurate scans.

The M probe has a measurement depth from 25 mm to 65 mm beneath the skin. When the SCD is less than 25mm, the measured region will contain liver tissue only, and so an accurate assessment occurs. But if the SCD is greater than 25mm, the measured region will contain non-liver tissue that is between the skin and the liver. This might include: the liver capsule, which is fibrous and likely to produce higher stiffness values compared to the liver parenchyma; the bony ribs and intercostal muscles, which would similarly produce higher stiffness results; and the subcutaneous fat, which probably has lower stiffness compared to liver leading to an underestimation. The effect is that the LSM will be assessed based on the stiffness of these non- liver tissues, leading to inaccurate results.

When non-liver tissue is measured, typical elastograms may contain A waves, E waves and angled waves (see section 1.7.2). These are sometimes undetected by the internal quality control software algorithm of the Fibroscan. Thus, achieving accurate measurements is reliant upon the operator to recognize incorrect scans when they are not automatically detected and discarded by the Fibroscan.

This issue may be resolved by using a probe that can obtain measurements to a greater depth underneath the skin. It is notable that the 3 aforementioned studies that reported inaccurate LSMs in

obese patients had only the M probe available for use. It is probable that the routine use of the M probe on obese subjects, who would tend to have excessive SCD, resulted in numerous occasions where non liver tissue was measured instead of the liver, resulting in systematic overestimation or underestimation. In the 6 other studies that did not find LSM to be overestimated in obese patients, their operators may have been more able to limit excessive overestimation/underestimation. The operator's experience, technique and ability to identify aberrant waves on the elastogram when the software fails to reject them are important factors in obtaining accurate Fibroscans.

2.4.5 Invalid and unreliable measurements can occur when using the M probe in obese patients with excessive skin to liver capsule distance

The M probe being insufficient for obese patients also provides an explanation for the occurrence of invalid scans (defined as unable to obtain minimum 10 measurements) and unreliable scans (defined as the IQR/M ratio $\leq$ 30% required if final stiffness result is < 7.1kpa) in our study.

Our findings of invalid scans being associated with a higher mean BMI: 36.7 (SD 6.7) kgm^{-2} vs 30.1 (SD 6.9) kgm^{-2}, p=0.016, is consistent with other studies (1, 117, 118, 124-129). This can be explained by obese patients having an excessive SCD. Scans were invalid because the measurement region of the M probe was not deep enough to measure the liver parenchyma only. This would be reflected as a non-linear A mode signal, leading to rejection of the elastogram by the internal software (see 1.7.2).

Unreliable scans were associated with having a much higher mean LSM: 13.7 (SD 8.6) kPa vs 6.8 (SD 2.8). The higher value of the LSM is not causative of unreliable scans, but rather a manifestation of the underlying reason. By definition, unreliable scans occur when the set of measurements for the LSM have a wide range, specifically when the interquartile range is greater than 30% of the median measurement. Nearby tissue types (such as ribs and the fibrous liver capsule) are much "harder" than the liver. There is a bigger difference in elasticity between the liver and these tissue types than compared to the different regions within the liver itself. Thus a wide range of measurements likely indicates that non liver tissues along with liver tissue are being concurrently scanned. These measurements are included within the set of valid measurements and contribute to calculating the median score. As hard structures will have high stiffness measurements, this leads to a higher median score being assigned as the LSM, along with a wider range of values within the set of measurements taken. The result is that the LSM is overestimated and has high IQR/M ratio more likely to be greater than 30% - the exact scenario that was observed in unreliable scans.

Unreliable scans due to measuring non-liver tissue is again most likely related to obese patients who have an excessive SCD. Those with unreliable scans were 11kg heavier (90.7kg vs 79.7kg) and BMI 3.5 kgm^{-2} greater (33.1 kgm^{-2} vs 29.6 kgm^{-2}). This difference is probably great enough to systematically cause an excessive SCD leading to unreliable scans, in spite of the difference not being statistically significant. Small numbers in the unreliable scan group (n=11) probably prevented statistical significance being reached.

2.4.6 Study limitations and further research with the XL probe

The rates of an invalid scan or an unreliable scan in our study were 9% and 11% respectively. Scans were both valid and reliable only in 79.4% of subjects. Other studies have reported similar rates of invalid scans using the M probe: 8 – 14% (1, 117, 118, 124-129). None have highlighted the rate of obtaining both a valid and reliable scan. The low rate means inaccurate assessments occur in approximately 1 in 5 patients which severely limits the usefulness of Fibroscan in NAFLD. Obese patients are more likely to be affected by NAFLD, but are also the group of patients in which Fibroscan is least able to obtain valid, reliable and accurate readings.

The limits of the M probe in inadequate measurement depth for obese individuals was recognized and addressed by development of the XL probe. The XL probe had not yet been developed at the time of this study, nor was available for any of the publications that reported high failure rates. The measurement depth for the M probe is 25-65mm beneath the skin, while the XL probe is able to measure between 35-75mm underneath the skin (131). The potential advantages of the XL probe are reviewed in the meta-analysis presented in the next chapter.

This study has several limitations including small number of patients, absence of reference liver biopsy, and other relevant clinical data, such as fasting insulin and glucose levels to calculate the HOMA-IR. Limitations of the use of the M probe in this cohort are also recognized.

Nonetheless at the time of research, little data existed for the use of Fibroscan for NAFLD. The study was intended to be used as a pilot to determine further studies, and a much larger and comprehensive study of TE in diabetic patients was later performed (132) which overcame many of the limitations of this pilot including the lack of XL probe, while also assessing the controlled attenuation parameter (CAP) a novel feature of Fibroscan that assesses steatosis. This is presented in chapter 4 of this thesis.

2.5 CONCLUSION

The prevalence of advanced fibrosis in patients with type II diabetes appears to be high. Higher liver stiffness was observed in diabetics with obese range BMI and who required insulin therapy. These clinical features could be used to target at risk patients for screening. Fibroscan may be of limited use in obese individuals due to overestimation of high liver stiffness values and a high rate of invalid and unreliable scans. Further studies with a larger study population and use of the XL probe are needed to clarify these issues.

This page has been left intentionally blank

CHAPTER 3: SYSTEMATIC REVIEW WITH META-ANALYSIS: NON-INVASIVE ASSESSMENT OF NON-ALCOHOLIC FATTY LIVER DISEASE-THE ROLE OF TRANSIENT ELASTOGRAPHY AND PLASMA CYTOKERATIN-18 FRAGMENTS

This chapter was published as:

Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease - the role of transient elastography and plasma cytokeratin-18 fragments.

Kwok R, Tse YK, Wong GL, Ha Y, Lee AU, Ngu MC, Chan HL, Wong VW

Alimentary pharmacology & therapeutics 2014; 39(3):254-69.

Raymond Kwok is the corresponding author and was primarily responsible for the literature review, data extraction, and writing the drafts. The study was co-designed and analysis co-interpreted with the co-authors.

3.0 CHAPTER SUMMARY

Background: Non-alcoholic fatty liver disease (NAFLD) affects 15-40% of the general population. Those who have steatohepatitis (NASH) and progressive fibrosis and would be candidates for monitoring and treatment.

Aims: To review current literature on the use of non-invasive tests to assess the severity of NAFLD.

Methods: Systematic literature searching identified studies evaluating non-invasive tests of NASH and fibrosis using liver biopsy as the reference standard. Meta-analysis was performed for areas with adequate number of publications.

Results: Serum tests and physical measurements like transient elastography (TE) have high negative predictive value in excluding advanced fibrosis in NAFLD patients. The NAFLD fibrosis score comprise 6 routine clinical parameters and has been endorsed by current AASLD guidelines as a screening test to exclude low-risk individuals. The pooled sensitivities and specificities for TE to diagnose F≥2, F≥3 and F4 disease were 79% and 75%, 85% and 85%, and 92% and 92%, respectively. Liver stiffness measurement often fails in obese patients, but the success rate can be improved with the use of the XL probe. A number of biomarkers have been developed for the diagnosis of NASH, but few were independently validated. Serum/plasma cytokeratin-18 fragments have been most extensively evaluated and have a pooled sensitivity of 66% and specificity of 82% in diagnosing NASH.

Conclusions: Current non-invasive tests are accurate in excluding advanced fibrosis in NAFLD patients and may be used for initial assessment. Further development and evaluation of NASH biomarkers are needed.

3.1 INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) is currently the most common chronic liver disease worldwide, affecting 15-40% of the general population (95, 133). Depending on the presence of necroinflammation and hepatocyte ballooning, NAFLD is further divided into non-alcoholic fatty liver (NAFL) and non-alcoholic steatohepatitis (NASH) (97). NASH is the active form of NAFLD. It occurs in 10-20% of NAFLD patients and may progress to cirrhosis and hepatocellular carcinoma (HCC) (17, 58, 59). Since NAFLD is highly prevalent but the majority of patients only has NAFL and run a benign course, it is important to identify patients with NASH and NAFLD fibrosis efficiently.

Traditionally, liver biopsy is the primary method to assess the severity of NAFLD. However, it is an invasive procedure and carries a small but definite risk of complications. Furthermore, it is unrealistic to perform liver biopsy for 15-40% of the general population. In recent years, a number of blood tests and physical assessments have been developed to aid evaluation of NAFLD patients. Therefore, it is timely to appraise the diagnostic performance of these non-invasive tests.

Our review will focus on the diagnosis of fibrosis and NASH in the NAFLD spectrum. The diagnosis of simple steatosis was not chosen as a primary focus of review because of its generally innocuous nature. Some non-invasive tests have been much more widely studied and used compared with others. After considering expert opinion and existing reviews on the topic, an in-depth assessment on the performances of transient elastography (TE), cytokeratin-18 fragments (CK18) and acoustic radiation force impulse (ARFI) was deemed appropriate. TE and CK18 are amongst the most widely studied modalities in NAFLD fibrosis and NASH, while ARFI has generated much recent interest. Thus in addition to a systematic review of the variety of non-invasive diagnostic methods in NAFLD fibrosis and NASH, we performed a meta-analysis on the use of TE, CK18 and ARFI.

3.2 METHODS

3.2.1 Literature Search

A systematic web-based literature search of all publications in MEDLINE (via OvidSP), PUBMED (NLM) and EMBASE was conducted on 13th June 2013 from the date of inception for each of the databases. Our primary search strategy for identifying studies comprised of using free-text words (fatty liver, non-alcoholic fatty liver disease, non-alcoholic steatohepatitis, transient elastography, Fibroscan, liver stiffness measurement, elastography imaging techniques, acoustic radiation force impulse, keratin 18, cytokeratin 18). Two reviewers (RK and VWSW) performed literature search separately and agreed upon the final selection of studies. Search limits included English language, abstracts, and publication in peer-reviewed journals. A secondary search was performed to locate any potential studies missed by electronic search strategies. A comprehensive search of MEDLINE was performed for locating any existing systematic reviews on transient elastography, ARFI and CK18 in the diagnosis of NAFLD. Manual searching of reference lists from relevant reviews and primary studies and was performed. No additional suitable studies were found.

3.2.2. Meta-analysis

All candidate articles from our primary search had its abstract or full text scrutinized to determine whether it was a primary study. Subsequently the full text was further assessed to check for fulfillment of the inclusion/exclusion criteria. Disagreements were resolved through consensus. Inclusion/exclusion criteria for primary studies required the following features:

(1) Detailed description of adult human subjects under study

(2) Description of TE, ARFI or CK18 as an index test

(3) Description of liver biopsy as the reference standard. The definition of NASH was taken as the NALFD activity score ≥ 5. Fibrosis staging based on the Brunt's or Kleiner's system: F0 = no fibrosis; F1 = perisinusoidal or portal; F2 = perisinusoidal and portal/periportal; F3 = septal or bridging fibrosis; and F4 = cirrhosis (60, 61).

(4) A minimum number of NAFLD subjects ≥ 20

(5) Results describe number of cases of fibrosis for each stage or NASH using liver biopsy, the sensitivity, specificity and nominated cut off values of the index test so that a 2x2 table could be created. Corresponding authors were asked to provide study level data if adequate information could not be extracted from the published article.

(6) Different articles from a primary study that contained overlapping data cohorts were only counted once. The most suitable article to use was determined by seeking clarification from the authors, or by using the most updated manuscript that contained all the required data.

Both prospective and retrospective studies were acceptable. Studies in which subjects had other causes of chronic liver disease apart from NAFLD were included so long as discrete data for NAFLD population could be extracted. Studies which reported other noninvasive comparators were also allowed if the discrete data for TE, ARFI and CK18 could be extracted.

A final number of 9 articles for TE (1, 117, 118, 124-129), 11 articles for CK18 (134-144) and 2 articles for ARFI (128, 145) were assessed to be suitable for inclusion in the meta-analysis. There were too few studies for statistical analysis on the ARFI data.

Figure 14 outlines the stepwise evaluation and selection process for all the candidate studies.

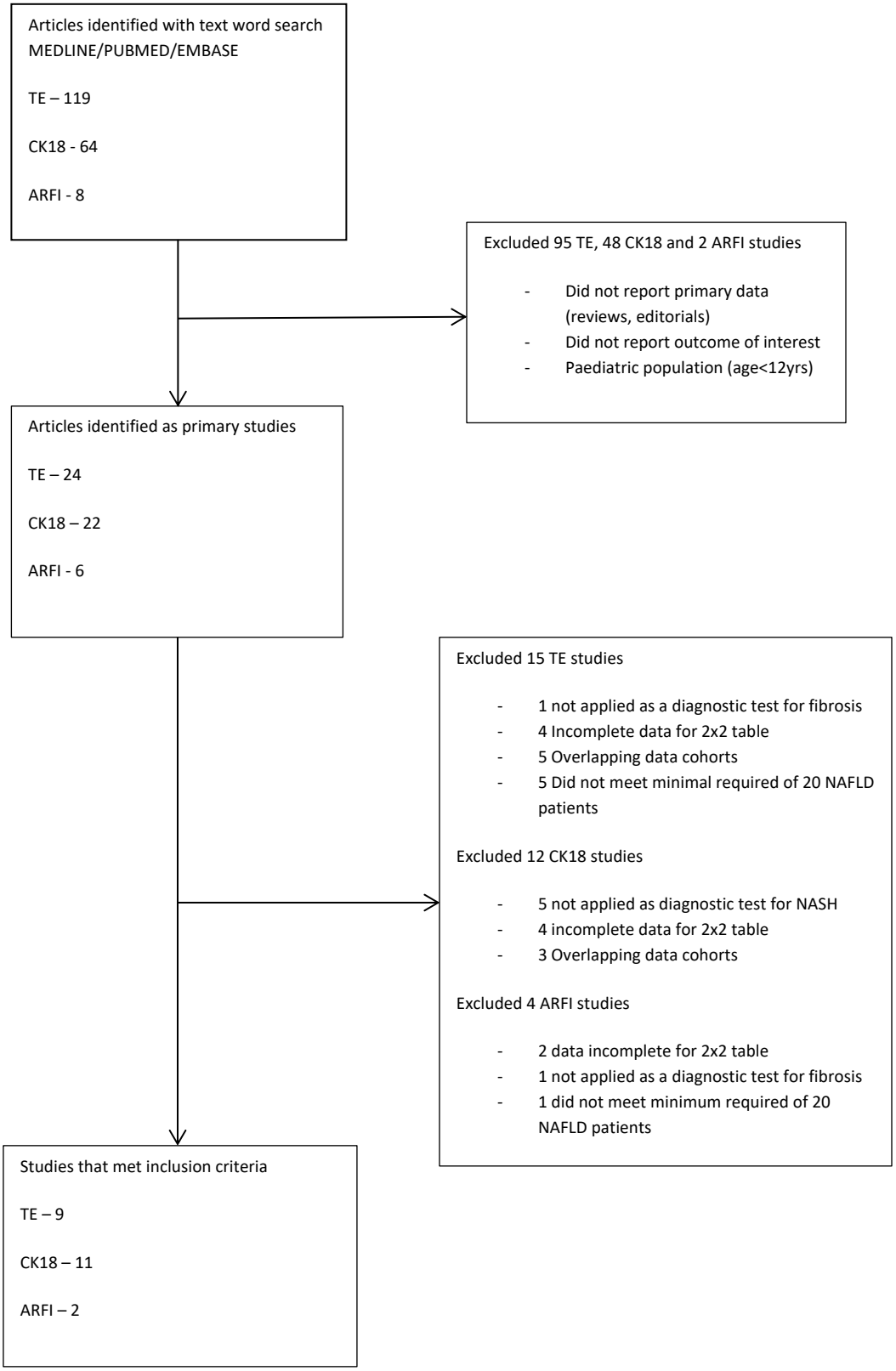


Figure 14: Summary of literature search and selection

3.2.3 Quality assessment

Each study's quality was analyzed by independent reviewers (RK, YKT). A modified version of the QUADAS (146) was used to assess the quality of the studies included for meta-analysis ***(Table 12/Supplementary Table 1)***. Consensus was reached in disagreements by referral to a third reviewer (VWSW).

TE studies overall scored highly on the QUADAS assessment ***(Figure 17/Supplementary Figure 1***). Two studies scored 12/13 whereas the rest scored 13/13. CK18 studies had a mean QUADAS score of 11.2 (range 9-13) ***(Figure 18/Supplementary Figure 2)***. The most common components in which studies lost points were an unclear description of the quality of liver biopsies (36% studies had high quality data), whether the histopathologist was blinded to other results (45%), and unclear descriptions of when serum was obtained for CK18 analysis in relation to the timing of liver biopsy (64%).

3.2.4 Data extraction

Two reviewers (RK, YJH) independently extracted the required information from primary studies. A data extraction pro-forma was created and variables included for collection were: patient age, sex, ethnicity, BMI, transaminase levels, results of the index and reference tests and accompanying diagnostic thresholds (cut-offs). Where available, other biochemical and blood parameters, presence of metabolic syndrome components and risk factors (other anthropometric measures, diabetes, hypertension, hypercholesterolemia and hypertriglyceridemia) were recorded. A 2×2 table was created for each modality and its reported cut-off for diagnosing each category.

3.2.5 Data synthesis and statistical analysis

From the 2×2 tables we calculated sensitivity and specificity. The estimates of sensitivity and specificity and their associated 95% confidence intervals (CIs) were presented graphically by plotting in paired forest plots. Summary estimates of sensitivity and specificity, along with 95% CIs, were obtained by using the bivariate random-effects modeling approach (minimum 4 studies) (147). Besides accounting for study size and between-study heterogeneity using a random effects model, the bivariate analyses enable correctly dealing with any possible negative correlation that might arise between the sensitivity and specificity. Moreover, we constructed a hierarchical summary receiver operating characteristic (HSROC) curve plotting sensitivity versus specificity (148). The HSROC curve illustrates the summary tradeoff between sensitivity and specificity across the studies.

To examine the potential sources of heterogeneity, we predefined the following covariates: body mass index (< 30kg/m^2 *vs* ≥ 30kg/m^2, for TE), and study quality factors (yes *vs* unclear *vs* no, for individual QUADAS item as described above). Separate bivariate models were simply performed to different subgroups of studies because sufficient data were not available (at least 10 studies) to allow adding

covariates to the hierarchical model by means of meta-regression. The two studies by Yoneda et al. had much higher cutoffs for F4, compared with the other studies included in the meta-analysis ***(Figure 16)***. In order to assess the effect on the pooled results, a post-hoc sensitivity analyses was conducted to calculate pooled estimates of sensitivity and specificity in the bivariate model by excluding these two studies. Statistical analyses were performed using STATA 10.0 (StataCorp, College Station, TX, USA), particularly the metandi (149) commands and Review Manager (150) software. All statistical tests were two-sided, with a p value < 0.05 indicating statistical significance.

3.3 NON-INVASIVE DIAGNOSIS OF NASH

NASH is the active form of NAFLD with necroinflammation and hepatocyte ballooning. With ongoing liver injury, NASH may progress to cirrhosis and HCC. In long-term follow-up studies, histological features of NASH predict future liver complications (104, 151). Previously, NAFL and NASH were considered distinct entities. However, recent longitudinal studies with paired liver biopsies suggest that some patients with NAFL may progress to NASH (17, 152). In any case, assessment of disease severity is important for prognostication and treatment monitoring.

3.3.1 Serum biomarkers

3.3.1.1 Cytokeratin-18 fragments

Cytokeratins are keratin-containing proteins that form intermediate filaments and comprise the structure of cytoskeletons of epithelial cells. CK18 is found predominantly in glandular epithelia of the digestive, respiratory and urogenital tracts. It is the major intermediate filament protein of the liver. During apoptosis of hepatocytes, capsases cleave CK18 generating fragments that can be detectable using immunoassays [153]. It is one of the most widely investigated biomarkers for NASH as a standalone test or as part of prediction models. The two main enzyme assays of CK18 that have been studied are M30 and M65, which supposedly measure hepatocyte apoptosis and total cell death, respectively.

Meta-analysis on CK18

We performed a meta-analysis of 11 studies with a total pool of 822 patients, in which 389 had histological NASH ***(Table 13/Supplementary Table 2)***. Since M30 and M65 had similar performance and M30 was more widely studied, we decided to focus on M30. The studies were further grouped according to whether a separate 'high sensitivity' and 'high specificity' cut-off (6 studies) was chosen, and/or a single 'best' overall cut-off level (7 studies) was used to diagnose NASH. In the 6 studies that chose separate cut-offs, for 'high sensitivity', the CK 18 cut-off chosen ranged 111.6 – 380.0 U/L (77-90% sensitivity and 34-94% specificity) ***(Figure 15)***. For 'high specificity', the cut-offs chosen ranged 261.4 – 670 U/L (24-86% sensitivity and 91-100% specificity). The AUROC for these 6 studies were 0.71-0.93. For the 7 studies that reported a single 'best' overall cut-off, the range of chosen cut-offs was 121.6-338.0 U/L, with 60-88% sensitivity, 66-97% specificity and AUROC 0.70-0.87.

In the pooled estimates of diagnostic accuracy, the 7 studies which used a single 'best' overall cut-off level showed 66% sensitivity and 82% specificity. In the 6 studies using separate 'high sensitivity' and 'high specificity' cut-offs, the pooled estimates were 82% sensitivity, 65% specificity and 58% sensitivity and 98% specificity, respectively. Pooled estimates of diagnostic accuracies remained stable when only studies with high quality were analyzed ***(Table 14/Supplementary Table 3)***. ***Figure 19/Supplementary Figure 3*** showed the HSROC plots of CK18.

Discussion on CK18

Our findings suggest that CK18 has moderate accuracy overall for diagnosing NASH (66% sensitivity, 82% specificity). When optimal cut-offs are used, sensitivity improves to 82%, while specificity is 98%. However, there is considerable variability in the suggested cut-offs and their respective diagnostic accuracy among studies. In clinical practice, this makes choosing which threshold to use very difficult. The variability may be partly explained as by choosing an optimal threshold to maximise either sensitivity or specificity, the accuracy of the other is greatly sacrificed. Other possible causes of heterogeneity include intervals between blood tests and liver biopsy, inadequate description of liver biopsy assessment and blinding, and inadequate reference test description. However, none of these was found to be significant, with only small differences in overall sensitivities and specificities in these subgroups ***(Table 14/Supplementary Table 3).***

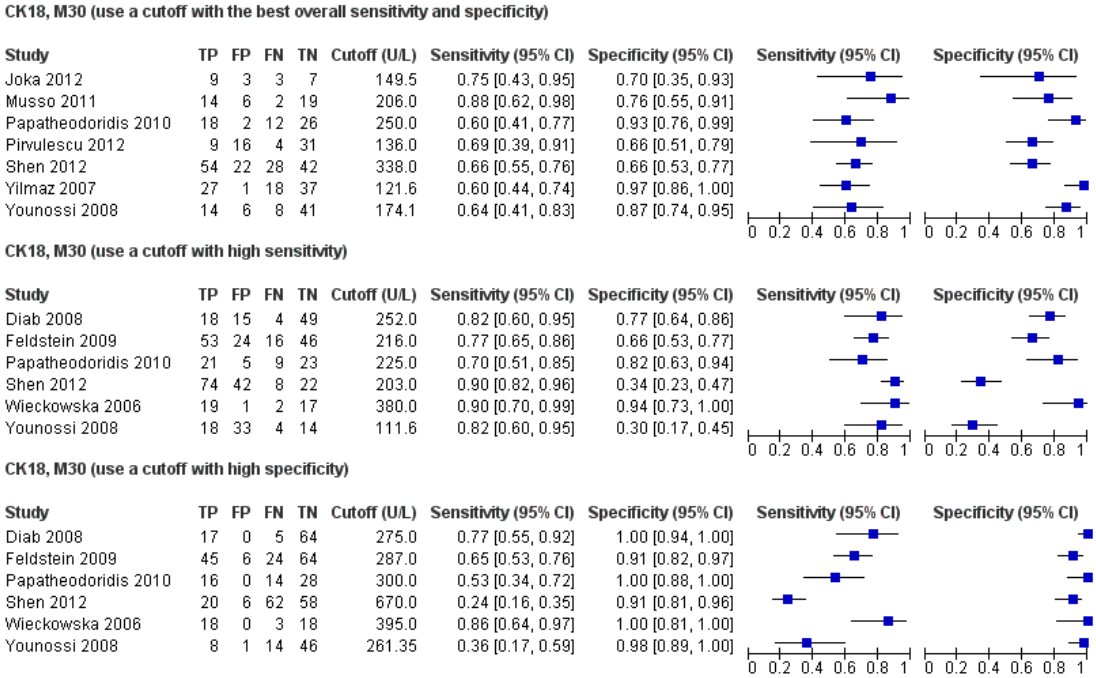


Figure 15: Forest plot from meta-analysis for CK18

Forest plot from meta-analysis of sensitivities and specificities for CK18 to diagnose NASH using a random-effect model. Cutoffs with the best overall accuracy, sensitivity and specificity in individual studies were adopted. TP:True Positives, FP: False Positives, FN: False Negatives, TN: True Negatives.

3.3.1.2 Other biomarkers

Soluble sFas (sFAS) is a death receptor from the TNFR family that has been implicated in apoptosis and is upregulated in NASH in animal models. An apoptosis panel combining CK18 with sFAS was found to have greater AUROC than either alone (154).

Tumor necrosis factor-alpha (TNF-α) is a proinflammatory cytokine which has been proven to play important roles in pathogenesis of NAFLD. Several studies demonstrated that TNF-α contributed to NASH development in that NASH patients or animal models exhibit elevated serum TNF-α (155-158). However, its diagnostic performance of differentiating NASH from NAFL has not been fully elucidated.

Another cytokine, interleukin-6 (IL-6) was elevated or upregulated in serum or liver tissue of NASH patients as stated by some independent studies (155, 159, 160), but did not show any difference between NASH and NAFL in other studies (161, 162). Grigorescu *et al* evaluated the accuracy of IL-6 as a noninvasive test for discriminating NASH from 79 NAFLD patients (160). At a cut-off of 6 pg/ml, the sensitivity and specificity were reported as 64% and 80%, respectively. However, the clinical utility of sole measurement of IL-6 for NASH diagnosis is probably of little value because of the discrepancies above mentioned.

Concerning insulin resistance which characterizes NASH (163), Shimada et al conducted accuracy analyses of homeostasis model assessment of insulin resistance (HOMA-IR). In accordance with the fact that HOMA-IR could be normal in the early stage NASH, they reported that HOMA-IR differentiated early stage NASH from NAFL with a sensitivity of 51% at a cut-off of 3 (specificity of 95%, PPV 98%, NPV 31%, and AUROC 0.76). In another study, HOMA-IR was found to be significantly associated with NASH and was an independent predictor (164). However, there was no baseline difference in HOMA-IR between normal versus NAFLD and NAFL versus NASH; only between normal subjects and NASH was significant difference.

High-sensitivity C-reactive protein (hsCRP) is an acute phase reactant which can detect lower grade inflammation. Yoneda et al was the first to show the usefulness of elevated hsCRP in distinguishing biopsy-proven NASH patients with nonprogressive steatosis subjects at an AUROC of 0.83 (165). However, the results were not reproduced by others (162, 166). In particular, Haukeland et al demonstrated that CC-chemokine ligand-2 (CCL2) but not hsCRP was elevated in NAFLD and was significantly higher in NASH than NAFL (162).

CCL2, also known as monocyte chemoattractrant protein-1 (MCP-1) is a potent chemokine which is responsible for hepatic recruitment of macrophages during liver inflammation (167). In other study of 104 subjects, high CCL2 level was associated with elevated ALT (168). In addition, CCL2 level was significantly higher in patients diagnosed with NAFLD by ultrasound. However, there are no studies which have validated CCL2 with liver biopsy and so further research is required before.

In a series of 70 patients with biopsy-proven NAFLD and 10 healthy controls, significantly higher pentraxin-3 level was found in NASH than non-NASH cases (169). The AUROC for separating NASH from

non-NASH with pentraxin-3 was 0.76. The sensitivity, specificity, PPV, and NPV were 66.7%, 78.6%, 82.4%, and 61.1%, respectively at the cut-off of 1.6 ng/mL. There is a possibility of utilization of pentaxin-3 for not only differentiating NASH from non-NASH but also assessing degree of fibrosis, in that there was a stepwise increase in the level of this marker according to the histological stage of fibrosis. However, because pentraxin-3 is primarily an acute phase reactant responding inflammation, the sole measurement of this marker may not represent underlying pathology.

Serum prolidase enzyme activity (SPEA) reflects hepatic prolidase enzyme activity (170). Kayadibi et al reported that SPEA was significantly elevated in patients with NASH than NAFL with an AUROC of 0.85, a sensitivity of 84%, a specificity of 82%, a PPV of 82%, and a NPV of 84% (cut off 1134 u/l) (171). Potential advantage is that SPEA could predict fibrosis as well as steatohepatitis. However, further investigation and validation is needed as for other biomarkers.

Soluble receptor for advanced glycation endproducts (sRAGE) has been known to be associated with some components of metabolic syndrome (172, 173). A case control study involving 57 NAFLD patients and 14 healthy controls showed significantly decreased level of sRAGE in NASH group (174). In differentiating NASH from NAFL, the AUROC of sRAGE was 0.77. The sensitivity was 75.0% and specificity was 71.4% at a cut-off of 1309 pg/mL. Although the level of sRAGE might be decreased in NASH, it is not unique to NASH (175). Hence, they would possibly be useful when added to NASH diagnostic panels after further investigations.

Oxidative stress has been recognized as an important mechanism in the pathogenesis of NASH. Markers from different oxidation pathways were investigated for use in NASH diagnosis but failed to show solid and consistent results (176-179). In addition, the serum or plasma measurement of oxidative markers may not necessarily reflect the activity of different oxidation pathways in the liver. Therefore, the use of oxidative stress markers in clinical practice is still questionable.

3.3.2 Clinical Models

A thorough medical history to assess for metabolic syndrome risk factors and exclude alcohol and secondary causes of fatty liver is crucial in establishing NAFLD. However symptoms are not helpful in discerning which patients have NASH, as there is usually an absence until a considerable degree of cirrhosis develops (180). As for physical examination, a specific pattern of fat distribution, dorsocervical lipohypertrophy, was shown to be associated with severity of steatohepatitis but this sign is non-specific and consistent recognition can be difficult when it is subtle (181). In addition, the performance of routine laboratory parameter has not reached satisfactory levels of sensitivity and specificity (182).

Diagnostic performance can be improved when clinical and laboratory parameters are incorporated into prediction models (***Table 10)***. Poynard's NashTest consists of 13 parameters including some metabolic factors (183). From a cohort of patients diagnosed with NAFLD via the SteatoTest (also developed by Ponyard), NashTest, was assessed in its ability to differentiate NASH from simple steatosis. A specificity of 94% but the sensitivity only reached 33%. A later attempt was performed to validate this test in another French cohort (184). However, there were only 15 NashTest-positive cases and 19 biopsy-confirmed NASH among more than 250 patients, hence further study is warranted.

NASH Diagnostics, which incorporates cleaved cytokeratin-18 (CK-18), intact CK-18 minus cleaved CK-18, adiponectin, and resistin yielded a sensitivity of 72.1%, specificity of 91.4%, and overall area under the receiver-operating curve (AUROC) of 0.85 (137). A later study conducted by the same group however, demonstrated lower AUROC of 0.70 for the same panel (185). In that study, the authors newly constructed a model called the NASH model as a part of the NAFLD Diagnostic Panel. It consists of 6 clinical or apoptosis- and necrosis-related parameters: type 2 diabetes mellitus, gender (male being negative impact), body mass index, triglyceride, cleaved CK-18 and CK-18 minus cleaved CK-18. In set of 79 NAFLD patients, the authors found AUROC of 0.81, which was superior to the NashTest AUROC of 0.70. The discrepancy in results along with small sample sizes calls for external validation.

The Nice Model is a scoring system incorporating 3 independent variables which predict non-alcoholic fatty liver disease activity score (NAS) ≥5: alanine aminotransferase (ALT), CK-18, and the presence of metabolic syndrome (186). Using ALT, CK-18, and the presence of metabolic syndrome alone, an AUROC of 0.78, 0.74, and 0.74 was obtained respectively for detection of definitive NASH. Combining these 3 variables increased AUROC to 0.88 in the training group and 0.83 in the validation group. The reported sensitivity of logarithmic transformation of this scoring system was 84%, with a specificity of 86% and negative predictive value (NPV) of 98%. Yet, the positive predictive value (PPV) of this model is quite low.

OxNASH is a risk score model which incorporates 13-hydroxyl-oactadecadienoic acid (13-HODE)/linoleic acid (LA) ratio, age, body mass index (BMI), and aspartate aminotransferase (AST) (187). In addition to the variables which were included in other models such as age, BMI, and AST, the rationale for oxNASH in clinical diagnosis of NASH is based on the finding that oxidative stress is an important mechanism of

pathogenesis in NAFLD (188). Although this model showed an acceptable AUROC, it has not been externally validated and blood markers for oxidation products are not easy to perform in most centers.

HAIR (hypertension, increased ALT, and insulin resistance) had been introduced in 2001 and its performance characteristics for NASH was relatively high (189). However, this scoring system included highly selective patients who were suffering from severe obesity (BMI > 35 kg/m^2) and to date, no external validations have been carried out.

Table 10: Clinical models for predicting NASH

Study	Name	Component/Formula	Study Population	Results	Comment
Poynard et al, 2006 [183]	NashTest	1. Age 2. Sex 3. Height 4. Weight 5. Triglyceride 6. Cholesterol 7. α2-MG 8. Apolipoprotein A1 9. Haptoglobin 10. GGT 11. ALT 12. AST 13. Total bilirubin - undisclosed formula	160 - training group 97 - validation group 383 - controls	AUROC 0.79 Se 33%, Sp 94% PPV 66%, NPV 81%	Validated in 274 patients with morbid obesity - Se 21%, Sp 96%, PPV 27%, NPV 94% (calculated)
Younossi et al, 2008 [137]	NASH Diagnostics	1. Cleaved CK-18 2. CK-18 minus cleaved CK-18 3. Adiponectin 4. Resistin - undisclosed formula	69 - training group 32 - validation group	AUROC 0.85 Se 72%, Sp 91% (threshold 0.4320)	Reevaluated in 79 patients by same group - AUROC 0.70, Se 61%, Sp 69%, PPV 68%, NPV 63% (threshold 0.389)
Younossi et al, 2011 [185]	NASH Model of NAFLD Diagnostic Panel	1. Type 2 diabetes mellitus 2. Gender 3. BMI 4. Triglyceride 5. Cleaved CK-18	79 NAFLD patients	AUROC 0.81 Se 91%, Sp 47%, PPV 61%, NPV 86% (threshold 0.2210) Se 44%, Sp 92%, PPV 83%, NPV 65% (threshold 0.6183)	

		6. CK-18 minus cleaved CK-18			
Anty et al, 2010 [186]	Nice Model	1. ALT 2. CK-18 3. Metabolic syndrome	464 morbidly obese patients 310 - training group 154 - validation group	AUROC 0.83-0.88 Se 84%, Sp 86%, PPV 44%, NPV 98% (logarithmic transformation, threshold 0.1400)	Model = -5.654 + 3.780E-02 x ALT x 2.215E-03 x CK-18 = 1.825 x (presence of metabolic syndrome = 1) Logarithmic transformation = 1/1(1+EXP(-Nice Model))
Feldstein et al, 2010 [187]	oxNASH	1. 13-HODE/LA ratio 2. Age 3. BMI 4. AST	73 - training group 49 - validation group	AUROC 0.74-0.83 Se 81-84% (threshold 55) Sp 63-97% (threshold 73)	Model = 100 x exp(z)/{(1+exp(z)) z = -10.051 + 0.0463 x age (years) + 0.147 x BMI + 0.0293 x AST + 2.658 x 13-HODE/LA ratio
Dixon et al, 2001[189]	HAIR	1. Hypertension 2. (increased) ALT 3. IR	105 morbidly obese patients	AUROC 0.90 Se 80%, Sp 89% (threshold 2)	Hypertension = 1 ALT > 40 IU/L = 1 IR index > 5.0 = 1

Legend: NASH, non-alcoholic steatohepatitis; α2-MG, alpha2 macroglobulin; GGT, gamma-glutamyl-transpeptidase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUROC, area under the receiver-operating curve; Se, sensitivity; Sp, specificity; PPV, positive predictive value; NPV, negative predictive value; CK-18, cytokeratin-18; BMI, body mass index; NAFLD, non-alcoholic fatty liver disease; 13-HODE, 13-hydroxyl-oactadecadienoic acid; LA, linoleic acid; IR, insulin resistance

3.4 NON-INVASIVE DIAGNOSIS OF FIBROSIS AND CIRRHOSIS

Fibrosis and cirrhosis is the common pathway of chronic liver diseases. Fibrosis is a natural response to tissue injury. With ongoing liver injury, there is accumulation of fibrous tissue. Eventually, the liver architecture is disrupted, and multiple nodules are formed and separated by thick fibrous septa. This marks the development of cirrhosis. A number of serious complications can occur with the onset of cirrhosis. Although HCC has been reported in patients with non-cirrhotic NAFLD (190, 191) cirrhosis is still the most important risk factor of HCC (58, 59). Other complications include ascites, spontaneous bacterial peritonitis, variceal bleeding, hepatic encephalopathy and hepatorenal syndrome. Therefore, it is important to stage the degree of fibrosis and cirrhosis. Liver biopsy is the reference standard for determining the stage of fibrosis, but due to its limitations (as described in chapter 1), non-invasive assessment methods need to be developed.

3.4.1 Biomarkers and prediction scores

Biomarkers of fibrosis are divided into 2 types. Class I biomarkers measure fibrogenesis and fibrinolysis directly. Class II biomarkers do not measure fibrosis directly but are clinical parameters associated with fibrosis. For example, patients with higher aminotransferases are more likely to have active disease and therefore fibrosis, but aminotransferases are not a measurement of fibrosis and the association is not absolute (192). Moreover, it is important to note that fibrosis and cirrhosis are the results of years of disease activity. Thus, a single-time measurement of markers of disease activity would not have good correlation with the severity of fibrosis. In fact, when NAFLD reaches the stage of cirrhosis, steatosis and necroinflammation typically regress (193). NASH is currently believed to be the most important aetiology underlying cryptogenic cirrhosis (194, 195).

As none of the available biomarkers has sufficient accuracy in diagnosing fibrosis as a standalone test, there have been a number of prediction scores ***(Table 11)***. In general, the scores were derived using liver histology as the reference standard. Clinical parameters and biomarkers associated with different fibrosis stages were identified, and a score was constructed based on the relative importance of each factor. Some of the scores were developed and validated in NAFLD patients only, while the majority were first developed for patients with other liver diseases such as chronic hepatitis C and later adopted for NAFLD.

The NAFLD fibrosis score is one of the most extensively tested prediction scores (196). It comprises age, hyperglycemia, body mass index (BMI), platelet count, albumin, aspartate aminotransferase (AST) and alanine aminotransferase (ALT). The score was derived from 480 patients in the training cohort and further tested in 253 patients in the validation cohort. Using a pair of high and low cutoffs, the score had 82% positive predictive value and 88% negative predictive value in diagnosing F3 disease. Around 30% of patients had score between the 2 cutoffs and thus indeterminate results. The latest AASLD guideline supports the use the NAFLD fibrosis score to risk stratify NAFLD patients (97). Since 90% of the original

cohort for the development of the NAFLD fibrosis score were Caucasians (196), the score has been independently validated in the Chinese population. NAFLD Fibrosis score was found to still have a high negative predictive value of 91% in Chinese, but few patients had high scores suggestive of advanced fibrosis. The phenomenon may be partly because Asian patients tend to develop metabolic complications at a lower BMI (197).

Other scores have not been as extensively studied, but the FIB-4 index appears to have the highest accuracy in diagnosing fibrosis in NAFLD patients when compared to other prediction scores. The FIB-4 index comprises age, platelet count, AST and ALT. In 3 separate validation studies in America, Europe and Asia, the FIB-4 index had an area under the receiver-operating characteristics curve of over 0.80 in diagnosing F3-4 disease (127, 198, 199). The components and performance of other prediction scores are shown in ***Table 11***.

It is important to note that the prediction scores were validated against liver histology. Since liver histology is an imperfect reference standard with sampling variability, intraobserver and interobserver bias, there is a ceiling for the perceived accuracy in such validation studies (200). In other words, even if a score has 100% accuracy, assuming the accuracy of liver biopsy is 90%, the score will still disagree with histology in 10% of cases and classified as inaccurate results. In reality, however, the prediction scores are modeled against histology and therefore would suffer from a similar degree of case misclassification.

Table 11: Biomarkers and prediction scores of liver fibrosis in NAFLD

			F2		F3	
Score	Components	Class I or II biomarkers	Sensitivity	Specificity	Sensitivity	Specificity
Specific for NAFLD						
NAFLD fibrosis score[196]	Age, hyperglycemia, BMI, platelet, albumin, AST/ALT ratio (dual cutoffs)	II	-	-	0.77	0.96
BARD score [201]	BMI, AST/ALT ratio, diabetes	II	-	-	0.62	0.66
FibroMeter NAFLD [202]	Glucose, AST, ferritin, platelet, ALT, body weight, age	II	0.79	0.96	-	-
Not specific for NAFLD						
AST/ALT ratio [203]	AST, ALT	II	-	-	0.21	0.90
APRI [204]	AST, platelets (dual cutoffs)	II	-	-	0.65	0.97
ELF [205]	Hyaluronic acid, TIMP1, PIIINP (dual cutoffs)	I	0.80	0.67	0.80	0.90
FIB-4 [206]	Age, AST, platelet, ALT (dual cutoffs)	II	-	-	0.74	0.98
FibroTest [207]	Total bilirubin, GGT, α_2-macroglobulin, ApoA1, haptoglobin (dual cutoffs)	I and II	0.71	0.98	0.88	0.99
Hepascore[208]	Age, gender, bilirubin, GGT, hyaluronic acid, α-2 macroglobulin	I and II	0.51	0.88	0.76	0.84

Legend: ALT, alanine aminotransferase; ApoA1, apolipoprotein A1; APRI, AST-to-platelet ratio index; AST, aspartate aminotransferase; BMI, body mass index; ELF, enhanced liver fibrosis panel; GGT, gamma-glutamyl transpeptidase; NAFLD, non-alcoholic fatty liver disease; PIIINP, procollagen III amino-terminal peptide; TIMP1, tissue inhibitor of matrix metalloproteinase 1

3.4.2 Physical measurements

3.4.2.1 Ultrasound, computed tomography and magnetic resonance imaging

Ultrasound is the most commonly performed imaging test in patients with liver disease. A recent meta-analysis found that ultrasound is able to diagnose NAFLD when hepatic steatosis exceeds 33% at a good accuracy (84.8% sensitivity and 93.4% specificity) (209). The drawbacks are that it is a qualitative measure and its lack of ability to detect minor steatosis. Also, it is affected by intraobserver and interobserver variability (kappa= 0.54-0.92, 0.44-1.00) and is unable to distinguish NASH from simple steatosis (209, 210). Cirrhosis can be diagnosed in advanced cases when the liver is small and shrunken, or when there are signs of portal hypertension such as ascites, splenomegaly, varices and recanalization of the umbilical vein. However, the diagnosis can be difficult in early cirrhosis when signs of portal hypertension are absent. It follows that fibrosis is certainly impossible to assess with ultrasound (210). Furthermore, hepatomegaly and increased liver echogenicity in patients with NAFLD would make ultrasonographic features of cirrhosis inconspicuous. As a result there have been various attempts to develop ultrasound quantitative measures based on the greater echogenicity of the liver in NAFLD compared to other organs. The Ultrasonographic Fatty Liver Indicator (US-FLI) and the Hepato-Renal index are two such methods (211, 212) but require further evaluation as only small studies have been performed.

Computed tomography is superior to ultrasound in detecting focal steatosis, but otherwise has a similar diagnostic performance to ultrasound. CTs should be non-contrast because contrast affects the attenuation of the liver causing different thresholds. It is accurate in diagnosing hepatic steatosis that is at least moderate in severity (82% sensitivity, 100% specificity) (213). However, CT misdiagnoses fatty liver when there are other diffuse liver conditions such as haemachromatosis (214). Although CT can evaluate nodular liver and other features such as ascites and varices that may suggest cirrhosis, it cannot assess early cirrhosis or fibrosis, and it also cannot distinguish NASH from simple steatosis (215). There is also the additional drawback of radiation exposure. Thus it is not the modality of choice for routine diagnosis given the high prevalence of NAFLD.

In prospective studies using liver biopsy as the gold standard, conventional MRI performed better than ultrasound in detecting minor steatosis (216), but is poor at diagnosing NASH and assessing fibrosis (210). Many varieties of MRI technique have been developed to improve its performance in the diagnostic spectrum of NAFLD. Magnetic resonance spectroscopy (MRS) is emerging as a very promising modality. This directly measures the signal from hydrogen atoms and can distinguish between its different molecular bonds. The spectra pertaining to methyl groups in triglyceride molecules can be detected. Hence MRS is able to directly diagnose hepatic triglycerides, and can also quantify its content (HTC). MRS shows good diagnostic accuracy for all grades of steatosis (AUROC 0.87-0.89) (217). MRS also has the advantages of being able to assess the entire volume of liver. As more refined software and technique algorithms are being developed, it is challenging liver biopsy as a possible new gold standard in diagnosing steatosis (218).

Magnetic resonance elastography (MRE) is an MRI modality with promising results for diagnosing fibrosis. MRE is phase-contrast-based MRI technique that produces an image of a propagating shear wave. In the Mayo clinic protocol, a constant mechanical wave is produced from a disc shaped driver that is attached to the patient's anterior right chest wall (219). The data acquired allows the MRI to generate an image map of the liver that depicts the quantitative tissue elasticity. Early studies of MRE suggest that it is superior to TE in diagnosing each stage of fibrosis (220) and has good accuracy for diagnosing NASH (221). The disadvantages of MRI techniques are that they are expensive and not widely available. Further external validation is also required.

3.4.2.2 Transient elastography

TE (Fibroscan, Echosens Paris France) enables non-invasive assessment of liver fibrosis using ultrasonic elastography principles. The Fibroscan probe consists of an ultrasound fitted on the axis of an electrodynamic transducer. The probe is placed on the skin overlying the liver, and generates a low-amplitude 50Hz mechanical pulse which creates a shear wave. The velocity of the shear wave is directly related with the stiffness of the liver. Ultrasound signals at low energy 3.5MHz emitted from the probe measure the shear wave velocity and can directly calculate the elastic modulus. This is expressed in kilopascals and is known as liver stiffness measurement (LSM). TE has been validated as a measure of fibrosis across a wide spectrum of chronic liver disease and has overall a good accuracy. It has the advantage of being quick, easy to learn, well tolerated by patients, and assesses a volume of liver around 100-200 times the size of a liver biopsy. There have been many studies examining its use in NAFLD patients, and there is ongoing debate regarding its diagnostic accuracy and feasibility especially in obese patients.

Meta-analysis on TE

Nine studies including a total pool of 1047 NAFLD patients from different ethnic backgrounds were identified as suitable for meta-analysis ***(Table 15/Supplementary Table 4)***. Data on M probe included 854 NAFLD patients. Data was grouped according to whether the M probe or the XL probe was used, and then further sub-grouped according to the fibrosis stage that it was being compared. Eight studies had suitable data for the M probe, whereas 1 study had suitable data only for the XL probe. There were 7, 8 and 6 TE studies reported that its performance compared to liver biopsy for F≥2, 3 and 4 respectively ***(Figure 16)***. For F≥2, the LSM cut-off ranged from 6.7-7.7 kPa, with 67-94% sensitivity, 61-84% specificity and AUROC 0.79-0.87. For F≥3, the LSM cut-off was 8.0-10.4 kPa, with 65-100% sensitivity, 75-97% specificity and AUROC 0.76-0.98. For F4, the LSM cut-off was 10.3-17.5 kPa, with 78-100% sensitivity, 82-98% specificity and AUROC 0.91-0.99. In the pooled estimates of diagnostic accuracy TE had overall for F≥2 79% sensitivity, 75% specificity; F≥3 85% sensitivity, 85% specificity and F4 92% sensitivity, 92% specificity ***(Table 14/Supplementary Table 3)***. ***Figure 20/Supplementary Figure 4*** showed the HSROC plots of TE.

Discussion on TE

The overall results suggest that TE is excellent in diagnosing F $\geq$ 3 (85% sensitivity, 82% specificity) and F4 (92% sensitivity, 92% specificity) and has moderate accuracy for F $\geq$ 2 (79% sensitivity, 75%

specificity). Our analysis of 854 NAFLD patients in eight studies is the largest so far and most updated. The quality of data in the included studies was excellent, with all studies obtaining at least 12/13 on the modified QUADAS ***(Figure 17/Figure S1)***, and hence no subgroup analysis between high- and low-quality studies was performed. In addition, analysis of whether BMI and ALT was a factor in heterogeneity could not be performed because of wide range of these factors in each of the included studies.

Obesity is the main reason for failed LSM, and the problem can be largely overcome using the XL probe (222, 223). The largest study of 193 patients reported the ability to obtain 10 measurements in 93% of patients with BMI > 30kg/m^2 with AUROCS of 0.80, 0.85 and 0.91 for F≥2,3 and 4 respectively, although lower LSM cut-offs need to be used (129). Pooled statistical analysis could not be performed for the XL probe performance due to insufficient number of studies (224, 225). All TE studies had similar baseline characteristics, used similar cut-offs and there were no heterogeneity factors identified. TE studies had high quality data and subgroups and post-hoc sensitivity analysis did not show that this affected the overall summary estimates.

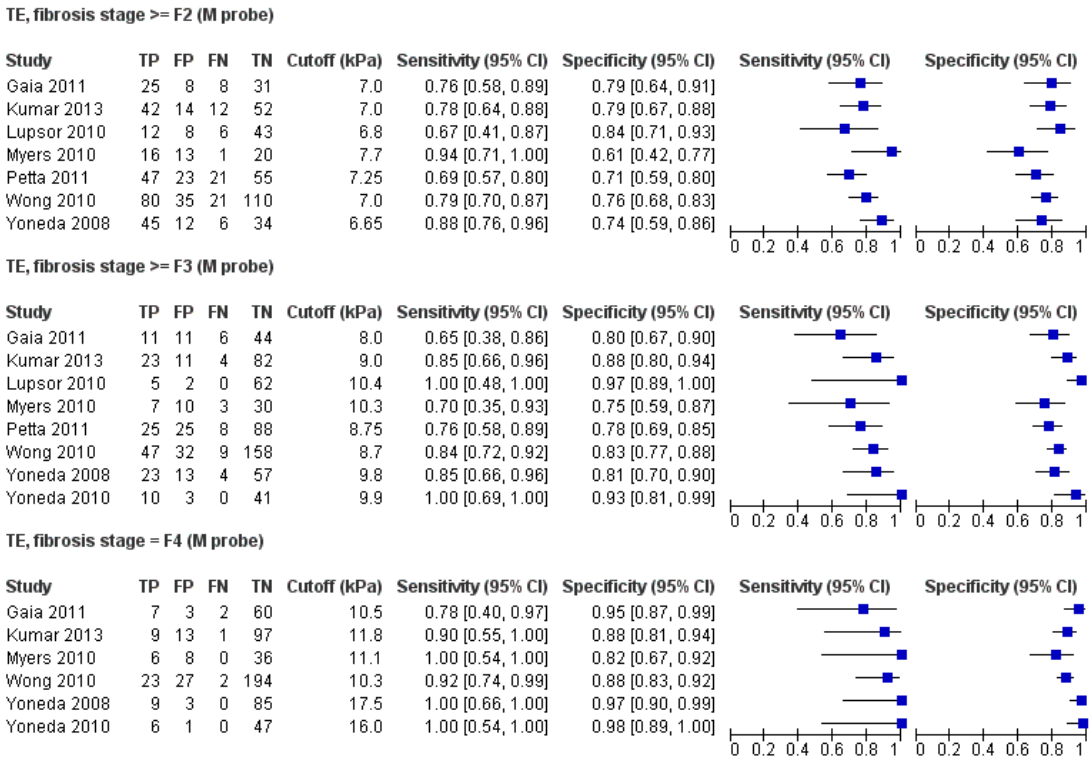


Figure 16: Forest plot from meta-analysis for Transient Elastography

Forest plot from meta-analysis of sensitivities and specificities for TE to diagnose different fibrosis stages using a random-effect model. TP:True Positives, FP: False Positives, FN: False Negatives, TN: True Negatives.

3.4.2.3 Acoustic radiation force impulse (ARFI)

Acoustic radiation force impulse imaging (ARFI) is a form of tissue elastography that is integrated into in a conventional high end ultrasound machine (Siemens S2000). A region-of-interest (ROI) in the liver is targeted using short-duration acoustic pulses with a fixed frequency of 2.67 MHz Shear-waves are generated away from the region of excitation that are tracked using an ultrasonic, correlation-based method. The shear wave speed of the tissue within a ROI is measured and can be used to calculate the elasticity of the liver. Like TE, the result is expressed in kilopascals. ARFI has the advantage of being a feature existing on an ultrasonography machine. This allows for the convenience of assessing for structural abnormalities, steatosis as well as fibrosis in a single sitting.

Summary estimates for ARFI were not possible in this review due to insufficient data being available. Only 2 studies fit our selection criteria (128, 145) , although a further 2 articles (226, 227) could have been included if attempts to contact study authors were successful. The AUROCS reported in our candidate studies (128, 145, 226, 227) ranged from 0.74-0.97 for the diagnosis of F≥3 in NAFLD. From a recent meta-analysis on the performance of ARFI across a heterogeneous range of liver disease, the mean AUROCS were 0.87, 0.91 and 0.93 for the diagnosis of F≥2, 3 and 4 respectively [228]. ARFI appears to be is a promising modality for NAFLD, but availability of this feature on ultrasound devices is currently limited.

3.4.2.4 Liver scintigraphy

Technetium-99 m-2-methoxy-isobutyl-isonitrile (Tc 99-MIBI) is a lipophilic cationic agent that was initially designed for myocardial perfusion imaging utilizing the property of Tc99-MIBI uptake and retention being related to mitochondrial function. In NASH, the precise mechanism is unclear, but it has been observed that the liver: heart ratio and the liver: spleen ratio uptake of Tc99-MIBI is decreased in NASH compared to simple steatosis (229, 230). Further studies are needed.

3.5 CONCLUSION

NAFLD is a disease that affects 15-40% of the general population. Accurate identification of patients with progressive or advanced disease is one of the most urgent clinical needs. At present, serum tests and physical measurements such as TE come close as highly accurate non-invasive tests to exclude advanced fibrosis and cirrhosis in NAFLD patients. CK18 has moderate accuracy in diagnosing NASH while other biomarkers have not been extensively studied. Further studies are needed to explore the optimal test combinations and the role of these tests in prognostication and treatment monitoring.

3.6 Supplementary Material

Table 12: Modified QUADAS

Supplementary Table 1. Modified Quality assessment of Diagnostic Accuracy Studies Checklist (Modified QUADAS)

1. Was the spectrum of patients representative of the patients who will receive the test in practice? (Generalizability item)
2. Were selection criteria clearly described? (Clarity item)
3. Is the reference standard likely to correctly classify the target condition? (Validity item)
4. Is the time between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests? (Validity item)
5. Did the whole sample or a random selection of the sample receive verification using a reference standard of diagnosis? (Validity item)
6. Did patients receive the same reference standard regardless of the index test result? (Validity item)
7. Was the reference standard independent of the index test (ie, the index test did not form part of the reference standard)? (Validity item)
8. Was the execution of the index test described in sufficient detail to permit replication of the test? (Clarity item)
9. Was the execution of the reference standard described in sufficient detail to permit its replication? (Clarity item)
10. Were the reference standard results interpreted without knowledge of the results of the index test? (Validity item)
11. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice? (Validity item)
12. Were uninterpretable/intermediate test results reported? (Clarity item)
13. Were withdrawals from the study explained? (Clarity item)

Table 13: Characteristics of studies on cytokeratin-18 fragments

Supplementary Table 2. Characteristics of studies on cytokeratin-18 fragments

Study	Location	Patients (n)	Age (yrs)	Gender (%male)	BMI (kg/m^2)	ALT (U/L)	%NASH (n)
Wieckowska 2006	Cleveland USA	39	50.8±11.1	46.1	31.5±4.0	73	53.8 (21)
Yilmaz 2006	Bursa Turkey	83	48.9±9.1	54.2	30.3±4.8	42	54.2 (45)
Diab 2008	Cleveland USA	86	51.0	20.9	48	21.5	25.6 (22)
Younassi 2008[a]	Inova Fairfax USA	69	37.4±8.3, 42.5±10.4, 39.3±9.8	6.7, 40.1, 9.4	45.7±4.8, 48.2±8.7 47.0±9.1	22.1±12.2, 47.9±32.1 21.9±8.1	31.9 (22)
Feldstein 2009	Cleveland USA	139	48.0	26.7	34.2	66.0	49.6 (69)
Papatheodoridis 2010[c]	Athens Greece	58	47±16 47±12	59, 47	28±5 30±4	79, 76	51.7 (30)
Musso 2011	Turin, Italy	41	37±2[d]	70[d]	25.4±0.5[d]	70±5[d]	39.0 (16)
Grigorescu 2012[b]	Cluj-Napoca Romania	79	39.1±10.7 48.3±11.4	70, 71.2	28.6±3.8 30.5±3.8	48.6±26.2 86.3±49.0	74.7 (59)
Joka 2012[b]	Hannover Germany	22	49.9±3.2 45.6±3.3	70, 66.7	26.0±0.9, 27.8±1.1	52.8±8.4 94.4±10.4	45.5 (10)
Pirvelescue 2012[b]	Bucharest, Romania	60	45.9±10.6 44.9±9.4	29.8, 30.8	39.6±11 49.4±7.6	21.3±11.8 42.3±15.2	21.7 (13)
Shen 2012	Hong Kong China	146	48.1±9.7	66.6	27.4±3.9	71±42	56.2 (82)

NA: Not Applicable

Variables with ± represent mean±standard deviation. Variables without ± indicate that it is the median value for that variable

a. Variables reported in subgroup order: SS, NASH, controls
b. Variables reported in subgroup order:Non NASH, NASH
c. Variables reported in subgroup order: M30<250, M30≥250
d. Overall data for NAFLD patients not available. Data for the largest subgroup of (14/16) NASH patients with AA allele for LOX-1 IVS4-14 A→6 gene reported

Table 14: Overall and subgroup analyses transient elastography and cytokeratin-18 fragments

Supplementary Table 3. Overall and subgroup analyses transient elastography and cytokeratin-18 fragments[†]

	Number of studies	Sensitivity (95% CI)	Specificity (95% CI)
Transient elastography (M probe)			
Overall			
≥ F2	7	0.79 (0.72 to 0.84)	0.75 (0.71 to 0.79)
≥ F3	8	0.85 (0.73 to 0.92)	0.85 (0.79 to 0.90)
= F4	6	0.92 (0.82 to 0.97)	0.92 (0.86 to 0.96)
Adequate index test description (Yes)			
≥ F2	6	0.79 (0.73 to 0.85)	0.74 (0.70 to 0.78)
≥ F3	7	0.82 (0.73 to 0.88)	0.83 (0.78 to 0.87)
Adequate reference test description (Yes)			
≥ F2	6	0.77 (0.72 to 0.82)	0.76 (0.72 to 0.80)
≥ F3	7	0.86 (0.74 to 0.93)	0.86 (0.80 to 0.91)
= F4	5	0.92 (0.80 to 0.97)	0.93 (0.88 to 0.96)
Cytokeratin-18 (M30)			
Overall			
The best overall	7	0.66 (0.59 to 0.72)	0.82 (0.69 to 0.90)
High sensitivity	6	0.82 (0.75 to 0.87)	0.65 (0.43 to 0.82)
High specificity	6	0.58 (0.38 to 0.76)	0.98 (0.89 to 0.97)
Acceptable delay between tests (Yes)			
The best overall	4	0.65 (0.57 to	0.74 (0.59 to

		0.73)	0.85)
High sensitivity	4	0.83 (0.74 to 0.90)	0.74 (0.48 to 0.90)
Adequate index test description (Yes)			
The best overall	6	0.66 (0.59 to 0.73)	0.81 (0.65 to 0.91)
High sensitivity	5	0.81 (0.73 to 0.87)	0.71 (0.51 to 0.85)
High specificity	5	0.62 (0.40 to 0.81)	0.98 (0.83 to 1.00)
Adequate reference test description (Unclear)			
The best overall	4	0.69 (0.60 to 0.78)	0.76 (0.64 to 0.85)
High sensitivity	4	0.85 (0.77 to 0.90)	0.58 (0.28 to 0.84)
High specificity	4	0.54 (0.28 to 0.78)	0.94 (0.89 to 0.97)
Blinding for index test results (Yes)			
The best overall	4	0.63 (0.55 to 0.71)	0.85 (0.62 to 0.95)

[†] We were unable to provide pooled estimates of diagnostic accuracy for specific patient subgroups because no bivariate analyses could be fitted on less than four studies or 2 x 2 data contained one or more zero values.

Table 15: Characteristics of 9 studies on transient elastography

Supplementary Table 4. Characteristics of 9 studies on transient elastography

Study	Location	NAFLD patients	Age (yrs)	Gender (%male)	BMI (kg/m^2)	ALT (U/L)
Yoneda 2008	Yokohama Japan	97	51.8±13.7	41.2	26.6±4.2	80±62.3
Lupsor 2010	Cluj-Napoca, Romania	69	42[a]	70.8	28.71[a]	80[a]
Wong 2010	Pessac, France and Hong Kong, China	246	51±11	54.9	28.0±4.5	75±54
Yoneda 2010	Yokohama, Japan	54	M:48.3±13.5 F:52.5±11.4	46.3	M:28.2±5.0 F:26.2±4.4	M:66.4±29.1 F:54.9±33.1
Myers 2010	Multicentre, Canada	50	49[ab]	66.2	26[ab]	61[ab]
Petta 2011	Palermo, Italy	146	44.1±13.2	71.2	29.1±4.1	80.9±57.8
Gaia 2011	Turin, Italy	72	48[a]	72.2	27.5[a]	58[a]
Kumar 2013	New Dehli, India	120	39.1±12.8	75.0	26.1±3.6	62.5
Wong 2012[c]	Pessac, France and Hong Kong, China	193	52±11	57.0	28.9±4.8	73±76

Unless stated, variables with ± represent mean and standard deviation

a. Median Value reported
b. Values refer to entire cohort of chronic liver disease patients. Specific values for NAFLD patients not reported.
c. Data for XL probe included only.

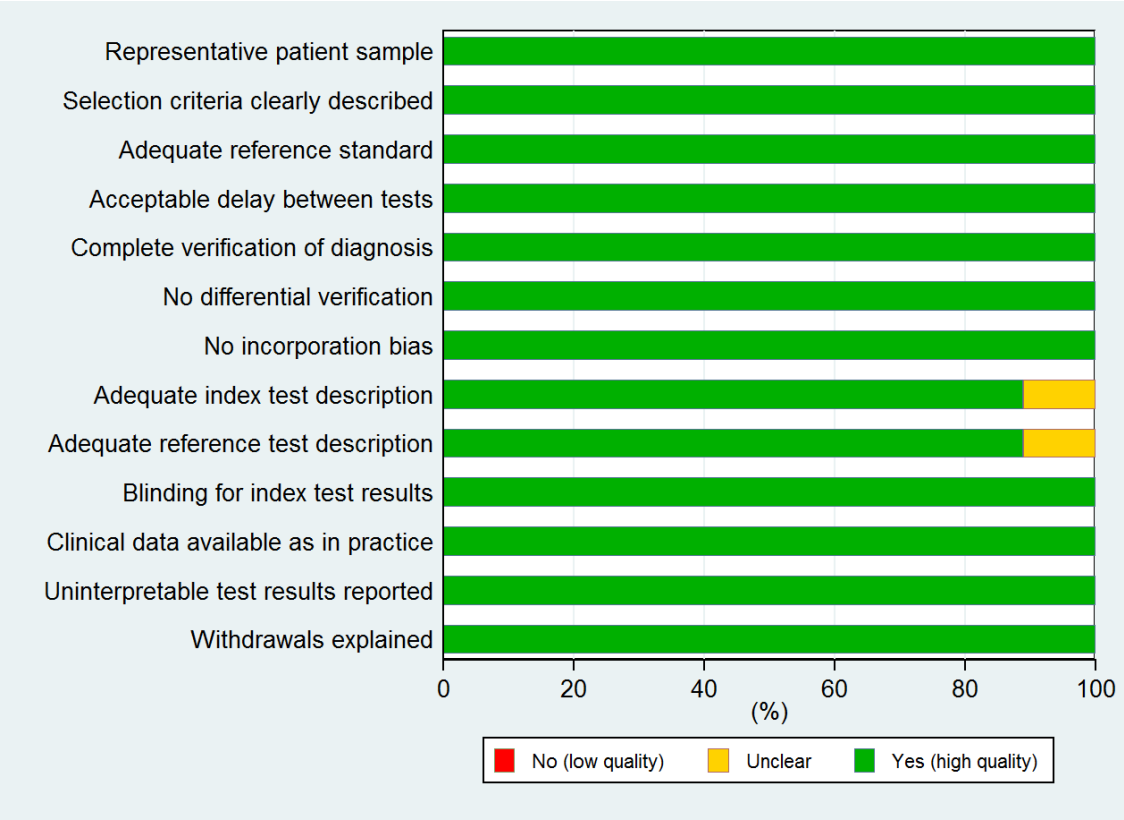


Figure 17: QUADAS assessment of 9 studies on transient elastography

Supplementary Figure 1. QUADAS assessment of 9 studies on transient elastography

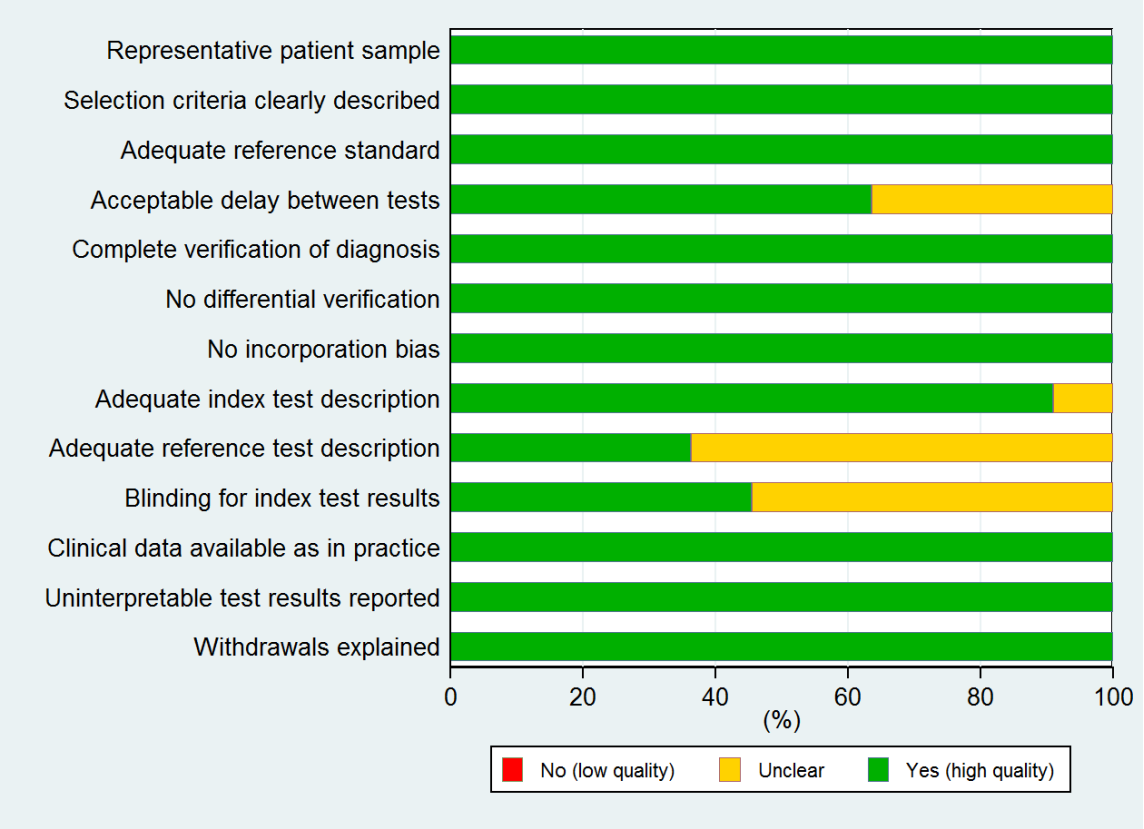
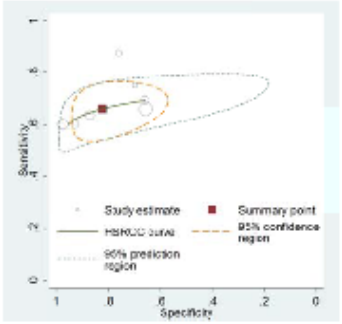


Figure 18: QUADAS assessment of 11 studies on cytokeratin-18 fragments

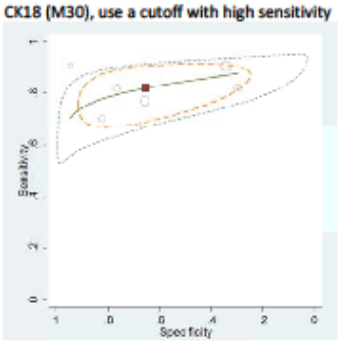
Supplementary Figure 2. QUADAS assessment of 11 studies on cytokeratin-18 fragments

Supplementary Figure 3. Hierarchical summary receiver-operating characteristic (HSROC) graphs with 95% confidence region and 95% prediction region for CK18 (M30). The size of the circles is proportional to the number of patients included in the study.

CK18 (M30), use a cutoff with the best overall sensitivity and specificity



CK18 (M30), use a cutoff with high sensitivity



CK18 (M30), use a cutoff with high specificity

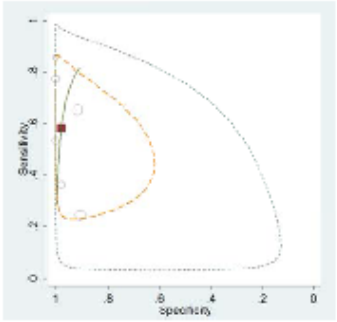
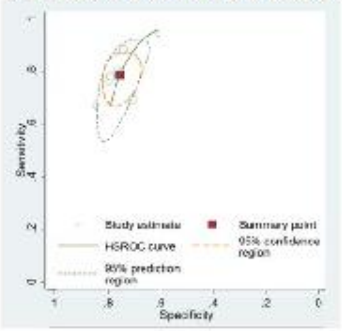


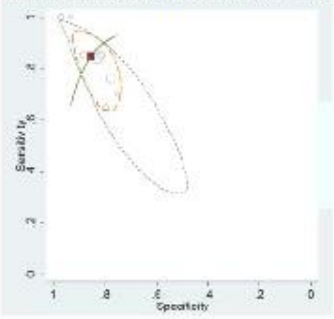
Figure 19: HSROC graphs for CK18

Supplementary Figure 4. Hierarchical summary receiver-operating characteristic (HSROC) graphs with 95% confidence region and 95% prediction region for transient elastography (M probe), fibrosis stages ≥ F2, ≥ F3 and = F4. The size of the circles is proportional to the number of patients included in the study.

Transient elastography (M probe), fibrosis stage ≥ F2



Transient elastography (M probe), fibrosis stage ≥ F3



Transient elastography (M probe), fibrosis stage = F4

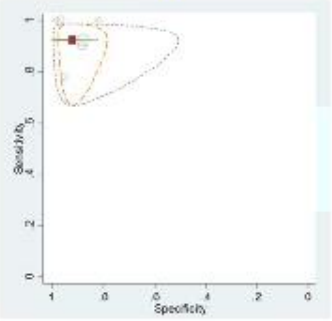


Figure 20: HSROC graphs for TE

This page has been left intentionally blank

CHAPTER 4: SCREENING DIABETIC PATIENTS FOR NON-ALCOHOLIC FATTY LIVER DISEASE WITH CONTROLLED ATTENUATION PARAMETER AND LIVER STIFFNESS MEASUREMENTS: A PROSPECTIVE COHORT STUDY

This chapter has been published as:

Screening diabetic patients for non-alcoholic fatty liver disease with controlled attenuation parameter and liver stiffness measurements: a prospective cohort study.

Kwok R, Choi KC, Wong GL, Zhang Y, Chan HL, Luk AO, Shu SS, Chan AW, Yeung MW, Chan JC, Kong AP, Wong VW

Gut 2016 Aug; 65(8):1359-68

Raymond Kwok is the corresponding author and was primarily responsible for recruiting subjects, performing the study, collecting data, interpreting the analysis and writing the drafts. The study was co-designed with the co-authors.

4.0 Chapter Summary

Objective: Type 2 diabetes is an important risk factor for non-alcoholic fatty liver disease (NAFLD), but current guidelines provide conflicting recommendations on whether diabetic patients should be screened for NAFLD. We therefore studied the strategy of screening diabetic patients by FibroScan.

Design: Liver steatosis and fibrosis were assessed by controlled attenuation parameter (CAP) and liver stiffness measurements (LSM) respectively using Fibroscan, at a diabetic center for patients from primary care and hospital clinics. Probe-specific LSM cutoffs were used to detect advanced fibrosis.

Results: Of 1918 patients examined, 1799 (93.8%) had valid CAP and 1884 (98.2%) had reliable LSM (1770 with the M probe and 114 with the XL probe). The proportion of patients with increased CAP and LSM was 72.8% (95% confidence interval 70.7-74.8%) and 17.7% (95% confidence interval 16.0-19.5%), respectively. By multivariable analysis, female gender, higher body mass index, triglycerides, fasting plasma glucose and alanine aminotransferase, and non-insulin use were associated with increased CAP. Longer duration of diabetes, higher body mass index, alanine aminotransferase, spot urine albumin-creatinine ratio, and lower high-density lipoprotein-cholesterol were associated with increased LSM. 94 patients with increased LSM underwent liver biopsy: 56% had steatohepatitis; 50% had F3-4 disease.

Conclusion: Diabetic patients have a high prevalence of NAFLD and advanced fibrosis. Those with high BMI and dyslipidaemia are at particularly high risk and should be the target for liver assessment. Our data supports screening for NAFLD and/or advanced fibrosis in patients with type 2 diabetes.

What is already known about this subject?

- Non-alcoholic fatty liver disease (NAFLD) is common in patients with type 2 diabetes.
- Diabetes is an important risk factor of steatohepatitis and cirrhosis in NAFLD patients.
- Current major hepatology organisations (AASLD, EASL and APASL) provide little guidance on NAFLD screening in diabetic patients because of the paucity of data.
- It is possible to measure liver fat and fibrosis quickly by Fibroscan. Its application as a screening tool in high-risk patients has not been systematically studied.

What are the new findings?

- Around 70% of diabetic patients from primary care and hospital clinics had increased controlled attenuation parameter suggestive of NAFLD.
- Around 18% of diabetic patients had increased liver stiffness of ≥9.6 kPa.
- Patients with high body mass index, dyslipidaemia and increased alanine aminotransferase were at highest risk of increased controlled attenuation parameter and liver stiffness.
- Among patients with increased liver stiffness, 56% had steatohepatitis, 21% had advanced fibrosis and 29% had cirrhosis.

How might it impact on clinical practice in the foreseeable future?

- There is a high prevalence of NAFLD and NAFLD with advanced fibrosis in patients with type 2 diabetes.
- Screening with Fibroscan is a convenient initial assessment for patients with type 2 diabetes. It may result in early detection of fibrosis and cirrhosis. However, its accuracy does not allow confident diagnosis of advanced disease. Further improvements in non-invasive assessment of liver fibrosis are needed.

4.1 INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) is the most common chronic liver disease, affecting 15–40% of the population worldwide (231). Of these NAFLD patients, 20-30% has non-alcoholic steatohepatitis (NASH), the progressive form of NAFLD which is associated with liver fibrosis. In 10-20% of NASH patients, this eventually progresses to cirrhosis and a high risk of hepatocellular carcinoma.(17, 58, 59)

Type 2 diabetes is a major risk factor for NAFLD. Both feature insulin resistance as a core component of their pathophysiology. As such, up to 90% of diabetic patients in some populations also have NAFLD.(232) However, due to insufficient data, current guidelines offer conflicting recommendations on whether diabetic patients should be screened for NAFLD.(97, 233) The main arguments against screening include uncertainties surrounding diagnostic tests and treatment options and lack of knowledge related to the long-term benefits of screening. On the other hand, screening may identify patients with NAFLD-related cirrhosis who would benefit from hepatocellular carcinoma and varices surveillance.

Knowledge on the epidemiology of NAFLD is incomplete because of the limitations of various diagnostic modalities. Liver biopsy is considered the reference standard, but is impractical to apply to a large study population. Conventional abdominal ultrasonography is easily accessible but is only qualitative. It is poor in detecting minor steatosis and suffers from intra/inter observer and variation.

Transient Elastography is a non-invasive test of liver fibrosis that is quick, easy to perform, and has a high degree of patient acceptance.(234) It has high accuracy and reproducibility when used to detect advanced fibrosis and cirrhosis. In addition, the latest model measures a novel physical parameter called the controlled attenuation parameter (CAP). Since fat affects ultrasound propagation, CAP measurement has been shown to be accurate in estimating the amount of liver fat.(235-237) Using this non-invasive technique, it is now possible to measure liver fat and fibrosis in a large number of patients.

In this study, we aim to test the strategy of NAFLD and fibrosis screening in patients with type 2 diabetes. We also studied factors associated with increased CAP and liver stiffness to guide selection of patients for screening.

4.2 METHODS

4.2.1 Subjects

This is a prospective cohort study. From March 2013 to May 2014, we screened 2466 consecutive patients aged 18 years or above with type 2 diabetes who attended comprehensive diabetic complications screening at the Diabetes Mellitus and Endocrine Centre Prince of Wales Hospital Hong Kong. The sources of referrals include hospital and primary care clinics in the New Territories East Cluster, which serves a population of 1.2 million. Subjects with active malignancy, positive hepatitis B surface antigen or antibody against hepatitis C virus, secondary causes of fatty liver (e.g. consumption of amiodarone and tamoxifen), and congestive hepatopathy were excluded. Men who consumed more than 20 g and women who consumed more than 10 g of alcohol per day were also excluded. All patients provided informed written consent. The study protocol was approved by the Clinical Research Ethics Committee of The Chinese University of Hong Kong. The patients would be prospectively followed for 10 years for hepatic, cardiovascular and metabolic complications. This paper reports the results of the baseline hepatic assessment.

4.2.2 Clinical assessment

The patients underwent a comprehensive 4-h assessment for diabetes-related complications and risk factors according to the European DIABCARE protocol.(238) The assessment included an interview by diabetes nurses, vitals, anthropometric measurements, fundus examination, and podiatry assessment. The medical history, current drugs, smoking and alcohol consumption were recorded using standard questionnaires. Body mass index was calculated as body weight (kg) divided by body height (m) squared. Waist circumference was measured at a level midway between the lower rib margin and iliac crest with the tape all around the body. Blood pressure was measured on both arms in the sitting position after resting for at least 15 minutes. After fasting for 8 h overnight, blood was sampled for assays of fasting lipids, glucose, glycated hemoglobin (HbA_{1c}), and renal and liver function tests. The upper limit of normal for alanine aminotransferase (ALT) was 30 IU/l for men and 19 IU/l for women.(239) The abbreviated Modification of Diet in Renal Disease equation recalibrated for Chinese was used to calculate the estimated glomerular filtration rate (eGFR).(240) Spot urine for albumin-creatinine ratio was performed to detect albuminuria.

4.2.3 Fibroscan examination

During the diabetic complication assessment visit, liver stiffness measurement (LSM) and CAP were obtained using FibroScan®502 (Echosens, Paris, France) as described previously.(241) All patients were fasted for at least 8 h before the procedure. The LSM score was represented by the median of 10 measurements and was considered reliable only if at least 10 successful acquisitions were obtained and the interquartile range (IQR)-to-median ratio of the 10 acquisitions was ≤0.3. The CAP score was represented by the median value. The significance of the IQR-to-median ratio for CAP is not well-defined compared to LSM. Thus, it was not used as criteria for reliability for CAP measurements. The CAP was considered reliable and included in the final analysis based upon whether 10 successful acquisitions were obtained as the sole requirement. The M probe was used in the first instance for all patients so that both LSM and CAP could be obtained. If the M probe failed, the XL probe catering for obese patients was used.(242)

Hepatic steatosis was graded by CAP using the M probe according to published cutoffs (S1=222-232; S2=233-289; S3≥290 dB/m).(236) Probe-specific LSM cutoffs used to define advanced fibrosis and cirrhosis (M probe F3 = 9.6-11.4, F4≥11.5; XL probe F3=9.3-10.9, F4≥11.0 kpa) were derived from previous studies.(241, 242) Two operators performed the procedures. Both had more than 5 years of experience with Fibroscan, and had performed more than 2000 procedures.

4.2.4 Liver histology

Patients with suspected advanced fibrosis or cirrhosis based on the Fibroscan examination were invited to undergo liver biopsy. Percutaneous liver biopsy was performed using the 16G Temno needle. Liver histology was assessed by a single experienced pathologist (A.W.C.) who was blinded to the clinical data. Histological scoring was performed according to the NASH Clinical Research Network system.(61) Fibrosis was staged from 0 to 4: F0 = absence of fibrosis; F1 = perisinusoidal or portal; F2 = perisinusoidal and portal/periportal; F3 = septal or bridging fibrosis; and F4 = cirrhosis. NASH was defined by the presence of hepatic steatosis and inflammation with hepatocyte injury (ballooning) with or without fibrosis.(97)

4.2.5 Statistical analysis

Data was summarized and presented using appropriate descriptive statistics. The normality of continuous variables was assessed by skewness statistic and graphically by normal probability plot. Triglycerides, fasting plasma glucose, HbA1c , plasma creatinine and alanine aminotransferase were all natural log-transformed before being entered into inferential statistical analysis. Patient characteristics between those with and without valid CAP measurements were compared using independent t-test, chi-square or Fisher's exact tests as appropriate. Multivariable logistic regression analyses were conducted to identify patient characteristics independently associated with each of the outcomes: (1) NAFLD and (2) advanced fibrosis. Univariate analysis was first performed on each of the considered independent variables to select candidate variables for the multivariable analyses. Those factors with a P value < 0.25 in the univariate analyses were selected as candidate variables for backward stepwise multivariable logistic regressions to delineate factors independently associated with each of the outcomes.(243) Furthermore, in order to avoid collinearity, only the independent variable with the highest Wald statistic value (or equivalently the smallest p value) was selected as candidate variable among each set of highly inter-correlated (r >0.8) independent variables [including (1) body weight, BMI and waist circumference; (2) systolic and diastolic blood pressure; (3) total cholesterol and LDL- cholesterol]. Hosmer-Lemeshow test was used to assess the goodness-of-fit of the final multivariable logistic regression models obtained.(243) An insignificant result of Hosmer-Lemeshow test indicates a model that fits the data well. The results of significant factors identified were presented with their odds ratio (OR) and 95% confidence intervals (CI). All the statistical analyses were performed using SPSS 22.0 (SPSS Inc, Chicago, IL, USA). All statistical tests involved were two-tailed and statistical significant level was set at 0.05.

For a prevalence of NAFLD ranging from 30% to 90%, a sample size of 1800 subjects would estimate the prevalence of NAFLD with a margin of error ranging from 1.4 – 2.3% at 5% level of significance. For a prevalence of advanced fibrosis of 1-20%, the same sample size would estimate the prevalence with margin of error ranging from 0.5 – 1.9% at 5% level of significance.

4.3 RESULTS

A total of 2119 patients fulfilled the inclusion criteria and underwent Fibroscan assessment, and 1918 underwent complete diabetes complication screening. Of these, 1884 (98.2%) had a reliable LSM by either the M probe or XL probe, while 1799 (93.8%) had reliable CAP scores using the M probe ***(Figure 21)***.

The mean age was 61 years, and 54% were males ***(Table 16)***. A total of 6.6% had platelet count below 150×10^9/L (6.6% of the patients with valid CAP measurements and 6.7% of those without). Compared with patients who failed M probe examination, those with valid CAP data had shorter duration of diabetes, lower BMI, waist circumference, systolic blood pressure and HbA_{1c}, were more likely to smoke, and were less likely to be male and use anti-hypertensive and insulin. Patients with valid CAP also had less albuminuria.

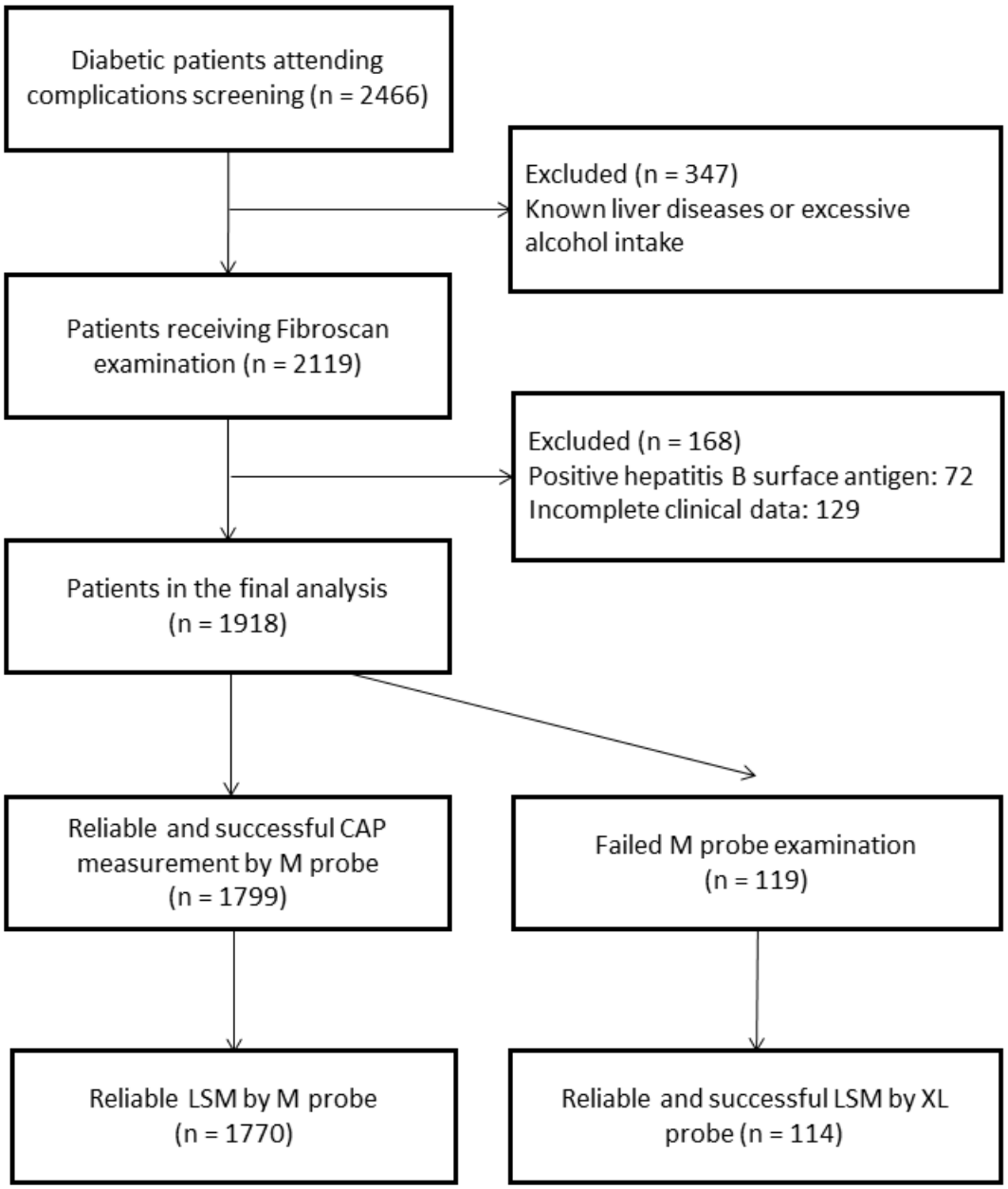


Figure 21: Study participant flowchart

CAP, controlled attenuation parameter; LSM, liver stiffness measurement

Table 16: Characteristics of the study population

	All (n=1918)	With valid CAP		P value
		No (n=119)	Yes (n=1799)	
Clinical and biochemical profile				
Age (years)	60.6 (11.5)	59.7 (12.4)	60.6 (11.5)	0.379
Years since diabetes diagnosed	10.9 (8.5)	13.5 (8.4)	10.7 (8.5)	0.001
Sex[a]				
Male	1041 (54.3%)	38 (31.9%)	1003 (55.8%)	<0.001
Female	877 (45.7%)	81 (68.1%)	796 (44.2%)	
Educational level[a]				
Primary or below	767 (40.3%)	47 (39.8%)	720 (40.3%)	0.289
Secondary	928 (48.7%)	53 (44.9%)	875 (49.0%)	
College or above	209 (11.0%)	18 (15.3%)	191 (10.7%)	
Full / part-time working[a]	771 (40.6%)	42 (35.6%)	729 (41.0%)	0.251
Current smoker[a]	210 (11.0%)	6 (5.1%)	204 (11.4%)	0.034
Regular alcohol drinker[a]	70 (3.7%)	2 (1.7%)	68 (3.8%)	0.317
Body weight (kg)	69.1 (14.2)	82.9 (18.1)	68.2 (13.4)	<0.001
Body mass index (kg/m^2)	26.6 (4.5)	32.8 (5.7)	26.2 (4.1)	<0.001
Waist circumference (cm)	92.9 (11.1)	106.8 (11.4)	92.0 (10.5)	<0.001
Systolic blood pressure (mmHg)	138.0 (19.2)	141.8 (20.4)	137.7 (19.1)	0.023
Diastolic blood pressure (mmHg)	76.2 (11.2)	76.6 (13.3)	76.1 (11.0)	0.693
Total cholesterol (mmol/L)	4.3 (0.9)	4.3 (0.9)	4.3 (0.9)	0.486
HDL-cholesterol (mmol/L)	1.3 (0.4)	1.3 (0.4)	1.3 (0.4)	0.284
LDL-cholesterol (mmol/L)	2.3 (0.7)	2.3 (0.7)	2.3 (0.7)	0.646
Triglycerides (mmol/L)[b]	1.3 (0.9 – 1.9)	1.4 (1.1 – 2.0)	1.3 (0.9 – 1.9)	0.128
Fasting plasma glucose (mmol/L)[b]	7.2 (6.1 – 8.9)	7.8 (6.2 – 10.0)	7.2 (6.1 – 8.9)	0.089
Glycated hemoglobin (%)[b]	7.4 (6.7 – 8.6)	8.0 (6.8 – 9.3)	7.4 (6.7 – 8.5)	0.008
Alanine transferase (IU/L)[b]	23.0 (17.0 – 32.0)	21.0 (17.0 – 34.0)	23.0 (17.0 – 32.0)	0.716
Platelet count (× 10^9/L)[b]	227 (192 – 269)	234 (194 – 289)	226 (191 – 268)	0.218
Estimated glomerular filtration rate (ml/min)	100.7 (34.6)	94.7 (39.6)	101.1 (34.2)	0.086
Urine albumin-creatinine ratio (mg/mmol)	2.6 (0.7 – 13.9)	4.6 (1.6 – 23.6)	2.4 (0.7 – 12.8)	<0.001
Micro or macroalbuminuria[a]	888 (46.9%)	76 (64.4%)	812 (45.8%)	<0.001
Medication use				
Oral anti-diabetic drugs[a]	1661 (86.6%)	105 (88.2%)	1556 (86.5%)	0.589
Anti-hypertensives[a]	1340 (69.9%)	102 (85.7%)	1238 (68.8%)	<0.001
Lipid lowering drugs[a]	1297 (67.6%)	90 (75.6%)	1207 (67.1%)	0.054
Insulin[a]	744 (38.8%)	72 (60.5%)	672 (37.4%)	<0.001

Variables marked with [a] are presented as frequency (%), [b] are presented as median (interquartile range), and the rest are presented as mean (standard deviation).

CAP, controlled attenuation parameter; HDL-cholesterol, high density lipoprotein cholesterol; LDL-cholesterol, low density lipoprotein cholesterol; LSM, liver stiffness measurement

4.3.1 Proportion of patients with increased CAP and LSM

Among 1799 patients with valid CAP data, the median CAP score was 266 dB/m (IQR 216-313). 1309 (72.8%; 95% confidence interval [CI] 70.7-74.8%) had increased CAP of 222 dB/m or more ***(Figure 22/2A)*****.** The number of patients with grade 1, 2 and 3 steatosis as suggested by CAP values was 92 (5.1%), 533 (29.6%) and 684 (38.0%), respectively.

Among 1770 patients with reliable LSM by the M probe, the median LSM was 6.3 kPa (IQR 4.9-8.3). 303 of them (17.1%) had LSM ≥9.6 kPa by M probe suggestive of advanced fibrosis or cirrhosis; 199 (11.2%) also had LSM ≥11.5 kPa suggestive of cirrhosis ***(Figure 22/2B)***. In addition, 114 patients failed M probe examination but had successful LSM by the XL probe. Their median LSM was 6.9 kPa (IQR 4.8-9.6). 31 of them (27.2%) had LSM ≥9.3 kPa by XL probe suggestive of advanced fibrosis or cirrhosis; 25 (21.9%) had LSM ≥11.0 kPa suggestive of cirrhosis ***(Figure 22/2C)***. Taken together, the proportion of patients with increased LSM by either the M probe or the XL probe was 17.7% (95% CI 16.0-19.5%).

As expected, increased LSM was mainly observed in patients with increased CAP. 269 of 1309 (20.6%) patients with increased CAP had increased LSM, compared with 34 of 490 (6.9%) patients with normal CAP (P<0.001). The median LSM of patients with normal CAP but increased LSM was 14.2 kPa (IQR 11.1-16.5). Some of these patients might have burnt out NASH.

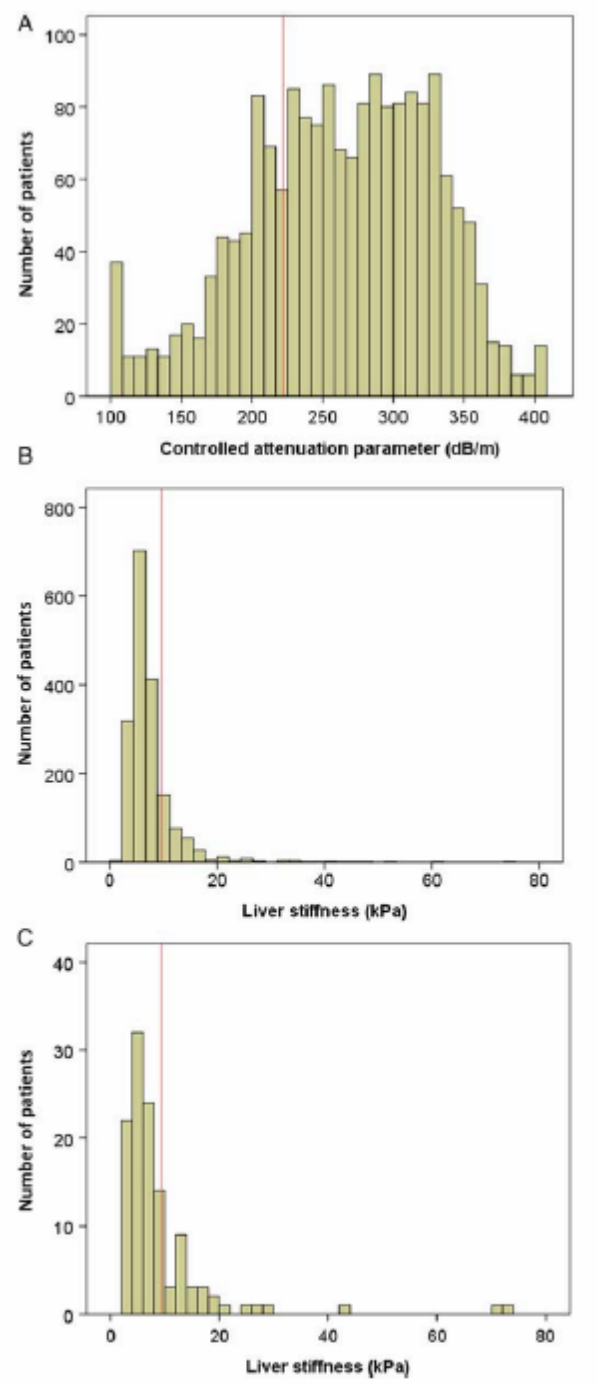


Figure 2 Distribution of (A) controlled attenuation parameter (n=1799), (B) liver stiffness by the M probe (n=1770) and (C) liver stiffness by the XL probe (n=114) in the entire cohort. The red lines represent the threshold for abnormal values: (A) 222 dB/m for controlled attenuation parameter, (B) 9.6 kPa for liver stiffness by the M probe and (C) 9.3 kPa for liver stiffness by the XL probe.

Figure 22: Distribution of CAP, LSM (M probe) and LSM (XL probe)

4.3.2 Factors associated with increased CAP

By univariate analysis, increased CAP ≥222 dB/m was associated with higher body weight, BMI, waist circumference, diastolic blood pressure, triglycerides, fasting plasma glucose and ALT ***(Table 17)*****.** It was also associated with lower high density lipoprotein (HDL)-cholesterol and the use of oral anti-diabetic drugs, anti-hypertensives and lipid lowering drugs. Fewer patients with increased CAP were on insulin treatment. By multivariable analysis, female gender, higher BMI, triglycerides, fasting plasma glucose and ALT level, and not using insulin remained as independent factors associated with increased CAP (Hosmer-Lemeshow test; P=0.564). Increased CAP was found in 410 of 751 (54.6%), 609 of 736 (82.7%), and 280 of 296 (94.6%) patients with BMI <25, 25-30 and ≥30 kg/m^2, respectively (P<0.001) ***(Figure 23/ 3)***. When the different classes of oral anti-diabetic drugs were analyzed separately, none of them were associated with an increased or decreased risk of increased CAP in the multivariable analysis (data not shown).

276 of 335 (82.4%) men with ALT ≥30 IU/l had increased CAP, compared with 434 of 665 (65.3%) men with ALT <30 IU/l (P<0.001). Similarly, 409 of 502 (81.5%) women with ALT ≥19 IU/l had increased CAP, compared with 185 of 291 (63.6%) women with ALT <19 IU/l (P<0.001). The sensitivity, specificity, positive predictive value and negative predictive value of the combined ALT cutoffs to detect increased CAP were 52.5% (95% CI 49.8-55.3%), 68.9% (95% CI 64.6-73.0%), 81.8% (95% CI 79.1-84.4%) and 35.3% (95% CI 32.2-38.4%), respectively.

Table 17: Comparison of study population characteristics and increased controlled attenuation parameter

Characteristics	Increased CAP		OR_U	Univariate P value	OR_A (95% CI)	Multivariate P value
	No (n=490)	Yes (n=1309)				
Clinical and biochemical profile						
Age (years)	61.1 (11.2)	60.5 (11.6)	0.96[c]	0.327	NE	
Years since diabetes diagnosed	10.8 (8.9)	10.7 (8.4)	0.99[c]	0.841	NE	
Sex[a]						
Male (ref)	291 (29.0%)	712 (71.0%)	1		1	
Female	199 (25.0%)	597 (75.0%)	1.23	0.058	1.68 (1.31 – 2.15)	<0.001
Educational level[a]						
Primary or below (ref)	216 (30.0%)	504 (70.0%)	1		NS	
Secondary	220 (25.1%)	655 (74.9%)	1.28	0.030		
College or above	51 (26.7%)	140 (73.3%)	1.18	0.374		
Full / part-time working[a]						
No (ref)	296 (28.2%)	755 (71.8%)	1		NE	
Yes	188 (25.8%)	541 (74.2%)	1.13	0.268		
Current smoker[a]						
No (ref)	433 (27.3%)	1151 (72.7%)	1		NE	
Yes	54 (26.5%)	150 (73.5%)	1.05	0.794		
Regular alcohol drinker[a]						
No (ref)	463 (26.9%)	1255 (73.1%)	1		NE	
Yes	22 (32.4%)	46 (67.6%)	0.77	0.327		
Body weight (kg)	61.1 (10.8)	70.8 (13.3)	1.99[c]	<0.001	NE	
Body mass index (kg/m^2)	23.7 (3.2)	27.2 (4.0)	1.33	<0.001	1.30 (1.25 – 1.35)	<0.001
Waist circumference (cm)	85.6 (9.0)	94.4 (10.0)	2.65[c]	<0.001	NE	
Systolic blood pressure (mmHg)	136.6 (20.6)	138.1 (18.5)	1.04[c]	0.152	NE	
Diastolic blood pressure (mmHg)	74.9 (11.0)	76.6 (11.0)	1.16[c]	0.003	NS	
Total cholesterol (mmol/L)	4.2 (0.9)	4.3 (0.9)	1.10	0.128	NS	
HDL-cholesterol (mmol/L)	1.4 (0.4)	1.3 (0.4)	0.40	<0.001	NS	
LDL-cholesterol (mmol/L)	2.2 (0.7)	2.3 (0.7)	1.06	0.458	NE	
Triglyceride, TG (mmol/L)[b]	1.0 (0.8 – 1.4)	1.4 (1.0 – 2.0)	3.45	<0.001	2.03 (1.59 – 2.58)	<0.001
Fasting plasma glucose (mmol/L)[b]	6.8 (5.7 – 8.4)	7.4 (6.3 – 9.0)	2.34	<0.001	1.74 (1.18 – 2.58)	0.005
Glycated hemoglobin (%)[b]	7.2 (6.5 – 8.6)	7.4 (6.8 – 8.5)	1.37	0.264	NE	
Alanine transferase (IU/L)[b]	19.0 (15.0 – 24.0)	24.0 (18.0 – 35.0)	3.17	<0.001	2.14 (1.63 – 2.81)	<0.001
eGFR (ml/min)	98.8 (37.3)	101.9 (32.9)	1.03[c]	0.088	NS	

Urine ACR (mg/mmol)[b]	1.9 (0.6 – 11.8)	2.7 (0.8 – 13.6)	1.03	0.203	NS	
Medication use						
Oral anti- diabetic drugs[a]						
No (ref)	84 (34.6%)	159 (65.4%)	1		NS	
Yes	406 (26.1%)	1150 (73.9%)	1.50	0.006		
Anti-hypertensives[a]						
No (ref)	191 (34.0%)	370 (66.0%)	1		NS	
Yes	299 (24.2%)	939 (75.8%)	1.62	<0.001		
Lipid lowering drugs[a]						
No (ref)	189 (31.9%)	403 (68.1%)	1		NS	
Yes	301 (24.9%)	906 (75.1%)	1.41	0.002		
Insulin[a]						
No (ref)	284 (25.2%)	843 (74.8%)	1		1	
Yes	206 (30.7%)	466 (69.3%)	0.76	0.012	0.64 (0.50 – 0.82)	<0.001

Variables marked with [a] are presented as frequency (row %), [b] are presented as median (interquartile range) and log-transformed before being entered into association analysis; all others are presented as mean (standard deviation).

[c] The odds ratio was estimated for every 10-unit increase of the underlying factor.

eGFR, estimated glomerular filtration rate; ACR, albumin-creatinine ratio; ref, reference group of the categorical variable; OR_U, unadjusted odds ratio; OR_A, odds ratio adjusted for other significant factors obtained from backward stepwise logistic regression analysis using variables with p-value <0.25 in univariate analysis as candidate variables; NS, not statistically significant in multivariable analysis; NE, not entered into multivariable analysis owing to $p \geq 0.25$ or collinearity.

4.3.3 Factors associated with increased LSM

By univariate analysis, increased LSM (≥9.6 kPa by M probe or ≥9.3 kPa by XL probe) was associated with higher body weight, BMI, waist circumference, blood pressure, triglycerides, fasting plasma glucose, HbA_{1c}, ALT, and spot urine albumin-creatinine ratio ***(Table 18)*****.** It was also associated with lower total cholesterol, HDL-cholesterol, low density lipoprotein-cholesterol and eGFR, and the use of anti-hypertensives and lipid lowering drugs. By multivariable analysis, longer duration of diabetes, higher BMI, ALT and spot urine albumin-creatinine ratio, and lower HDL-cholesterol remained as independent factors associated with increased LSM (Hosmer-Lemeshow test; P=0.672). Increased LSM was found in 60 of 744 (8.1%), 141 of 754 (18.7%), and 131 of 370 (35.4%) patients with BMI <25, 25-30 and ≥30 kg/m^2, respectively (P<0.001) ***(Figure 23/3)*****.** When the different classes of oral anti-diabetic drugs were analyzed separately, none of them were associated with an increased or decreased risk of increased LSM in the multivariable analysis (data not shown).

108 of 345 (31.3%) men with ALT ≥30 IU/l had increased LSM, compared with 85 of 673 (12.6%) men with ALT <30 IU/l (P<0.001). Similarly, 116 of 540 (21.5%) women with ALT ≥19 IU/l had advanced fibrosis, compared with 24 of 320 (7.5%) women with ALT <19 IU/l (P<0.001). The sensitivity, specificity, positive predictive value and negative predictive value of the combined ALT cutoffs to detect advanced fibrosis were 67.3% (95% CI 61.9-72.3%), 57.2% (95% CI 54.7-59.7%), 25.3% (95% CI 22.5-28.3%) and 89.0% (95% CI 86.9-90.9%), respectively. The proportion of patients with increased CAP and LSM by ALT level is shown in ***Table 19.***

Table 18: Comparison of study population characteristics and increased liver stiffness

Characteristics	Increased LSM		OR_U	Univariate P value	OR_A (95% CI)	Multivariate P value
	No (n=1550)	Yes (n=334)				
Clinical and biochemical profile						
Age (years)	60.4 (11.3)	61.2 (12.5)	1.06[c]	0.255	NE	
Years since diabetes diagnosed	10.8 (8.6)	11.6 (8.3)	1.11[c]	0.149	1.20 (1.01 – 1.41)	0.034
Sex[a]						
Male (ref)	827 (81.0%)	194 (19.0%)	1		NS	
Female	723 (83.8%)	140 (16.2%)	0.83	0.116		
Educational level[a]						
Primary or below (ref)	625 (82.7%)	131 (17.3%)	1		NE	
Secondary	745 (81.6%)	168 (18.4%)	1.08	0.569		
College or above	168 (83.2%)	34 (16.8%)	0.97	0.868		
Full / part-time working[a]						
No (ref)	904 (81.5%)	205 (18.5%)	1		NE	
Yes	628 (83.1%)	128 (16.9%)	0.90	0.390		
Current smoker[a]						
No (ref)	1366 (82.0%)	300 (18.0%)	1		NE	
Yes	174 (84.1%)	33 (15.9%)	0.86	0.464		
Regular alcohol drinker[a]						
No (ref)	1486 (82.4%)	318 (17.6%)	1		NE	
Yes	53 (79.1%)	14 (20.9%)	1.23	0.492		
Body weight (kg)	67.5 (13.2)	76.7 (16.1)	1.55[c]	<0.001	NE	
Body mass index (kg/m^2)	26.1 (4.1)	29.3 (5.2)	1.17	<0.001	1.14 (1.10 – 1.17)	<0.001
Waist circumference (cm)	91.5 (10.5)	99.5 (11.5)	1.94[c]	<0.001	NE	
Systolic blood pressure (mmHg)	137.3 (19.4)	141.4 (18.2)	1.11[c]	0.001	NS	
Diastolic blood pressure (mmHg)	75.8 (11.1)	77.8 (11.7)	1.16[c]	0.004	NE	
Total cholesterol (mmol/L)	4.3 (0.9)	4.2 (0.9)	0.86	0.033	NE	
HDL-cholesterol (mmol/L)	1.3 (0.4)	1.2 (0.4)	0.36	<0.001	0.65 (0.44 – 0.95)	0.027
LDL-cholesterol (mmol/L)	2.3 (0.7)	2.2 (0.7)	0.77	0.004	NS	
Triglyceride, TG (mmol/L)[b]	1.3 (0.9 – 1.8)	1.5 (1.1 – 2.2)	1.83	<0.001	NS	
Fasting plasma glucose (mmol/L)[b]	7.2 (6.0 – 8.7)	7.7 (6.3 – 9.6)	1.91	0.001	NS	
Glycated hemoglobin (%)[b]	7.4 (6.7 – 8.5)	7.5 (6.8 – 9.1)	1.98	0.027	NS	
Alanine transferase (IU/L)[b]	22.0 (16.0 – 29.0)	32.0 (22.0 – 48.0)	4.48	<0.001	4.08 (3.12 – 5.34)	<0.001
eGFR (ml/min)	101.7 (34.2)	96.1 (36.3)	0.95[c]	0.007	NS	
Urine ACR (mg/mmol)[b]	2.1 (0.7 – 11.3)	6.2 (1.6 – 30.2)	1.22	<0.001	1.17 (1.10 – 1.25)	<0.001

Medication use					
Oral anti-diabetic drugs[a]					
No (ref)	210 (84.3%)	39 (15.7%)	1		NE
Yes	1340 (82.0%)	295 (18.0%)	1.19	0.360	
Anti-hypertensives[a]					
No (ref)	501 (88.7%)	64 (11.3%)	1		NS
Yes	1049 (79.5%)	270 (20.5%)	2.02	<0.001	
Lipid lowering drugs[a]					
No (ref)	518 (85.1%)	91 (14.9%)	1		NS
Yes	1032 (80.9%)	243 (19.1%)	1.34	0.029	
Insulin[a]					
No (ref)	963 (83.4%)	192 (16.6%)	1		NS
Yes	587 (80.5%)	142 (19.5%)	1.21	0.114	

Variables marked with [a] are presented as frequency (row %), [b] are presented as median (interquartile range) and log-transformed before being entered into association analysis; all others are presented as mean (standard deviation).

[c] The odds ratio was estimated for every 10-unit increase of the underlying factor.

eGFR, estimated glomerular filtration rate; ACR, albumin-creatinine ratio; ref, reference group of the categorical variable; OR_U, unadjusted odds ratio; OR_A, odds ratio adjusted for other significant factors obtained from backward stepwise logistic regression analysis using variables with p-value <0.25 in univariate analysis as candidate variables; NS, not statistically significant in multivariable analysis; NE, not entered into multivariable analysis owing to $p \geq 0.25$ or collinearity.

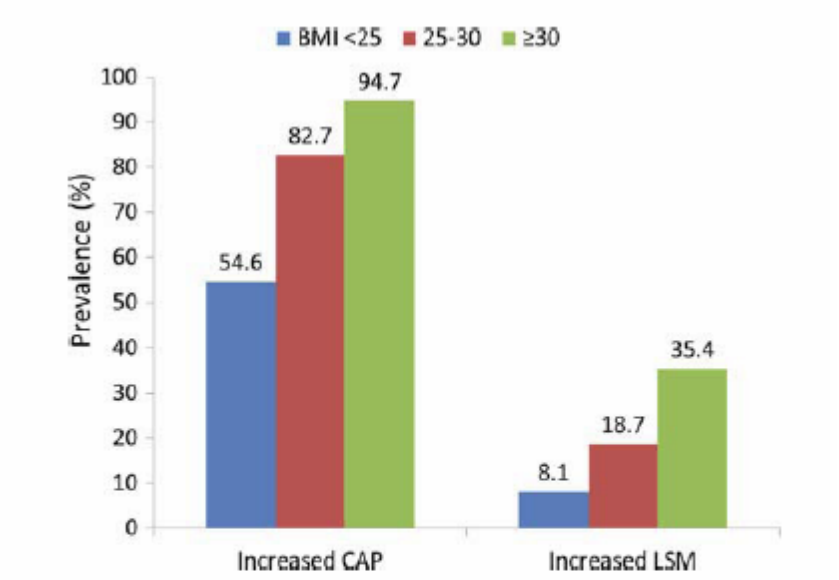


Figure 3 Proportion of patients with increased controlled attenuation parameter (CAP) and liver stiffness measurement (LSM) according to body mass index. A CAP value of ≥222 dB/m was used to estimate the proportion of patients with non-alcoholic fatty liver disease. An LSM of ≥9.6 kPa by the M probe or ≥9.3 kPa by the XL probe was used to estimate the proportion of patients with advanced fibrosis.

Figure 23: Prevalence of CAP and LSM by BMI

Table 19: Proportion of patients with increased CAP and LSM by ALT levels

	CAP <222 dB/m	CAP ≥222 dB/m	P
All patients (n=1799)			
Normal LSM	456 (93.1%)	1040 (79.4%)	<0.001
Increased LSM[a]	34 (6.9%)	269 (20.6%)	
Normal ALT[b] (n=956)			
Normal LSM	317 (94.1%)	539 (87.1%)	<0.001
Increased LSM[a]	20 (5.9%)	80 (12.9%)	
High ALT[b,c] (n=837)			
Normal LSM	138 (90.8%)	497 (72.6%)	<0.001
Increased LSM[a]	14 (9.2%)	188 (27.4%)	

[a]Increased LSM was defined as LSM ≥9.6 kPa by the M probe or ≥9.3 kPa by the XL probe.
[b]There were 6 patients with missing ALT values.
[c]High ALT was defined as ALT ≥30 IU/L in men and ≥19 IU/L in women.
ALT, alanine aminotransferase; CAP, controlled attenuation parameter; LSM, liver stiffness measurement

4.3.4 Liver histology

94 patients underwent liver biopsy ***(Table 20)*****.** The mean liver biopsy length was 20 (SD 7) mm and mean number of portal tracts was 8 (SD 3). Eighty seven patients had M probe examination with a median CAP of 320 dB/m (IQR 286-350 dB/m) and LSM 14.1 kPa (IQR 11.8-20.6 kPa). Seven patients only had XL probe examination, and the median LSM was 17.6 kPa (IQR 8.6-29.2 kPa). Seventy five of these 94 (80%) patients had increased LSM (≥9.6 kPa by M probe or ≥9.3 kPa by XL probe). The sensitivity, specificity, positive predictive value and negative predictive value of increased LSM to detect F3-4 disease were 94%, 34%, 59% and 84%, respectively. Eighty two of the 87 patients who had a median CAP had an increased CAP (≥222 dB/m). The sensitivity, specificity, positive predictive value and negative predictive value of increased CAP to detect steatosis (S≥1) was 95%, 14%, 93% and 20% respectively. Lobular inflammation was found in 89 (95%) patients, hepatocyte ballooning in 54 (57%), and fibrosis in 89 (95%) (Table 5). In particular, 60 (50%) patients had F3-4 disease. 78 (83%) patients had NAFLD activity score ≥3. 53 (56%) subjects had NASH.

Seven subjects had steatosis found in <5% of hepatocytes (liver specimen length 14-26 mm). Three had normal liver histology, 1 had mild lobular inflammation but no fibrosis, 2 had bridging fibrosis (F3) and 1 had cirrhosis (F4). All 3 patients with F3-4 disease had grade 1-2 lobular inflammation and no other etiologies of chronic liver disease. The picture was suggestive of burnt out NASH.

Table 5. Histological severity of 94 patients with liver biopsy

Table 20: Histological severity of 94 patients with liver biopsy

Histological features	Grade/Stage	n	%	CAP (dB/m)	LSM (kPa)[a]
Steatosis	0	7	7	238 (234-263)	17.3 (14.5-28.8)
	1	38	40	310 (284-349)	14.4 (11.6-26.7)
	2	38	40	324 (303-348)	13.8 (11.8-17.1)
	3	11	12	351 (310-393)	12.8 (7.3-15.3)
Lobular inflammation	0	5	5	263 (249-306)	14.9 (14.2-20.1)
	1	57	61	318 (283-351)	14.2 (11.8-22.2)
	2	32	34	326 (295-351)	13.7 (9.0-17.1)
	3	0	0	-	-
Ballooning	0	40	43	324 (274-358)	13.1 (9.3-16.8)
	1	47	50	317 (279-347)	14.3 (11.9-21.3)
	2	7	7	310 (299-364)	26.6 (15.3-36.8)
Fibrosis	0	5	5	235 (204-263)	8.8 (14.9-18.4)
	1	29	31	341 (310-359)	11.8 (7.8-12.9)
	2	13	14	306 (272-341)	12.6 (7.7-14.4)
	3	20	21	321 (296-350)	14.0 (12.7-16.3)
	4	27	29	323 (292-338)	23.8 (14.8-35.3)
NAFLD activity score[b]	0-2	16	17	269 (238-331)	14.5 (13.1-21.3)
	3-4	62	66	319 (288-349)	14.0 (11.7-21.8)
	≥5	16	17	333 (321-356)	14.1 (8.8-16.8)
Non-NASH		41	44	322 (270-357)	13.2 (9.9-17.0)
NASH		53	56	317 (288-347)	14.5 (11.9-23.8)

CAP and LSM were expressed in median (interquartile range).

[a]The LSM based on M probe examination is shown because only 7 patients in this cohort required XL probe measurements.

[b]The NAFLD activity score: 0-2 Non NASH, 3-4 Possible NASH, ≥5 Probable NASH

CAP, controlled attenuation parameter; LSM, liver stiffness measurement; NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis

4.4 DISCUSSION

In this large prospective hospital cohort, NAFLD was found in the large majority of diabetic patients. A significant proportion also had increased LSM. Patients with high BMI and additional metabolic factors were most likely to have increased CAP and LSM.

In previous community-based studies using abdominal ultrasonography, around 32-62% of diabetic patients were found to have NAFLD.(244-247) The prevalence of NAFLD is higher in our cohort for several reasons. First, abdominal ultrasonography is operator-dependent and is insensitive to mild steatosis.(210) Therefore, NAFLD may be underreported and missed in some cases. In comparison, CAP measurement by FibroScan can detect liver fat involving as little as 10% of the hepatocytes.(235-237) Our study confirmed the discriminating ability of CAP for detecting steatosis was confirmed in the 87 subjects who had a liver biopsy. The sensitivity and positive predictive value of CAP to detect S≥1 were 95% and 93% respectively.

Our study was conducted at the hospital setting. Although we had an open referral system for both hospital and primary care clinics, the metabolic burden of our patients was higher than that of patients in the community setting.

Since we could measure CAP and liver stiffness simultaneously with FibroScan, it was possible to assess not only the prevalence of NAFLD but also disease severity. According to past histological studies, diabetes is one of the most important risk factors of cirrhosis in patients with NAFLD (189, 248). At the population level, diabetes also doubles the risk of hepatocellular carcinoma in men.(249) A retrospective study of 1131 diabetes patients using FibroTest also reported a prevalence of between 2.8–5.6% for advanced fibrosis.(250) Similarly, studies from Europe and Australia found that 5-35% of diabetic patients had increased LSM by different cutoffs.(251-253) In our study, 18% of the patients were found to have increased LSM. We further performed liver biopsy in 94 patients and confirmed the presence of advanced fibrosis or cirrhosis in half of the patients. If we adjust for the positive predictive value of LSM, around 10% of our diabetic patients would still have advanced liver disease. This argues strongly in favor of liver assessment in diabetic patients.

Previous studies report an LSM failure rate of around 3% and unreliable measurements in 12-16%.(222, 223) In the current study, however, 98.2% of the subjects could have successful and reliable LSM. Differences in study design can explain the apparent discrepancies. First, the use of XL probe allowed successful LSM even in obese patients.(242) Second, recent data suggest that the success rate of LSM (number of valid acquisitions divided by the total number of acquisitions) does not affect the reliability of LSM as compared to liver histology.(55) Therefore, we did not include success rate as one of the reliability criteria in this study. Finally, all examinations were performed by experienced operators who had performed more than 2000 measurements before. Operator experience is pivotal in achieving successful and reliable LSM.(222)

Although FibroScan is easy to perform, it is unlikely that clinicians can apply it to all diabetic patients because of the large number of patients and the availability of assessment. Therefore, it is important to identify patients at risk of NAFLD and advanced liver disease. Traditional metabolic risk factors such as high BMI and dyslipidaemia were independent factors associated with NAFLD and increased LSM. In other words, patients with diabetes and additional metabolic diseases are at higher risk of having advanced liver disease and may benefit from liver assessment. That said, it is important to note that among diabetic patients with BMI less than 25 kg/m^2, 55% still had increased CAP and 8% had increased LSM. In previous population studies, NAFLD and advanced fibrosis are observed in a small but significant proportion of non-obese subjects. Such patients usually have other components of metabolic syndrome and despite relatively normal BMI often have recent weight gain.(231, 254) In contrast, ALT correlates poorly with histological severity.(192, 255) Metabolic factors are more important than ALT in guiding liver assessment in diabetic patients.

Thiazolidinediones improve liver histology in NASH patients, and use of metformin has been shown to be associated with a lower mortality rate in cirrhotic patients.(256, 257) Nevertheless, neither class of drugs was independently associated with NAFLD or advanced fibrosis in our study. It should however be noted that the current study is not a randomized controlled trial. The selection of different drugs was influenced by the underlying disease status. Unexpectedly, insulin use was associated with a lower risk of NAFLD. This is counterintuitive as patients requiring insulin treatment usually have failed treatment with oral anti-diabetic drugs and thus have poorer diabetic control. One possible explanation is that patients with cirrhosis are more insulin resistant and therefore more likely require insulin treatment. As a NAFLD patient progresses to cirrhosis, liver fat tends to disappear and the patient would thus be classified as having no fatty liver.(193) In fact, there was a trend that insulin use was associated with advanced fibrosis in our entire cohort ***(Table 18)***. Three of our patients who underwent liver biopsy also had features of burnt out NASH.

Our study has the strength of a large sample size and the use of one of the best and widely available non-invasive tests of liver steatosis and fibrosis. Compared to previous studies, ours also had comprehensive diabetic assessment. Nevertheless, there were several limitations. First, at the time of the study, CAP could only be measured by the M probe. More obese patients who required XL probe examination only had LSM but not CAP data. The true prevalence of NAFLD would therefore be even higher. This would be particularly relevant to Western countries where the prevalence of obesity is higher. In the Caucasian cohorts described above, 8-15% of the diabetic patients had failed M probe examination.(251-253) Nonetheless, the FibroScan programme has been updated to allow CAP measurement by XL probe. If validated, CAP measurement can also be done in obese patients. Second, liver biopsy was only performed in a subset of patients. However, it is unethical to biopsy patients with no apparent liver disease. A low LSM also has excellent negative predictive value in excluding advanced fibrosis.(241) Along the same line, the diagnostic performance of LSM in the histology subgroup of this study should be interpreted with caution because patients undergoing liver biopsy were selected based on increased likelihood of advanced disease. Third, our study was conducted on the local Hong Kong Chinese patients, which may limit the generalizability of the findings to other ethnicities. Finally, we only

reported the baseline assessment. The patients are currently under prospective follow-up for 10 years. The prognostic implications of the baseline liver assessment will be further unraveled.

In conclusion, diabetic patients at hospital and primary care clinics have a high prevalence of NAFLD and advanced liver fibrosis. Diabetic patients with high BMI and dyslipidaemia are at particularly high risk and may be the target for liver assessment. Our data support screening for NAFLD and/or advanced fibrosis in patients with type 2 diabetes. However, while LSM is good at ruling out advanced fibrosis, its accuracy in ruling in the disease remains limited.

This page has been left intentionally blank

CHAPTER 5: Assessment of liver fibrosis in chronic hepatitis B patients with transient elastography, and other non-invasive measures.

5.0 CHAPTER SUMMARY

Introduction: Staging liver fibrosis is useful in determining which chronic hepatitis B patients will benefit from antiviral therapy. Liver biopsy is the reference standard, but has several limitations. Non-invasive methods to stage liver disease are needed. Transient Elastography (TE) has been developed and has shown good correlation to date. The diagnostic accuracy and usefulness of TE is evaluated against liver histology. The performance of TE is compared with other easily performed non-invasive methods: FIB 4 index, Aspartate Platelet Ratio Index (APRI), Aspartate Alanine aminotransferase Ratio (AAR), Age Platelet Index (API), Fibrosis Index (FI). TE was also compared with Caffeine Breath Test (CBT) in a small cohort.

The accuracy of liver histology assessment is potentially affected by interobserver variability. Significant variability in histological assessment is therefore a factor which affects the validation of non-invasive methods. We sought to determine the degree of interobserver variability in the histological assessment of liver biopsy.

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 histopathology were collected. Transient elastography was performed. At least 10 successful measurements of LSM was required for a valid reading, with an interquartile range to median ratio of < 30% when LSM≥7.1 kPa. On liver biopsy specimens available, a second blinded assessment was performed. The fibrosis stage from the second assessment was compared to the original assessment.

Results: Seventy one patients were recruited. Liver Stiffness Measurement (LSM) Area Under Receiver Operator Characteristic (AUROC) curves for F≥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≥2 and F≥3 can be diagnosed or excluded with a very high degree of certainty (>90%) in 49.3% and 57.7% respectively. TE compared favourably against FIB 4, APRI, AAR, API and FI for every stage of fibrosis. In 7 patients, caffeine breath tests had the best performance for diagnosing fibrosis stage. Liver biopsies were classified differently in the second histological assessment in 44% of cases, with the intraclass correlation coefficient showing moderate agreement (K =0.457, p<0.001).

Conclusions: TE is a reliable and accurate non-invasive tool for diagnosing fibrosis stage. It has particularly high accuracy for F≥3 and F4, 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. However the inherent variability of liver biopsy can impede the true accuracy of non-invasive modalities.

5.1 BACKGROUND

5.1.1 Introduction to chronic hepatitis B – burden of disease worldwide and in Australia

Approximately 240 million people are chronically infected with hepatitis B virus (HBV) (93, 94). More than 75 % reside in the Asia and the Western Pacific (258).

Chronic Hepatitis B infection has a wide spectrum of disease and dynamic natural history. There are 4 recognized phases of infection. These are known as immune tolerance, immune clearance, inactive and immune reactivation. Different patients transition through these phases with a great deal of variability and irregularity.

Progression toward liver cirrhosis and development of hepatocellular carcinoma (HCC) complicate CHB infection. The annual incidence of cirrhosis is about 2% and HCC 3-6% (highest risk in cirrhotic patients) (259). The main risk factors (of HCC or cirrhosis) are male gender, hepatitis B e antigen (HBeAg) status, serum HBV viral load, alanine aminotransferase (ALT) level, HBV genotype, family history of hepatocellular carcinoma (HCC), and alcohol consumption (260). The cumulative lifetime risk of cirrhosis and HCC can range 15.9 - 76.2% and 4.4 - 61.8% respectively (260).

Chronic hepatitis B is the major cause of HCC accounting for 60-80% of the world's total cases. HBV results in over 780 000 deaths per year from end stage liver disease and HCC (93, 94). HBV comprise 5–10% of cases for liver transplantation (261-263)

In Australia, vaccination programs have significantly reduced the incidence of new HBV infection, resulting in a low prevalence of HBV. However, there are at risk communities where there is a much higher prevalence due to population movements. The latest estimate of CHB prevalence in Australia is approximately 1% (204,000 persons) (113). People who are born overseas account for 56.1% of the total cases of CHB, with 95% of new cases are due to immigrants from endemic countries. The migrants with the highest rates of CHB are those from the Asia/Pacific (3.55%) and Africa/Middle East (2.69%). The increasing incidence of CHB related HCC appears to be largely occurring in Asia-Pacific born Australians, and is projected to increase from 140 cases in 2005 to 250 cases per year by 2025 (264). At risk groups for people born in Australia are those who inject drugs (4%), Aboriginal and Torres Strait islanders (3.7%), and men who have sex with men (3%) (113).

5.1.2 Limitations of current chronic hepatitis B antiviral therapy

The goal of therapy is to prevent progression of CHB infection to cirrhosis, decompensation, HCC and death. Active HBV replication is the key driver of liver injury and disease progression. Current antivirals suppress viral replications, but are unable to eradicate the virus. In patients receiving antiviral therapy, rates of complete virological response, as defined as hepatitis B surface antigen (HBsAg) loss, remain low (265).

The 2 main forms for antiviral therapy for CHB infection are pegylated interferon and nucleotide/nucleoside analogues (NAs). Pegylated interferon offers finite treatment duration, but can be

associated with significant side effects. HBeAg seroconversion occurs in 32% (in those who are HBeAg positive at baseline), undetectable viral load in 14-19%, and ALT normalization in 41-59% (266).

The current first line NAs, entecavir monohydrate and tenofovir disoproxil fumerate, are potent antivirals with high genetic barrier to drug resistance and few side effects. Entecavir achieves viral suppression with undetectable DNA in 96% of patients, and a 0.4% rate of resistance over 4 years (267). Tenofovir therapy is associated with undetectable viral DNA in 99% of patients, with no resistance after 5 years (15).

The disadvantage of NAs is that viral suppression relies on sustained therapy, with relapse upon drug cessation. In HBeAg positive patients who achieve HBeAg seroconversion during treatment, more than 90% have detectable viral levels within 4 years of stopping treatment (268, 269). In HBeAg negative patients, , viral relapse (defined as HBV DNA>2000 IU/ml after 3 or more consecutive undetectable levels taken 6 months apart) occurs in 91.4% within 1 year of NAs cessation (270). Hepatitis B surface antigen loss is recommended as the end point of therapy by local and international guidelines (259, 267, 271, 272). Rates of surface antigen loss are only up to 2% per year (273-275) and hence most patients require long term therapy. This can be associated with high costs, and since the highest rates of CHB infection occur in low income countries, long term antiviral therapy may be out of reach for majority of the population (276).This means it is crucial to accurately select patients who will derive the most benefit from antivirals.

5.1.3 The importance of assessing liver fibrosis in chronic hepatitis B

Regional and international guidelines universally recommend liver fibrosis staging to be considered as part of the assessment of CHB patients (259, 267, 271, 272, 277). Diagnosing the stage of liver fibrosis is important because it allows for the identification of CHB patients with advanced disease, which aids in determining which patients will benefit from antiviral therapy, and from surveillance of cirrhosis complications. Therapy should be considered in those with moderate (Metavir stage F2) or advanced fibrosis (Metavir stage F3) (259, 267, 271, 272, 277) so as to prevent progression to cirrhosis (Metavir stage F4). Early cirrhosis can have no clinical, biochemical or radiological features. Diagnosis with more accurate measures will allow for the optimal commencement of surveillance programs to detect HCC and treat gastro-oesophageal varices.

5.1.4 Transient Elastography in chronic hepatitis B

Liver biopsy is the reference standard test that is recommended by all guidelines in the assessment of fibrosis (259, 267, 271, 272). However, liver biopsy is not an ideal method of assessment for liver fibrosis due to reasons discussed earlier in this thesis (see section 1.5). In brief, liver biopsy is poorly accepted by patients due to pain and risk of complications. It is resource intensive which makes it impractical to

apply routinely in large volumes. There can also be variability in sampling and interobserver interpretation.

In 2009 when this study was initiated, Transient elastography (TE) was still only an emerging method of liver fibrosis assessment. Most of the TE data at the time were from Europe on Caucasian hepatitis C patients with little literature on hepatitis B patients.

Marcellin et al in 2008 (87) was first to perform a dedicated study in CHB patients that compared Liver stiffness measurement (LSM) and liver biopsy. In 173 patients, the diagnostic performance for TE was: AUROC= 0.81 (95% confidence intervals (CI) 0.73-0.86) for F≥2; 0.93 (95% CI 0.88-0.96) for F≥3; and 0.93 (95% CI 0.82-0.98) for F=4. Optimal LSM cutoff values were determined to be LSM≥7.2 kPa for F≥2, and LSM≥11.0 kPa for F=4. These LSM cutoffs were lower than those observed for hepatitis C patients (77, 78, 84). This was thought to be due to hepatitis C having relatively more micronodular cirrhosis (278, 279), where nodules are smaller and denser implying denser fibrosis.

The only other literature at the time was by Wong et al in 2008 (280), who examined the relationship of LSM and the distribution of the fibrosis in CHB patients via image and morphometric analysis. Liver Stiffness Measurement had a higher correlation with pericellular fibrosis (r=0.43), compared to periportal (r=0.21) or perivenular fibrosis (r=0.25). Pericellular fibrosis occurs in the latter stages of fibrosis progression. Hence the investigators concluded that LSM correlates better with advanced fibrosis and cirrhosis. The serum ALT was also noted to have an effect on the LSM. Elevated ALT was associated with higher LSM scores and higher cutoffs were required to diagnose the same stage of fibrosis compared to normal ALT subjects (280).

Since there had been only 2 validation studies for chronic hepatitis b, our goal was to provide further validation of the diagnostic performance of TE in CHB fibrosis staging. The performance of TE was compared to other non-invasive markers of fibrosis, and the effect of ALT on determining optimal LSM cut-offs was also examined. At the time the study's conception, these aims would have led to results that would have been amongst the first reported. As of the time of writing however, many studies have since reported the use of TE and other non-invasive measures of liver fibrosis in CHB patients. Although no longer as novel, our findings are still important to report for comparison and independent validation, particular for our local population. More recent studies are discussed in the context of our study's findings later in discussion sections 5.4.7 and 5.4.8 of this chapter.

5.1.5 Caffeine Breath Test, FIB-4 index, APRI, API, FI, AAR and other non-invasive measures of liver fibrosis

Other non-invasive measures of liver fibrosis for CHB include Caffeine breath test (CBT) (281, 282), Hui index (283), Zeng Index (284), age-spleen ratio index (285)and compensated cirrhosis index (286). Measures that were specifically for CHC include the following: FIB-4 index (287), aspartate platelet ratio index (APRI) (288), age platelet index (API)(289), fibrosis index (FI) (290), aspartate aminotransferase-alanine aminotransferase ratio (AAR) (291), cirrhosis discriminant score (292), Fibrotest (293), Forns

Index (294), European Liver Fibrosis score (295), Hepascore (296), Fibrometer (297), Go¨teborg University Cirrhosis Index (298), Fibroindex (299), and S index (300).

Amongst the measures mentioned, the FIB-4 index, APRI, API, FI and AAR are the panels with the greatest simplicity. These only require liver function and/or platelet count levels while other panels require more elaborate parameters such as specific imaging results (eg. spleen length) or fibrosis markers (eg. hyaluronic acid). Elaborate parameters are not routinely performed and so could not be obtained in this cross sectional study. Hence the FIB-4 index, APRI, API, FI and AAR were chosen to be comparators with TE.

Caffeine breath test measures the plasma caffeine clearance using ^{13}C- caffeine which is a reflection of hepatic dysfunction (281). A local study of 48 CHB patients found that fibrosis correlated with the ^{13}C-caffeine breath test ($r=-0.62$, $P < 0.001$) and independently predicted significant ($F\geq 2$) and advanced ($F\geq 3$) fibrosis. Improvement in the^{13}C-caffeine clearance level by 61% ($P<0.001$) was seen in those responsive to lamivudine. By comparison, values remained stable or deteriorated in those with persistent viraemia and elevated alanine aminotransferase (282). The performance of TE and other non-invasive measures has never been assessed against CBT, and thus we made this comparison in a small number of patients in this study.

5.1.6 Inter-observer variability in histological staging of liver fibrosis may limit the accuracy of non-invasive tests

Inter-observer variability in liver biopsy assessment is well described. In academic centres with pathologists who specialise in hepatic histopathology, reproducibility in scoring fibrosis is good. A high level of concordance was found amongst 10 specialist liver pathologist who were part of the METAVIR study group ($K=0.80$-0.91) (32). This variability is higher outside of these specialist centres with discordance of up to 49.9%. In this study, we sought to determine the degree of inter-observer variability in the assessment of liver fibrosis stage by re-evaluating liver biopsy slides and comparing the results with the original assessment. This was performed with the aim of clarifying the sensitivity and specificity against which the noninvasive tests are being measured against.

5.1.7 Hypothesis and specific objectives

The main hypothesis is that the LSM will show similar AUROCs to Marcellin's study of between 0.81 – 0.93 (87) for predicting the fibrosis stage in chronic hepatitis B patients. We believe the LSM will compare favourably to other non-invasive tests and will provide further validation for the use of TE to diagnose fibrosis stage in our local population of chronic hepatitis B. In detail, the following are the objectives:

- Objective 1: To assess the rate of scan failure and the reliability of TE scans
- Objective 2: To determine the diagnostic performance of TE and suggest LSM cut-offs for each stage of Fibrosis in CHB patients
- Objective 3: To examine the effect of ALT on LSM and to devise specific cut-offs for elevated and normal ALT
- Objective 4: To apply the cut-offs on the study population and evaluate the utility of using TE to diagnose moderate and advanced fibrosis
- Objective 5: To evaluate the effectiveness of TE for diagnosing clinically silent cirrhosis
- Objective 6: To compare the performance of TE against other non-invasive measures of APRI, AAR, FIB4, FI and API.
- Objective 7: To compare the performance of caffeine breath testing with Fibroscan and other non-invasive measures of fibrosis in a small cohort of patients.
- Objective 8: To determine the interobserver variability by comparing the initial histological assessment for liver biopsy compared to a second histological evaluation.

5.2 METHODS

5.2.1 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 2 tertiary centres were prospectively recruited. TE was performed within 6 months of the liver biopsy. Some patients already had TE and liver biopsy within a 6 month time frame as part of clinical management and were included retrospectively. Seventy three patients were recruited. All subjects were over 18yrs of age and gave written informed consent.

5.2.2 Transient Elastography Assessment

The performance of the Fibroscan has been described in detail earlier in this thesis (refer to section 1.7). In brief, 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 is ≥ 7.1 kPa (55). Two trained operators including this author performed all the TE scans.

Two patients of the total 73 had unsuccessful readings, so that 71 cases were included in the final analysis.

5.2.3 Data Collection

Clinical data were collated from medical records from the public hospital or private specialists' rooms. Proforma included the following information: age, gender, alcohol intake, any other documented chronic liver disease, details of any antiviral therapy, INR, liver function tests, platelet count, hepatitis B sAg status, hepatitis B c antigen status, hepatitis B DNA viral load, alpha fetoprotein and any imaging results. Lab results closest to the Fibroscan date and not exceeding 1 month were recorded.

5.2.4 Original histological assessment of fibrosis stage and a second reference histological assessment

Baseline liver biopsy was record including Metavir Fibrosis score, biopsy length and number.

Where slides were still available, the biopsies were reviewed and Metavir score assessed by two experienced histopathologists (B.P.C.L and J.T), who were blinded to the results of the original assessment. Discrepancies between the two assessors were resolved with further discussions to reach a consensus assessment. The Metavir Fibrosis score obtained from the second evaluation was used as the final reference score for comparisons with all the non-invasive tests in the study.

5.2.5 Non-invasive markers

Formulas for the non-invasive measures analysed in the study are:

- *FIB-4 = [age (yrs) x AST (U/L)] / [platelet count (x10^9/L) x square root(ALT(U/L))] (287)*
- *APRI = 100x(AST (U/L) / upper level of normal) / platelet count (x10^9/L).(204)*
- *API score = 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 (10^9/l): >225 = 0, 200–224 = 1, 175–199 = 2, 150–174 = 3, 125–149 = 4,<125 = 5] (289)*
- *AAR = AST (U/L) / ALT (U/L). (291)*
- *FI score (fibrosis index) = 8 - 0.01 × number of platelets (10^9/L) - albumin (g/dl)(290)*

5.2.6 Caffeine Breath Test

Seven subjects were recruited to undergo caffeine breath testing within 6 months of their liver biopsy. Subjects abstained from caffeine-containing products and limited alcohol consumption to 10 g for 24 h prior to testing. All nonessential medications were withheld for 48 h prior to testing. After an overnight fast, subjects ingested 2 mg/kg of [3-methyl-13C] caffeine (99% 13C), obtained as powder from Cambridge Isotope Laboratories (Cambridge, MA, USA) and dissolved in 30 mL of water, followed by a 40 mL water wash of the container. The quantity of caffeine consumed was equivalent to two cups of coffee. Subjects sat quietly for 15 min before and throughout the CBT as physical activity influences endogenous CO_2 production. Sixteen Paired breath samples were obtained simultaneously during prolonged expiration into 10 mL glass vials via straws. Samples were collected immediately prior to, and 60 min after, caffeine ingestion. The ^{13}C-enrichment of expired CO_2 was determined by continuous flow isotope ratio mass spectrometry17 using an Automated 13-Carbon Breath Analyzer (PDZ Europa, Cheshire, UK). ^{13}C enrichment was expressed as D, by subtracting the average pre-dose enrichment from the average post-dose measurement, with respect to the international (13C/12C) PDB standard, originating from the Pee-Dee-Belemnite, a fossil limestone of the Pee-Dee-Formation in South Carolina, with a 13C/12C isotope ratio of 0.0112372.18. This was then expressed per 100 mg of the original caffeine dose. The reference range was determined in previous studies (281, 301).

5.2.7 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 regression 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. All statistical tests involved were two-tailed and statistical significant level was set at 0.05.

5.3 RESULTS

5.3.1 Clinical data and characteristics of the study patients

The mean age of the study population was 46.1 yrs (SD 11.9) and were predmoninantly male (69%). The mean of the log_{10}HBVDNA viral load was 5.0 IU/ml (SD 2.5), ALT 121 IU/L (SD 284) and AST 79 IU/L (SD 161). HBeAg positive patients consisted of 33/71 (46.5%) subjects.

The mean bilirubin, albumin and platelet count were not elevated. The mean ETOH consumed per week was 4.8g (SD 10). Data for INR, alphafetoprotein, imaging, height and weight were incomplete and was not included in the analysis.

Metavir fibrosis stage used was based on the second assessment of the liver biopsy (as described in 5.2.4). A total of 54/71 biopsies were able to be re-assessed.. For the remainder 17 liver biopsies, the Metavir stage from the original biopsy 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.

The number of subjects who were already being treated with antivirals was 27/71 (38.0%), while 18/71 (25.4%) were to start treatment post Fibroscan. The remainder 25/71 (35.2%) continued to be monitored without antivirals. The various characteristics of the study population are summarised in ***Table 21***.

Table 21: Clinical data of the study patients

Clinical characteristic	
Male	49/71 (69%)
Age (yrs)[a]	46.1 (11.9)
Log_{10}HBVDNA (IU)[a]	5.0 (2.5)
HBe antigen positive[a]	33/71 (46.5%)
BR (µmol/L)[a]	13 (8)
ALT (IU/L)[a]	121 (287)
ALB (g/L)[a]	44 (5)
AST (IU/L)[a]	79 (161)
Platelets ($x10^9$)[a]	217 (50)
ETOH (g/week)[a]	4.8 (10)
Metavir Stage (after 2nd 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 Status	
Treatment current	27/71 (38.0%)
Treatment planned, yet to be initiated	18/71 (25.4%)
Treatment previous, since ceased	1/71 (1.4%)
No antivirals – monitoring only	25/71 (35.2%)

a. Represents the mean value (standard deviation). All other variables report the frequency with percentages in parentheses.

5.3.2 Liver Stiffness Measurement of the study population (objective 1)

There were 71/73 subjects who had a valid TE. The median LSM was 6.9 kPa (IQR 5.3-10.7 kPa). The mean IQR/M ratio was 0.20 (SD 0.24). Sixty six out of 71 (93.0%) scans were reliable, as defined as having an IQR/M ratio of greater of equal to 0.30, when the LSM ≥ 7.1 kPa (55). The findings are shown in

Table 22.

Table 22: Liver Stiffness Measurement characteristics of the study population

Fibroscan parameter	
LSM (kPa)[a]	6.9 (5.3-10.7)
IQR/M ratio[b]	0.20 (0.24)
Valid Scans[c]	71/73 (97.3%)
Reliable scans[c]	66/71 (93.0%)
Success rate[b]	90.4 (14.7)

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

5.3.3 Diagnostic performance of Liver Stiffness Measurement for Fibrosis Stage for all subjects (objective 2)

The Metavir stage of fibrosis and corresponding LSMs for each subject are shown as a boxplot in ***Figure 24.***

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 F0-3 vs F4 was calculated ***(Figure 25).*** All patients (normal and elevated ALT) in the study were included in this analysis.

The AUROCS for diagnosing F≥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.

For F≥1, 2, 3 and 4, 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. The optimal cut-offs for sensitivity was chosen by selecting the coordinate on the AUROC curve that was at least 90% and with best corresponding specificity. The optimal cut-offs for specificity were chosen similarly.

The LSM cut-offs with the best overall diagnostic accuracy for F≥1,2,3 and 4 were: ≥ 6.5 kPa (68.4% sensitivity, 92.9% specificity); ≥7.5 kPa (71.4% sensitivity, 77.8% specificity); ≥ 9.7 kPa (84.2% sensitivity, 92.3% specificity); ≥ 11.9 kPa (100% sensitivity, 84.8% specificity) respectively.

The optimal sensitivity and specificity cut-offs in assessing each fibrosis stage were: F1 ≥ 4.7 kPa (93% sensitivity, 33.6% specificity) and F1 ≥ 6.5 kPa (68.4% sensitivity, 92.9% specificity); F2 ≥ 5.2 kPa (91.4% sensitivity, 36.1% specificity) and F2 ≥ 9.7 kPa (sensitivity 51.4%, specificity 94.4%); F3 ≥ 6.0 kPa (sensitivity 94.7%, specificity 44.2%) and F3 ≥ 9.7 kPa (sensitivity 84.2% and specificity of 92.3%); F4 ≥11.9 kPa (sensitivity 100%, specificity 84.8%) and F4 ≥15.9 kPa (sensitivity 80.0%, specificity 97.0%). ***Table 23*** below summarises the findings.

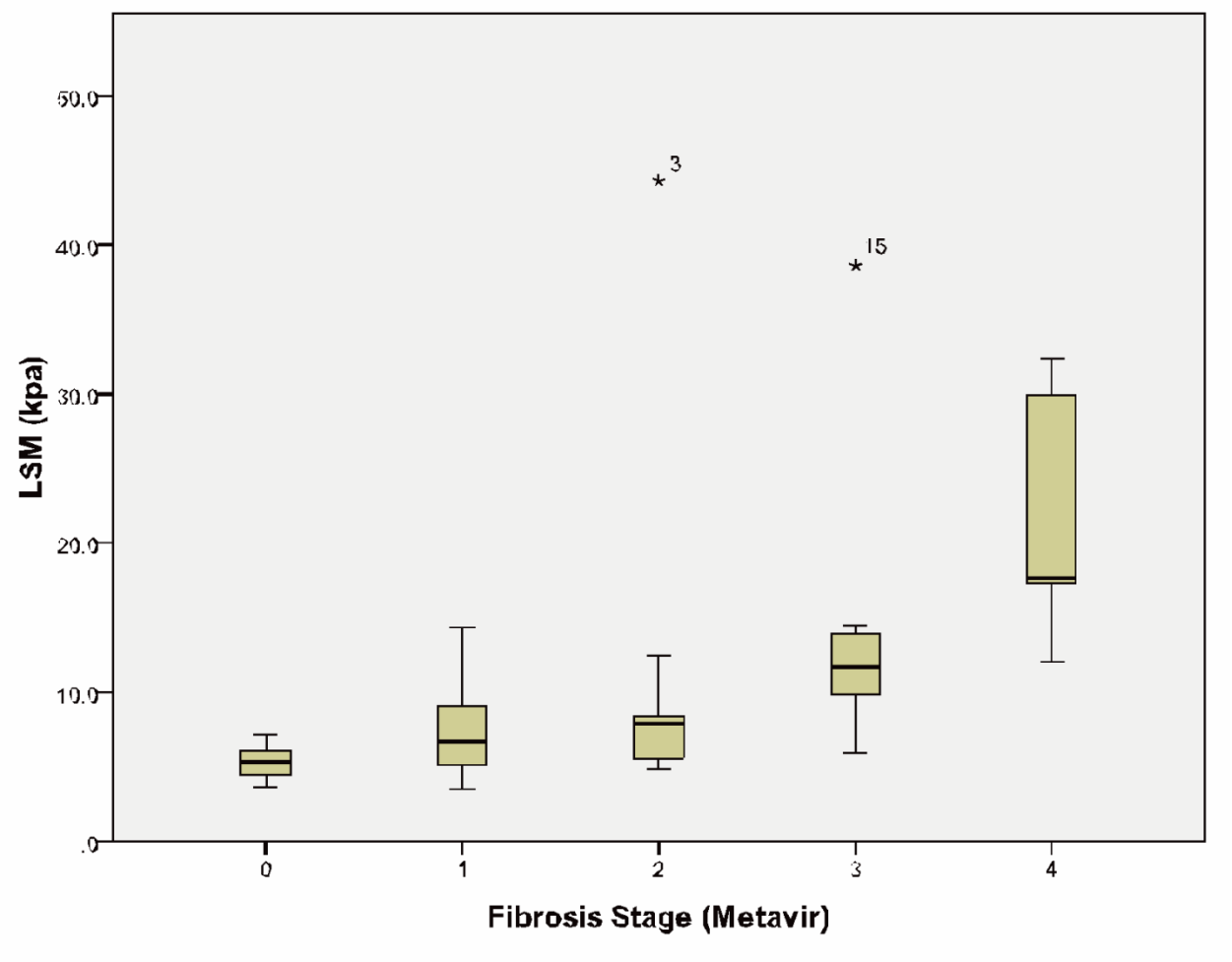


Figure 24: Boxplot of LSM and Fibrosis stage

For each liver biopsy, the fibrosis stage is plotted against the LSM as a boxplot. The box represents the set of LSM scores between the 25-75th percentile, while the horizontal line within the box represents the median. The whiskers extending from the box represent the highest and lowest LSM score.

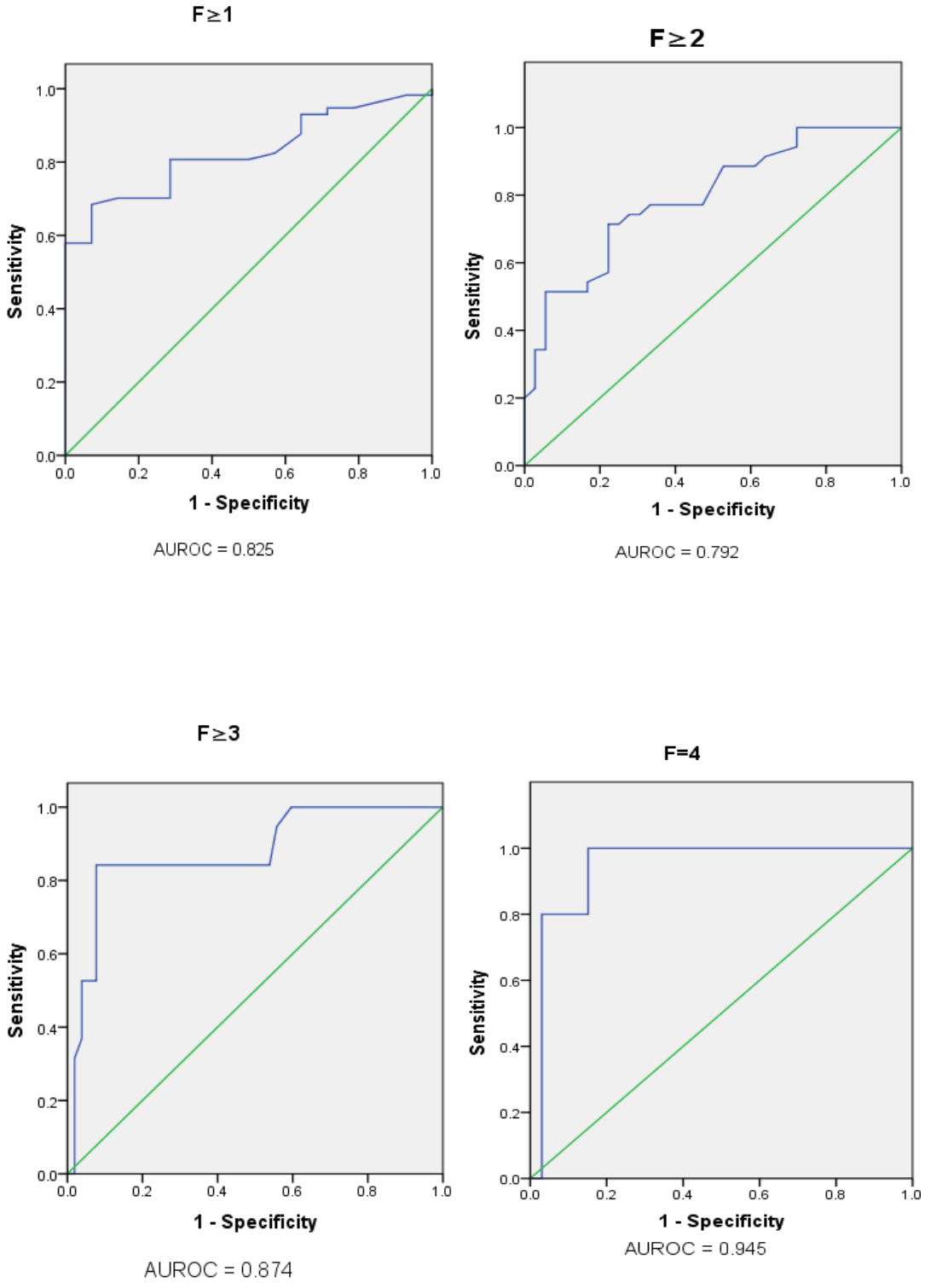


Figure 25: ROC curves of LSMs for F≥1, 2, 3 and 4 for patients with any ALT

Table 23: LSM AUROCs for diagnosing F≥1, 2, 3, 4 and Optimal LSM cut-offs in all subjects

F≥	AUROC	95% Confidence Interval	P value (vs AUROC = 0.5)	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)

5.3.4 Diagnostic performance of Liver Stiffness Measurement for Fibrosis Stage for normal ALT patients (objective 3)

The AUROC curves for LSM diagnosing F0 vs F≥1, F01 vs F≥2, F012 vs F≥3 and F0123 vs F4 was calculated in those with normal ALT and summarised in ***(Figure 26).***

The AUROCS for diagnosing F≥ 1, 2, 3 and 4 were: 0.792 (95% CI 0.632-0.952, p=0.027); 0.762 (95% CI 0.603-0.921, p=0.009); 0.956 (95% CI 0.868-1.000, p<0.001); and 0.909, but was not significant (95% CI 0.801-1.000 95%, p=0.169)

The LSM cut-offs with the best overall diagnostic accuracy for F≥1,2 and 3 were: ≥ 6.6 kPa (57.1% sensitivity, 100% specificity); ≥ 9.7 kPa (44.4% sensitivity, 100% specificity); ≥ 9.7 kPa (88.9% sensitivity, 100% specificity).

The optimal cut-offs for sensitivity and specificity respectively are as follows - F1: ≥ 4.5 kPa (9.6.4% sensitivity, 33.3% specificity) and ≥ 6.6 kPa (57.1% sensitivity, 100% specificity); F2: ≥ 4.7 kPa (100% sensitivity, 25% specificity) and ≥ 9.7 kPa (100% sensitivity 44.4% specificity); and F3: ≥ 6.0 kPa (100% sensitivity, 60% specificity) and ≥ 9.7 kPa (88.9% sensitivity 100% specificity). ***Table 24*** summarizes the findings.

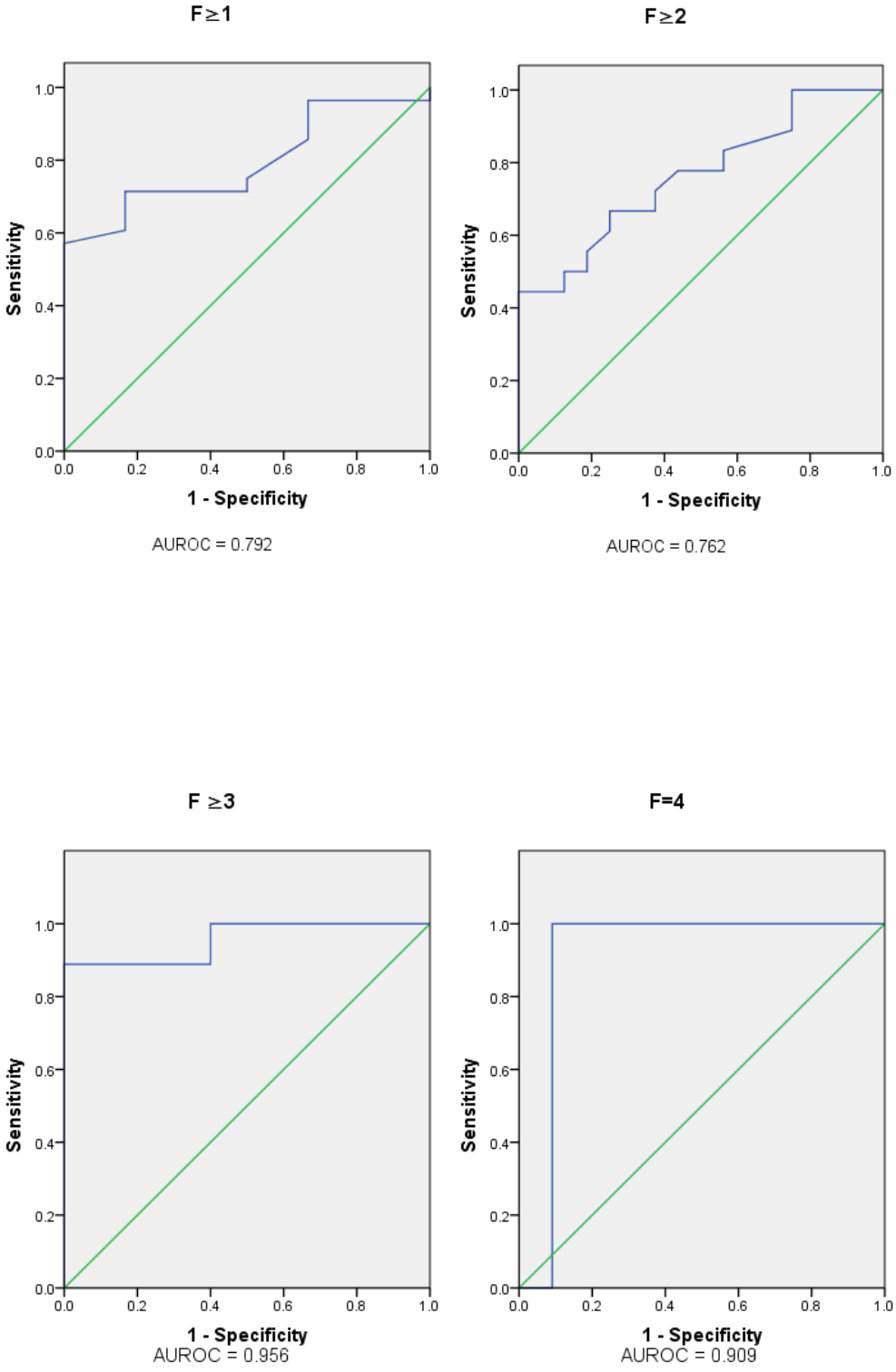


Figure 26: ROC curves of LSM for F≥1, 2, 3 and 4 in patients with normal ALT

Table 24: LSM AUROCs for diagnosing F≥1, 2, 3, 4 and Optimal LSM cut-offs in Normal ALT subjects

F≥	AUROC	95% Confidence Interval	P value (vs AUROC = 0.5)	Best Overall LSM kPa (Sn,Sp%)	Best LSM kPa for > 90% sensitivity, (Sn,Sp%)	Best LSM kPa for > 90% specificity: (Sn,Sp%)
1	0.792	0.632-0.952	0.027	6.6 (57.1, 100)	4.5 (96.4, 33.3)	6.6 (57.1,100)
2	0.762	0.603-0.921	0.009	9.7 (44.4, 100)	4.7 (100, 25)	9.7 (44.4, 100)
3	0.956	0.868-1.000	<0.001	9.7 (88.9, 100)	6.0 (100, 60)	9.7 (88.9, 100)
4	0.909	0.811-1.000	0.169	n/a	n/a	n/a

5.3.5 Diagnostic performance of Liver Stiffness Measurement for Fibrosis Stage for high ALT patients (objective 3)

The AUROC curves for LSM diagnosing F0 vs F≥1, F01 vs F≥2, F012 vs F≥3 and F0123 vs F4 was calculated in those with elevated ALT and summarised in ***(Figure 27)***.

The AUROCS for diagnosing F≥ 1, 2, 3 and 4 were: 0.886 (95% CI 0.751-0.982, p=0.002); 0.847 (95% CI 0.723-0.971, p<0.001); 0.817 (95% CI 0.639-0.994, p=0.003); and 0.939 (95% CI 0.858-1.000, p=0.005).

The LSM cut-offs with the best overall diagnostic accuracy for F≥ 1,2,3 and 4 were: ≥ 6.5 kPa (79.3% sensitivity, 87.5 % specificity); ≥ 7.6 kPa (88.2% sensitivity, 75% specificity); ≥ 10.5 kPa (80% sensitivity, 85.2% specificity); and ≥15.9 kPa (100% sensitivity, 93.9% specificity).

The optimal sensitivity and specificity cut-offs respectively were as follows. F1: ≥ 5.8 kPa (90% sensitivity, 62.5% specificity) and ≥ 7.6 kPa (69% sensitivity, 100% specificity); F2: ≥ 5.8 kPa (100% sensitivity, 40% specificity) and ≥ 12.3 kPa (52.9% sensitivity, 95% specificity); F3: ≥ 6.0 kPa (90% sensitivity, 29.6% specificity) and ≥ 12.5 kPa (70% sensitivity 92.6% specificity); and F4: ≥15.9 kpa (sensitivity 100.0%, specificity 93.9%).

Table 25 summarizes the findings.

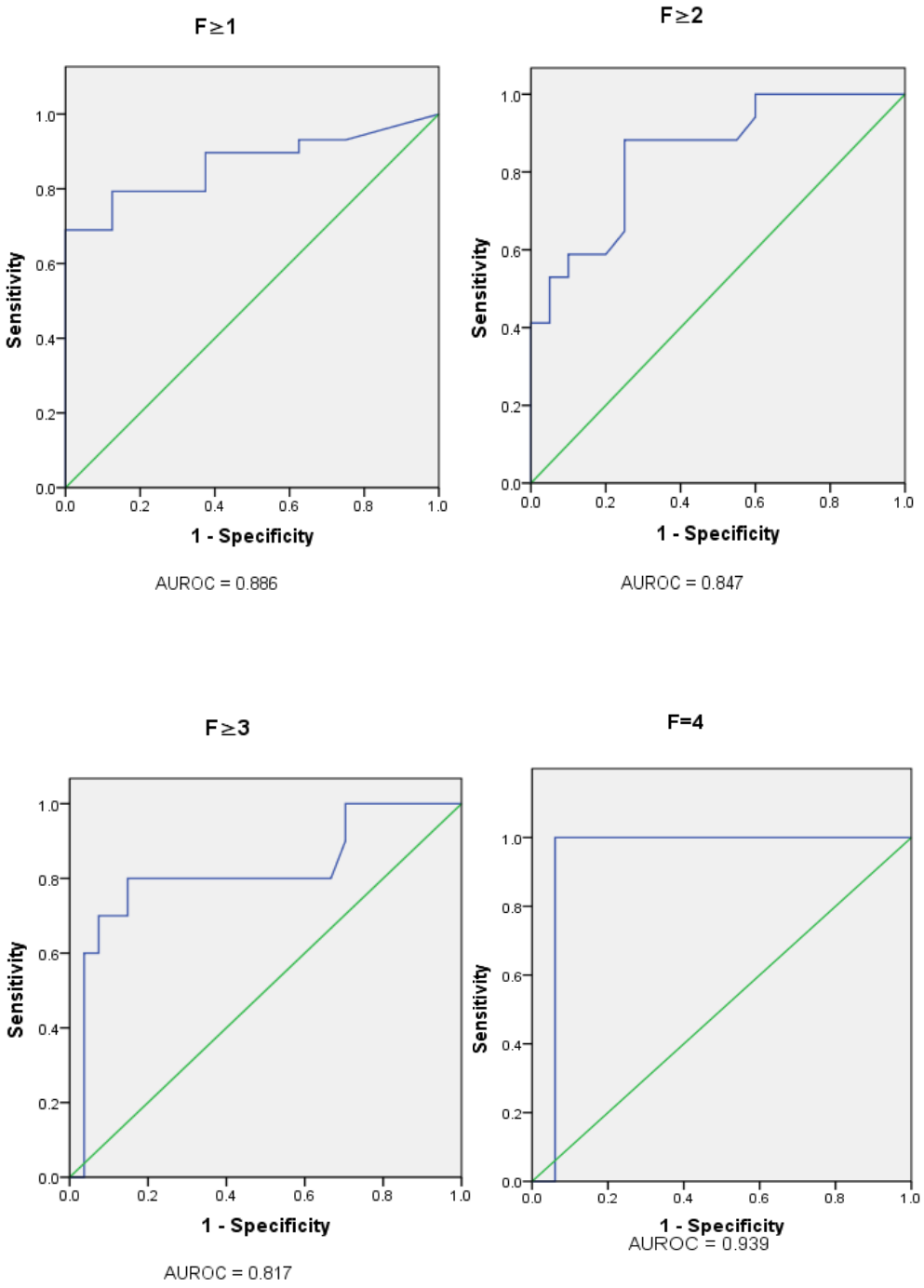


Figure 27: ROC curves of LSMs for F≥1, 2, 3 and 4 in patients with high ALT

Table 25: LSM AUROCs for diagnosing F≥1, 2, 3, 4 and Optimal LSM cut-offs in high ALT subjects

F≥	AUROC	95% Confidence Interval	P value (vs AUROC = 0.5)	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)

5.3.6 Optimal LSM cut offs for moderate and advanced fibrosis (objective 4)

Optimal cut offs were chosen to "rule in" and "rule out" F≥2 and F≥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 positive and negative predictive values were then calculated. The group of cut-offs selected was also specific to whether the ALT was normal or abnormal. This is summarised in ***Table 26***

Table 26: Optimal LSM cut-offs for F≥2 and F≥3 according to normal or elevated ALT

	LSM (kPa) for F≥2 [a]		LSM (kPa)for F≥3 [a]	
	Rule out	Rule in	Rule out	Rule in
Normal ALT subjects	4.7 (100, 25.0, 60.0, 100)	9.7 (44.4, 100, 100, 61.5)	6.0 (100, 60, 47.4, 100)	9.7 (88.9, 100, 90.0, 70.3)
High ALT subjects	5.8 (100, 40.0, 58.6, 100)	12.3 (52.9, 95.0, 90.0, 70.4)	5.8 (100, 29.6, 34.5, 100)	12.5 (70.0, 92.6, 77.8, 89.3)

a. corresponding sensitivity, specificity, positive predictive value and negative predictive value in parentheses

Normal ALT subjects have F2 and F3 ruled out if the LSM ≤ 4.7 kPa (100% NPV) and ≤ 6.0 kPa (100% NPV) respectively. If LSM ≥ 9.7 kPa, then F2 or F3 (100% PPV for both) was ruled in. For subjects whose LSM fell in between these cut-offs (ie. 4.8 – 9.6 kPa for F2 and 6.1 – 9.6 kPa for F3), the PPV ranges between 60-100% (F2) and 47.4-90% (F3), and the NPV ranges 61.5-100% (F2) and 70.3-100% (F3) respectively. Hence these LSM values are within a "grey zone" for F2 and F3, where there is not a high degree of certainty for an accurate diagnosis.

Likewise, the "grey zone" for high ALT subjects was 5.9 – 12.2 kPa for F2 (PPV 58.6-90.0%, NPV 70.4-100%) and 5.9 – 12.4 kPa for F3 (PPV 34.5-77.8%, NPV 89.3-100%) respectively.

In summary:

For F≥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≥2
- 35/71 subjects (49.3%) could not be determined with a high degree of accuracy and are considered in the "grey zone"

For F≥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≥3
- 30/71 subjects (42.3%) could not be determined with a high degree of accuracy and are in the “grey zone”.

5.3.7 Clinical, biochemical, imaging features of cirrhotic subjects compared with LSM (objective 5)

There were 5 patients with histologically proven cirrhosis. The results of their LSM ranged from 12.0 kPa to 32.4 kPa. All 5/5 subjects had a derived F stage of F4 using the LSM. Childs Pugh score was 5 out of 15, and thus a classification of Childs A cirrhosis applied for all 5 subjects. None had any abnormalities of the bilirubin, albumin, INR, or platelet count. Two subjects had recent ultrasound imaging available that showed no features of associated with cirrhosis, such as nodularity, portal vein dilatation, hypersplenism and hepatofugal flow. The results are shown in ***Table 27***.

Table 27: Comparison of clinical, biochemical, ultrasound imaging and LSM scores in histologically proven cirrhosis patients

Subject	Clinical features of cirrhosis	Br (µmol/L)	ALB (g/L)	INR	Childs Pugh score (302)	Platelets ($x10^9$)	US	LSM (kPa)	LSM derived F score
1	None	12	42	1.1	5	173	No cirrhotic features	12.0	4
2	None	12	43	1.1	5	196	n/a	32.4	4
3	None	16	40	1.0	5	164	n/a	17.6	4
4	None	4	44	1.0	5	250	n/a	29.9	4
5	None	13	40	1.0	5	178	No cirrhotic features	17.3	4

5.3.8 Summary of the comparison of the diagnostic performance of non-invasive measures for liver fibrosis (objective 6)

The diagnostic performance of FIB-4, APRI, API, AAR and FI was evaluated by calculating each measure's AUROC for detecting F≥1, 2, 3 and 4. For each non-invasive measure, the cut-offs that corresponded with the best diagnostic accuracy, minimum 90% sensitivity and minimum 90% specificity were calculated for each fibrosis stage.

The full details of these results are presented in sections 5.3.10 – 5.3.14. A summary of the results are described here in this section 5.3.9. The next section of new results is 5.3.15 – Caffeine breath test.

Overall, Fibroscan had the highest numeric AUROC for diagnosing each stage of fibrosis: F1, 2, 3, 4 AUROC=0.825, 0.792, 0.874 and 0.945 respectively. Statistically, the AUROCs for LSM are either superior or equal with the other non-invasive markers across all the F stages. Compared with FIB-4, the AUROC for LSM was significantly higher for F≥1 (p=0.043) and F≥3 (AUROC model for FIB-4 not valid), while there was no statistical difference for F≥2 (p=0.733) and F=4 (p=0.432). Compared to APRI, the LSM AUROC was superior for F≥1 (p=0.027), F≥3 (p=0.023), and F=4 (p=0.016), while there was no difference for F≥2 (p=0.134). For API, FI and AAR, the AUROC comparison was either statistically inferior, inferior by default due to the AUROC being invalid (not significant compared to a random classification model AUROC=0.5) except for API in the diagnosis of F=4 (0.945 vs 0.900) in which the AUROC was numerically lower, but not statistically significant (p=0.492). The next best performing test was the FIB4-I: F1, 2, 4 AUROC= 0.677, 0.711 and 0.912 respectively; followed by the APRI score: F1, 2, 3, 4 AUROC=0.672, 0.698, 0.720 and 0.821 respectively. API only demonstrated accuracy for diagnosing F=4, AUROC=0.900. AAR and FI only had a valid AUROC model for F2: 0.642 and 0.684 respectively and were poor tests overall.

Table 28 summarises the comparison of AUROCS for each stage of liver fibrosis for the non-invasive measures that have been analysed.

Table 28: Comparison of AUROCS of each non-invasive test for fibrosis stage

	LSM	FIB-4	APRI	API	FI	AAR
F ≥1 AUROC	**0.825**	**0.677**	**0.672**	**NS**	**NS**	**NS**
P value Vs LSM	-	0.043	0.027	n/a	n/a	n/a
Vs FIB-4	0.043	-	0.989	n/a	n/a	n/a
Vs APRI	0.027	0.989	-	n/a	n/a	n/a
Vs API	n/a	n/a	n/a	-	n/a	n/a
Vs FI	n/a	n/a	n/a	n/a	-	n/a
Vs AAR	n/a	n/a	n/a	n/a	n/a	-
F ≥2 AUROC	**0.792**	**0.711**	**0.698**	**NS**	**0.642**	**0.684**
P value vs LSM	-	0.733	0.134	n/a	0.034	0.224
Vs FIB-4	0.733	-	0.278	n/a	0.051	0.260
Vs APRI	0.134	0.278	-	n/a	0.420	0.894
Vs API	n/a	n/a	n/a	-	n/a	n/a
Vs FI	0.034	0.051	0.420	n/a	-	0.471
Vs AAR	0.224	0.260	0.894	n/a	0.471	-
F ≥3 AUROC	**0.874**	**NS**	**0.720**	**NS**	**NS**	**NS**
P value Vs LSM	-	n/a	0.023	n/a	n/a	n/a
Vs FIB-4	n/a	-	n/a	n/a	n/a	n/a
Vs APRI	0.023	n/a	-	n/a	n/a	n/a
Vs API	n/a	n/a	n/a	-	n/a	n/a
Vs FI	n/a	n/a	n/a	n/a	-	n/a
Vs AAR	n/a	n/a	n/a	n/a	n/a	-
F =4 AUROC	**0.945**	**0.912**	**0.821**	**0.900**	**NS**	**NS**
P value Vs LSM	-	0.432	0.016	0.492	n/a	n/a
Vs FIB-4	0.432	-	0.040	0.929	n/a	n/a
Vs APRI	0.016	0.040	-	0.137	n/a	n/a
Vs API	0.492	0.929	0.137	-	n/a	n/a
Vs FI	n/a	n/a	n/a	n/a	-	n/a
Vs AAR	n/a	n/a	n/a	n/a	n/a	-

NS: AUROC model not significant, compared to a random classification model
n/a: not applicable. No statistical comparison made as model was not significant

5.3.9 Diagnostic performance of FIB-4 index for Fibrosis Stage (objective 6)

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

The AUROCS for diagnosing F≥ 1, 2 and 4 were: 0.677 (95% CI 0.536-0.817, p=0.042); 0.711 (95% CI 0.660-0.883, p<0.001); and 0.912 (95% CI 0.843-0.981, p=0.002). The AUROC for diagnosing F≥3 was not significant.

The FIB4-I cut-offs with the best overall diagnostic accuracy for F≥ 1, 2 and 4 were: ≥ 1.4733 (43.9% sensitivity, 92.9 % specificity); ≥ 1.3427 (71.4% sensitivity, 77.8% specificity); and ≥ 1.8342 (100% sensitivity, 86.6% specificity).

The optimal sensitivity and specificity cut-offs respectively were as follows. F1: ≥ 0.7298 (90% sensitivity, 28.8% specificity) and ≥ 1.4733 (43.9% sensitivity, 92.9% specificity); F2: ≥ 0.7761 (91.4% sensitivity, 25% specificity) and ≥ 1.6128 (48.6% sensitivity, 91.7% specificity); and F4: ≥ 1.8342 (100% sensitivity, 86.6% specificity) and ≥ 2.174 (60% sensitivity, 91.9% specificity). The summary of the AUROC results and cut-offs are shown in ***Table 29*** and the corresponding AUROC curves are shown in ***Figure 28***.

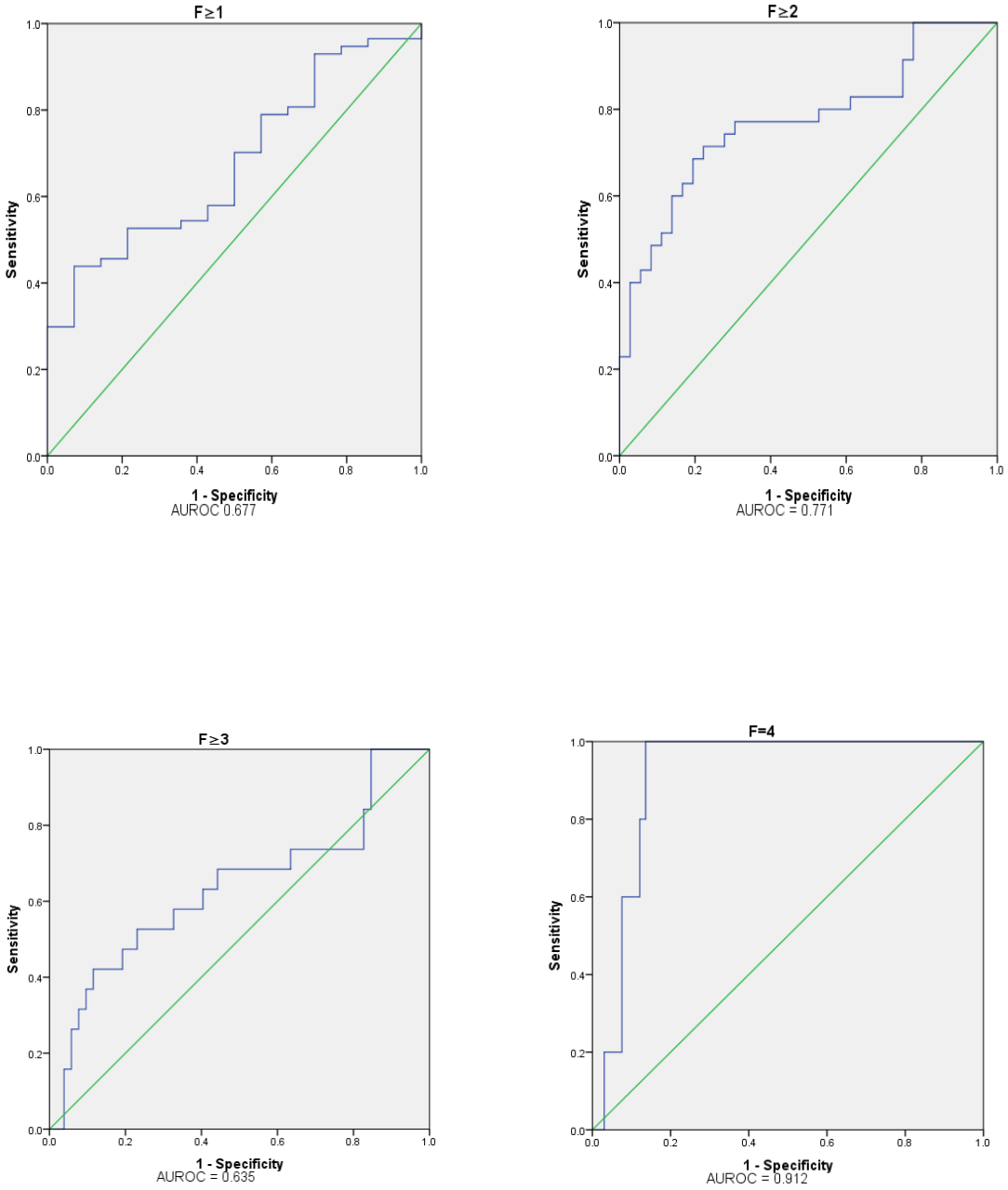


Figure 28: ROC curves of FIB4-Index for F≥1, 2, 3 and 4

Table 29: FIB4-Index AUROCs for diagnosing F≥1, 2, 3, 4 and Optimal cut-offs

F≥	AUROC	95% Confidence Interval	P value (vs AUROC = 0.5)	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)

5.3.10 Diagnostic performance of Aspartate Platelet ratio index (APRI) for Fibrosis Stage (objective 6)

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

The AUROCS for diagnosing F≥ 1, 2, 3 and 4 were: 0.672 (95% CI 0.534-0.809, p=0.048); 0.698 (95% CI 0.576-0.821, p=0.004); 0.720 (95% CI 0.596-0.844, p=0.005); and 0.821 (95% CI 0.712-0.930, p=0.017). T

The APRI cut-offs with the best overall diagnostic accuracy for F≥ 1, 2, 3 and 4 were: ≥ 0.5185 (49.1% sensitivity, 85.7% specificity); ≥ 0.559 (51.7% sensitivity, 83.3% specificity); ≥ 0.559 (68.4% sensitivity, 75.0% specificity); and ≥ 0.6106 (100% sensitivity, 74.2% specificity).

The optimal sensitivity and specificity cut-offs respectively were as follows. F1: ≥ 0.2329 (90% sensitivity, 14.3% specificity) and ≥ 0.559 (43.9% sensitivity, 92.9% specificity); F2: ≥ 0.286 (91.4% sensitivity, 33.3% specificity) and ≥ 1.174 (17.1% sensitivity, 91.7% specificity); F3: ≥0.3493 (90% sensitivity, 42.3% specificity) and ≥ 1.2929 (15.8% sensitivity, 90.4% specificity); and F4: ≥ .6106 (100% sensitivity, 74.2% specificity) and ≥ 1.3429 (20% sensitivity, 90.9% specificity). The ROC curves for APRI for each stage of fibrosis is shown in ***Figure 28***. The summary of the AUROC results and cut-offs are shown in

Table 30.

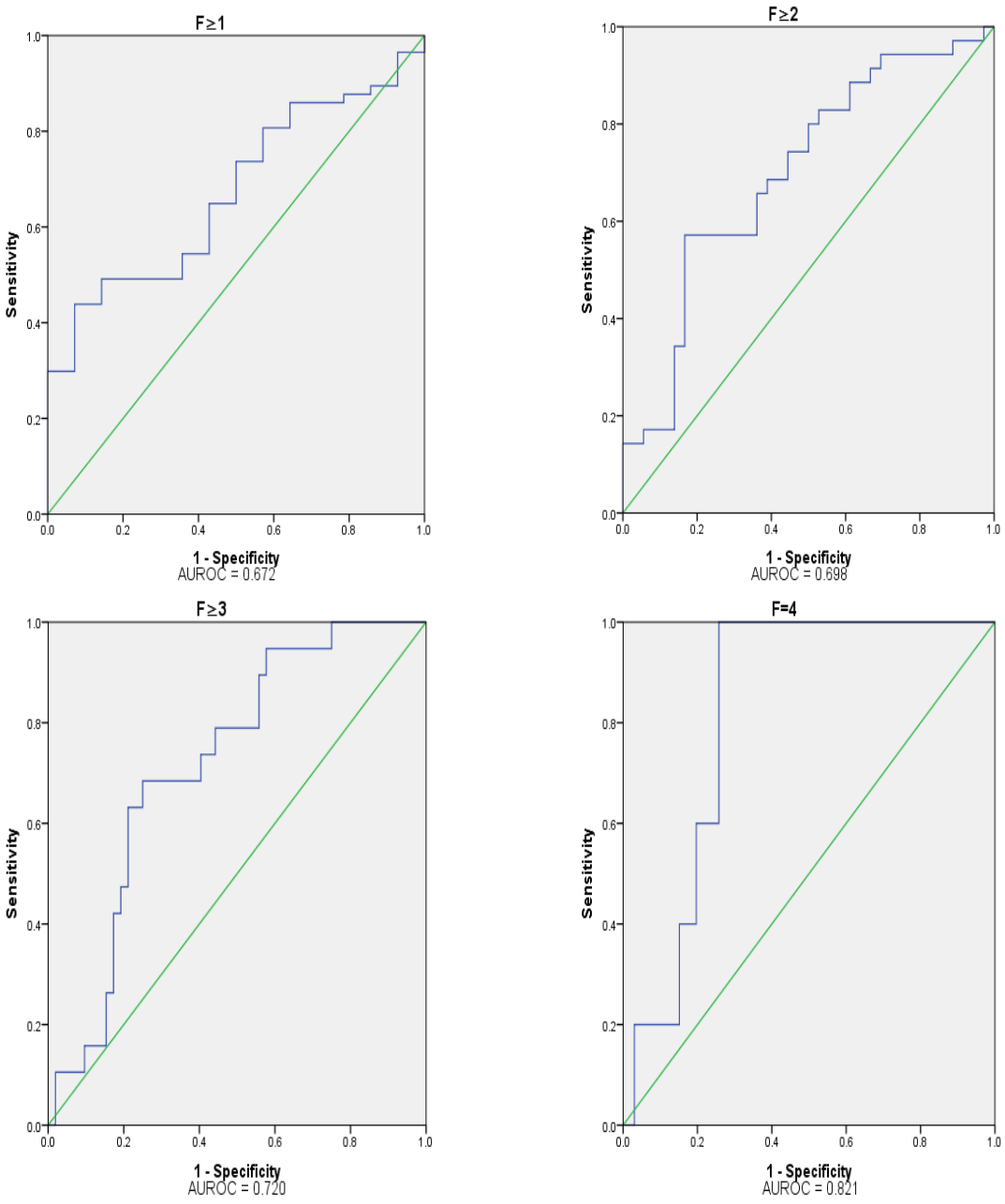


Figure 29: ROC curves of APRI for F≥1, 2, 3 and 4

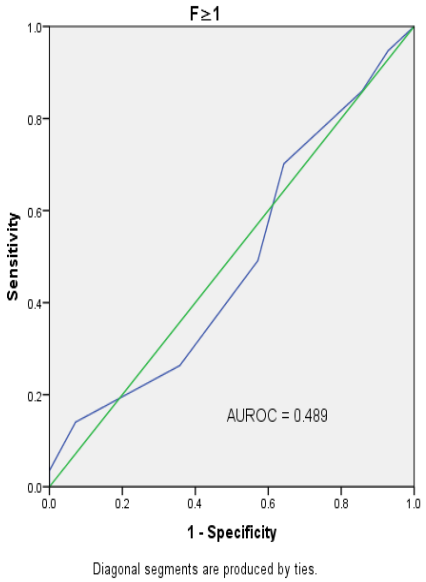
Table 30: APRI AUROCs for diagnosing F≥1, 2, 3, 4 and Optimal cut-offs

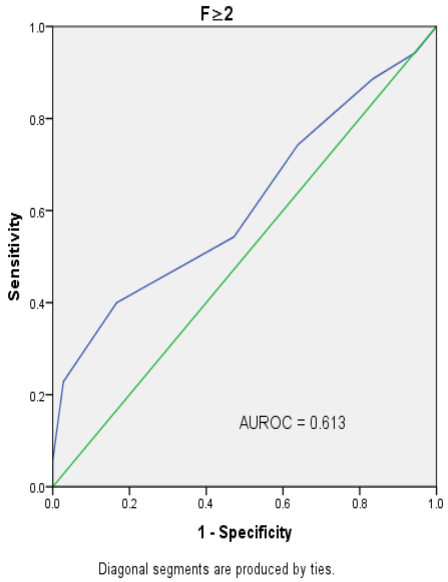
F≥	AUROC	95% Confidence Interval	P value (vs AUROC = 0.5)	Best Overall APRI (Sn,Sp%)	Best APRI for > 90% sensitivity, (Sn,Sp%)	Best APRI for > 90% specificity: (Sn,Sp%)
1	0.672	0.534-0.809	0.048	0.5185 (49.1, 85.7)	0.2329 (90.0, 14.3)	0.559 (43.9, 92.9)
2	0.698	0.576-0.821	0.004	0.559 (57.1, 83.3)	0.286 (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)

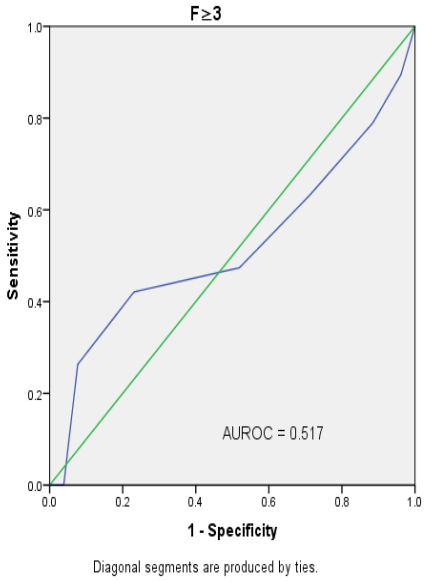
5.3.11 Diagnostic performance of Age Platelet Index (API) for Fibrosis Stage (objective 6)

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

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 was API≥5 (100% sensitivity, 77.3% specificity). The optimal cut-offs for sensitivity and specificity respectively were API≥5 (sensitivity 100.0%, specificity 77.3%) and API≥ 6 (sensitivity 60.0%, specificity 90.9%). The ROC curves for API for each stage of fibrosis is shown in ***Figure 30***. The summary of the AUROC results and cut-offs are shown in ***Table 31***.







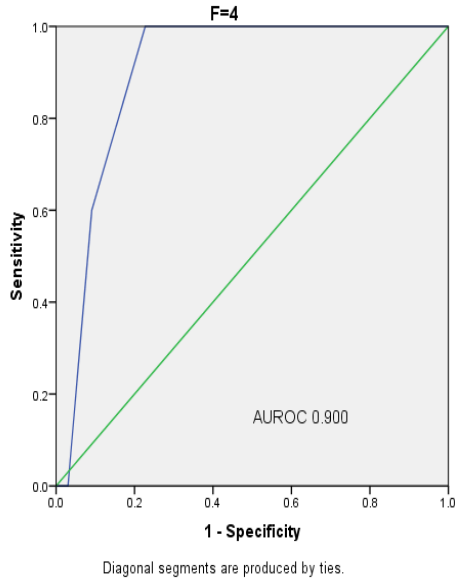


Figure 30: ROC curves of API for F≥1, 2, 3 and 4

Table 31: API AUROCs for diagnosing F≥1, 2, 3, 4 and Optimal cut-offs

F≥	AUROC	95% Confidence Interval	P value (vs AUROC = 0.5)	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)

5.3.12 Diagnostic performance of Fibrosis Index (FI) for Fibrosis Stage (objective 6)

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

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

Only F≥2 had a significant AUROC, which was 0.642 (95% CI 0.511-0.774, p=0.039). The best overall cut-off was ≥ 1.08 (91.4% sensitivity, 47.2% specificity). The optimal cut-offs for sensitivity and specificity respectively was ≥ 1.08 (91.4% sensitivity, 47.2% specificity) and ≥ 2.105 (22.9% sensitivity, 91.7% specificity). The ROC curves for FI for each stage of fibrosis is shown in ***Figure 31***. The summary of the AUROC results and cut-offs are shown in ***Table 32***.

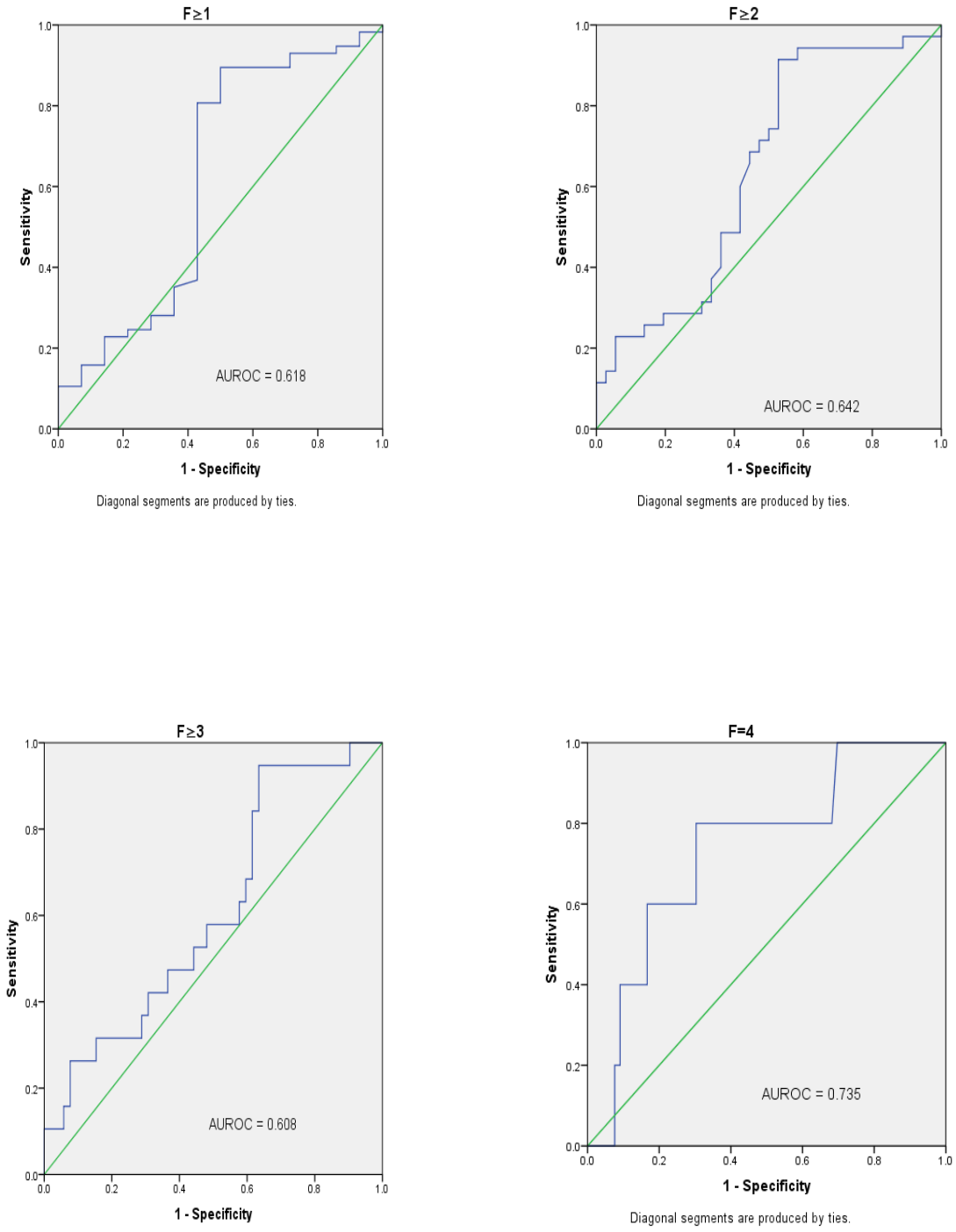


Figure 31: ROC curves of FI for F≥1, 2, 3 and 4

Table 32: FI AUROCs for diagnosing F≥1, 2, 3, 4 and Optimal cut-offs

F≥	AUROC	95% Confidence Interval	P value (vs AUROC = 0.5)	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

5.3.13 Diagnostic performance of Aspartate aminotransferase and alanine aminotransferase ratio (AAR) for Fibrosis Stage (objective 6)

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

Only F≥2 had a significant AUROC, which was 0.684 (95% CI 0.559-0.808, p=0.008). The best overall cut-off was ≥ 0.6954 (77.1% sensitivity, 63.9% specificity). The optimal cut-offs for sensitivity and specificity respectively were ≥ 0.5353 (91.4% sensitivity, 30.6% specificity) and ≥ 1.0571 (71.4% sensitivity, 91.7% specificity). The AUROCS for AAR diagnosing F≥1, 3 and F=4 were not significant. ***Table 33*** summarises the findings and ***Figure 32*** show the ROC curves.

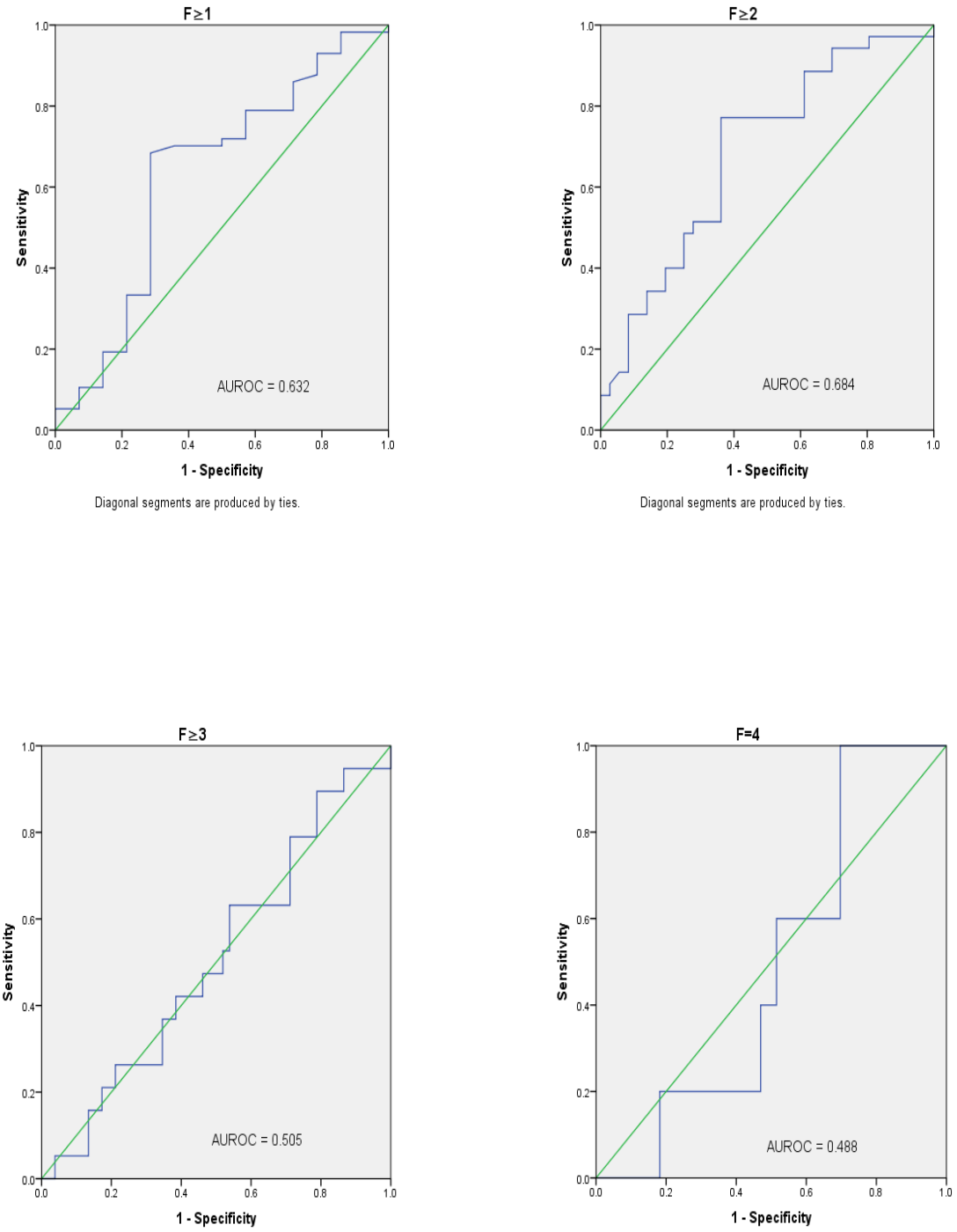


Figure 32: ROC curves of AAR for F≥1, 2, 3 and 4

Table 33: AAR AUROCs for diagnosing F≥1, 2, 3, 4 and Optimal cut-offs

F≥	AUROC	95% Confidence Interval	P value (vs AUROC = 0.5)	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

5.3.14 Diagnostic Performance of Caffeine breath test (objective 7)

Each of the 7 subjects whom received a caffeine breath test had their histological fibrosis stage compared along with the derived fibrosis stage for each non-invasive test. Analysis in section 5.3.13 and 5.3.14 showed AAR and FI to be failed tests due to poor AUROCs and so these were omitted. The derived fibrosis stage for TE, APRI, FIB4-I and API, would be determined using optimal specificity (>90%) cut-offs that were reported earlier in the previous sections of this chapter. For CBT, this was derived from previous studies (281, 301). Only F4 could only be determined with the API. ***Table 34*** summarises the comparisons.

The small sample size prevented statistically evaluation. A descriptive analysis of how well each non-invasive measure classifies F stage was performed.

Caffeine breath test correctly classified 4/7 cases, with 1 case misclassified by 1 stage and 2 cases misclassified by 2 stages. Transient Elastography correctly classified 3/7 cases misclassifying 1 case by 1 stage, and 3 cases misclassified by 2 stages. FIB-4 had 2/7 correct classifications, misclassifying 2 cases by 1 stage, and 3 cases by 2 stages. No cases were correctly classified by APRI but 2, 3 and 2 cases were misclassified by 1, 2 and 3 stages respectively.

Table 35 outlines the correct and incorrect classifications.

Table 34: Comparison of fibrosis stage assessment in subjects with caffeine breath test

Subject #	Histological F stage (Metavir)	TE		CBT		APRI		FIB-4		API	
		LSM (kpa)	F[a]	Caffeine Clr	F[a]	Ratio	F[a]	index	F[a]	Index	F4
1	1	9.0	1	1.88	1	0.308	0	1.17	0	4	N
2	2	5.9	0	1.06	4	0.467	0	1.64	1	4	N
3	2	5.0	0	1.69	2	0.646	1	2.39	4	7	Y
4	2	6.7	1	0.75	4	0.342	0	1.07	0	3	N
5	2	5.3	0	1.76	2	0.493	0	0.81	0	5	Y
6	4	12.0	4	0.71	4	0.617	1	2.40	4	5	Y
7	4	32.4	4	1.32	3	0.612	1	1.88	4	2	N

a. The Fibrosis stage was derived based on the value of the corresponding non-invasive test. For TE, APRI, FIB4-I and API, these were derived from results in this study. For CBT, this was derived from previous studies (283, 303). Only F4 could be determined with the API.

Table 35: Misclassification of fibrosis stage by non-invasive markers compared with histology

	CBT	LSM	APRI	FIB-4	API
Correctly classified	4	3	0	2	1 out 2 cases F4 correct
Misclassified by 1 stage	1	1	2	2	
Misclassified by 2 stages	2	3	3	3	
Misclassified by 3 stages			2		

5.3.15 Comparison of liver fibrosis staging between the initial histological assessment with a second assessment (objective 8)

In the initial histology assessment of fibrosis of 71 cases, the number of F0, F1, F2, F3 and F4 were 12, 18, 24, 12 and 5 respectively.

The second assessment on 54 of these, the F0 increased from 12 to 14; F1 decreased from 18 to 12; F2 increased from 24 to 26; F3 increased from 12 to 14; F4 remained the same at 5 cases. This is shown in ***Table 36.*** In the remainder 17 cases, where slides were not available for review, the result of the initial assessment was retained.

Table 36: Frequency of fibrosis stage as assessed by Initial and second histological examination

Metavir Score	Number of cases on Initial Examination	Number of cases after second examination
0	12	14
1	18	12
2	24	26
3	12	14
4	5	5

A total of 24 of 54 (44.4%) cases were restaged a different fibrosis stage. A difference of one stage was found in 19/54 (35.2%), two stages in 2 (3.7%) cases and three stages was 1 case (1.9%). Two cases (3.9%) cases were considered inadequate samples on re-evaluation.

The intraclass correlation coefficient was calculated using the Cohen's K to determine the level of agreement between original and reference histological assessment of the Metavir Fibrosis score. There was only moderate agreement: $K = 0.457$, $p < 0.001$.

Construction of 2x2 tables for the initial and review histological assessment was done. The number of true positives, true negatives, false positives and false negatives was entered into the 2x2 table. The initial histological assessment was compared against the second assessment being used as the reference standard as discussed in section 5.2.4. The sensitivity, specificity, positive predictive value and negative predictive value of the first histological assessment for each stage of fibrosis were then derived.

The original histological assessment had high specificity (95%, 91.7% and 98%), for F0 (38/40: true negatives out of total negatives), F3 (44/48: true negatives out of total negatives) and F4 (49/50: true negatives out of total negatives) respectively. Specificity for F1 was good (84.4%: 27/32 true negatives out of total negatives), but only fair for F2 (77.8%: 35/45 true negatives out of total negatives).

Sensitivity for F2 was good at 88.9% (8/9 true positives over total positives). However, sensitivity for F0, F1, and F4 were poor at 64.3% (9/14 true positives over total positives), 54.5% (12/22 true positive over total positives) and 66.7% (2/3 true positives over total positives) respectively. The sensitivity for F3 was extremely poor at 16.7% (1/6 true positives over total positives. ***Table 37*** summarises the diagnostic performance of the original assessment against the reference histological assessment.

Table 37: Diagnostic performance of initial histological assessment using the 2nd assessment as reference

	F0	F1	F2	F3	F4
Sensitivity (%)	64.3	54.5	88.9	16.7	66.7
Specificity (%)	95.0	84.4	77.8	91.7	98.0
Positive Predictive value (%)	81.8	70.6	44.4	20.0	66.7
Negative Predictive value (%)	88.4	73.0	97.2	89.8	98.0

5.3.16 Summary of the Main Findings

- Successful Fibroscan can be performed to measure liver stiffness in hepatitis B patients with 97% having valid scans, and of those 93.7% having reliable scans.
- LSM had excellent diagnostic performance for F≥3 and F=4. AUROC curves for F≥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≥2 and F≥3 can be diagnosed or excluded with a very high degree of certainty (>90%) in 49.3% and 57.7% respectively.
- LSM was able to diagnose 5/5 cases of biopsy proven F4 cirrhosis, in which these patients had no clinical or imaging features.
- Fibroscan was the best non-invasive measure for every stage of fibrosis when compared with FIB-4I, APRI, API, AAR and FI.
- In a small sample of 7 patients, caffeine breath tests performed well compared with other non-invasive tests.
- There was only moderate agreement between the first and second histological assessments of Metavir fibrosis stage (K =0.457, p<0.001) There was a disagreement in 44.4% of liver biopsy assessments, with a difference of classification by 1 stage being the most common.

5.4 DISCUSSION

5.4.1 Transient Elastography in Chronic hepatitis B patients was performed reliably in 93.7% of cases

A valid LSM was obtained using TE in 97.3% of cases, with 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, having performed more than a combined 900 scans, further supporting other studies where operator experience is important in obtaining a high rate of valid, reproducible and reliable LSM's (53).

5.4.2 Transient Elastography has good to excellent diagnostic performance for fibrosis stage in chronic hepatitis B

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).

Marcellin 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) (87).

Many other studies have since assessed the diagnostic performance of TE in hepatitis B patients. A 2013 systematic review reported the range of AUROCS for F≥2 was 0.78-0.87; F≥3 0.87-0.92 and F4 0.80-0.96 (303). A separate meta-analysis in 2012 identified 18 studies with 2772 patients. The pooled AUROCS for F≥2, F≥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 (304). Our results are consistent with those observed in these reviews and meta-analysis.

The performance of TE is better for detecting advanced fibrosis and cirrhosis compared to mild and moderate fibrosis. The AUROCs are higher as the degree of fibrosis increases, suggesting that TE is more accurate for latter stages of fibrosis. This was seen in the results of our study, as well as the systematic review and meta-analysis (303, 304). One explanation may be that LSM has stronger correlation with pericellular fibrosis (280) compared with periportal and perivenular fibrosis. Pericellular fibrosis occurs more in latter stages of fibrosis when there is formation of septa.

Another explanation perhaps is that LSM is a better reflection of the volume of fibrosis, rather than the stage of fibrosis. This, along with its implications, is discussed in 5.4.11.

5.4.3 Patients with elevated ALT have higher LSM cut-offs by factor of 1.3

Patients who have elevated ALT have higher optimal LSM cut-offs to diagnose the same stage of fibrosis. The optimal cut-offs for F≥3 and F4 for all patients (normal and high ALT included) were 9.7 kPa and 11.9 kPa respectively. For high ALT patients, the optimal LSM cut-offs for F≥3 was 10.5 kPa and F4 15.9 kPa. Therefore, 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 (280).

Wong's study had reported a reduction in the diagnostic performance of TE in patients with high ALT compared to patients with normal ALT (280). In our study, we did not observe any consistent difference. The AUROC was superior for F≥1, F≥2 and F4 but inferior for F≥3 in patients with high ALT. Our study only had 37 subjects with high ALT and 34 subjects with normal ALT. The lack of conclusive findings may be due to the small study population.

5.4.4 Variation in optimal LSM cutoffs

Defining the cut-offs for fibrosis stages has remained an area of ongoing debate. The lack of universally accepted cut-offs has significant implications when interpreting TE results (305). In our study the optimal LSM cut-off values for F≥2, 3 and 4 were 7.5 kPa, 9.7 kPa and 11.9 kPa respectively. Other studies have reported cut-offs for F≥2, 3 and 4 that range: 5.2 -8.5 kPa, 8.1 - 10.5 kPa and 10.3 – 12.9 kPa respectively (306-311).

The variation in LSM cut-offs is likely due to differences in study factors. There are differences in the population studied, quality of the liver biopsies obtained, histology assessment and TE operator expertise.

This issue is not unique to TE and exists for many clinical investigations. For instance, it is common for the ALT level, to have slightly different reference ranges when reported by different laboratories. Over time, it is common practice for laboratories to revise reference values as updated data becomes available. The situation is analogous for LSM cutoffs diagnosing fibrosis stage and so the approach should be no different. LSM cut-offs that are locally derived are most likely to accurately represent the fibrosis stage for that particular population and should be applied where available. However, since liver biopsy studies are not easily performed, local experts should determine which cut-offs to adopt from available studies, taking into account factors such as the demographic and data quality.

There is no published data for Australian patients for LSM cut-offs in CHB fibrosis. Although there has been several studies already that compare TE with liver biopsy in CHB patients, there is still value in determining cut-offs in our local population. These are likely to be the most accurate for our local clinical practice.

5.4.5 Dual liver stiffness measurement cut-offs for diagnosing moderate and advanced fibrosis and the grey zones between cut-offs.

The most rational way to use noninvasive methods for diagnosing liver fibrosis is to determine the range of values where there is high accuracy, and to reserve the use of liver biopsy only when the accuracy is not at an acceptable level. To maximize the use of TE, a high sensitivity cut-off (which has a high NPV and effectively rules out the stage of fibrosis in question), and a high specificity cut-off (which has a high PPV and effectively rules in the stage of fibrosis in question), should be applied. Only LSM scores that fall in between low and high cutoffs, the so called "grey zone" cannot be classified with a high enough level of accuracy, and thus a liver biopsy should be considered.

Due to the fact that ALT affects LSM cut-offs, we devised separate sets of dual cut-offs and corresponding grey zones for patients with normal ALT and high ALT. For normal ALT, the grey zones for F≥2 were 4.8 - 9.6 kPa and F≥3 was 6.0 - 9.7 kPa. For high ALT patients, the grey zone for F≥2 was 5.8 - 12.3 kPa and F≥3 was 5.8 - 12.5 kPa. Patients with LSM scores that fall within these grey zones means it is not possible to rule in or rule out F≥2 or F≥3 with a high degree of certainty. In the study cohort, 50.7% and 42.3% of patients fell within the grey zone range for F2 and F3 respectively. Thus TE was able to rule in or rule out F2 and F3 accurately in 49.3% and 57.7% of patients respectively in the study, avoiding the need for a liver biopsy.

Other studies have also recognized the need for dual cut-offs. Chan et al (88) proposed grey zone cutoffs for F3 in normal ALT and high ALT patients to be 6.0 - 9.0 kPa and 7.5 - 12.0 kPa respectively. TE was able to rule in or rule out F3 in 62% and 58% of normal and abnormal ALT subjects respectively (88).

In our study, we used LSM cut-offs with slightly higher sensitivity and specificity, leading to the grey zones being wider. Hence we had a slightly greater proportion of subjects whose LSM scores were classified in the grey zone. If the grey zones in Chan's paper were applied, the proportion of those who can be accurately ruled in/out for F3 increases from 57.7% to 67.6%, but at the cost of lower sensitivity and specificity.

TE was able to diagnose F3 accurately in 57.7%, which meant 42.3% of subjects did not require a liver biopsy. In clinical practice, the proportion of CHB patients avoiding a liver biopsy may actually be even greater. Guidelines from AASLD, EASL and APASL recommend antiviral therapy when there is persistently high ALT (259, 267, 271, 272) which is irrespective of the fibrosis stage. A grey zone LSM would cause some uncertainty regarding whether these patients truly have F3, but performing a liver biopsy for clarification would not change the indication for antiviral therapy. Since liver biopsy is unlikely to change management, many clinicians may feel it is unnecessary in these circumstances.

The instances where a liver biopsy is still needed to determine need for antiviral therapy after TE has been performed was investigated by this author with other local researchers. After determining the number of patients with LSMs within Chan's grey zones and excluding those with persistently elevated ALT, only 9/47 (19.1%) patients would still require a liver biopsy to clarify whether antiviral treatment was indicated. TE assessment allows for the majority of CHB patients to avoid having a liver biopsy which means a reduction in associated costs estimated to be $AUD 74 214 per annum for the local institution (312).

5.4.6 Diagnosing compensated cirrhosis with transient elastography

All cases (5 out of 5) were diagnosed with cirrhosis by TE while having no overt clinical, biochemical or imaging abnormalities. These findings confirm the excellent diagnostic capability for F4 that has been reported by this study (AUROC = 0.945) and others (303). Identifying those with cirrhosis using TE will allow for the monitoring of the severe complications of HCC and varices and reduce mortality.

5.4.7 Transient Elastography compared with FIB4-I, APRI, API, AAR and FI

APRI, API, AAR and FI were originally developed for chronic hepatitis C patients (204, 288, 290, 313), while FIB4-I was first developed for a HCV-HIV co-infected cohort (287) . The diagnostic accuracy of TE was overall equal or superior compared to these other noninvasive measures for fibrosis stages 1 through to 4. This was followed by FIB-4 and APRI. The API was only able to diagnose F4, while AAR and FI were poor tests for fibrosis or cirrhosis in CHB- see ***Table 28***.

An early validation study for FIB4-I in 668 CHB patients demonstrated good results, with the reported following AUROCS: F≥2 = 0.865, F≥3 = 0.910 and F4 = 0.926 (314). However, these superior results have not been replicated. A meta-analysis of 34 studies and 8855 patients assessed the performance of FIB4-I to have AUROCS up to 0.18 lower - F2: AUROC =0.82 (95% CI 0.77-0.86), F3: 0.73 (95% CI 0.66 - 0.80) and F4: 0.84 (95% CI: 0.77-0.92)(315). Our study results for the diagnostic performance of FIB4-I are similar with the meta-analysis results – see ***Table 29.***

The aforementioned meta-analysis also examined the diagnostic performance of APRI for fibrosis stage in CHB patients. The reported AUROCs for F2, 3 and 4 were: 0.74 (95% CI 0.70-0.78); 0.78 (95% CI 0.75-0.82) and 0.73 (95% CI 0.68-0.80) respectively (315). Our study findings for the performance of APRI are also similar with the meta-analysis - see ***Table 30.***

The use of API in CHB patients has a reported diagnostic performance for F4: AUROCs= 0.77-0.93 (285, 314, 316). Only 2 other studies have reported that API has diagnostic validity to assess other fibrosis stages: AUROC= 0.77 for F2 (317) and AUROC = 0.90 for F3 (314). Our results which show API is only valid to diagnose F4 is consistent with the literature – see ***Table 31.***

AAR is an attractive marker because of its simplicity. However, a study of 1543 CHB patients reported AAR to have poor performance in diagnosing fibrosis. The AUROCs for F2, 3 and 4 were 0.57, 0.62 and 0.64 respectively (318). Another study of 406 CHB patients reported slightly better performance: AUROC = 0.772 for F4 (319). But another study of 380 CHB patients reported that AAR to has no significant relationship with the degree of fibrosis and was concluded to be not useful in estimating fibrosis stage (320). AAR was also found to be a poor test in our study – see ***Table 33.***

At the time of writing, no other studies have described the use of FI for diagnosing fibrosis in chronic hepatitis B. Our findings suggest that it is not a useful test for fibrosis –see ***Table 32.***

5.4.8 Caffeine breath test shows promising results

CBT performed well when compared to TE and the other non-invasive measures in our small cohort. However, the utility of CBT is restricted by being relatively more labour intensive, time consuming and requiring a supply of radioactive ^{13}C. CBT can have false positives due to recent smoking and intake of caffeine containing products. TE is a more practical tool, and so CBT may need to demonstrate a significantly superior performance either alone or in combination with TE in order to gain widespread acceptance.

5.4.9 Transient Elastography compared with other prominent non-invasive assessment of fibrosis in chronic hepatitis B patients – an updated overview

There have been several recent developments for non-invasive markers of fibrosis in CHB. Of all these, the Fibrotest (FT) has been the second most widely studied behind TE. A meta-analysis by Poynard in 2011 (321) 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 hepascore and other serum based markers found that FT was inferior and only had an AUROC of 0.72 compared with hepascore AUROC 0.83 (322). 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≥2, 3 and 4 respectively. This study reported combining the 2 markers by multiplying the FT and LSM produced the best AUROCs: 0.941, 0.931, and 0.929 for F≥2, 3, and 4, respectively (308).

Of the remaining potential non-invasive measures, a review by Chen (303) attempted to identify quality markers by setting criteria that it must have independent validation, and have a diagnostic accuracy of AUROC > 0.85. Tests that fulfilled these criteria include the Forns index, Hepascore, Fibrometer, Zeng Index, Hui Index, API and FIB4-I. The latter 2 are inferior to TE which was also demonstrated in our study (see 5.4.9). For the other markers, the range of their reported diagnostic performance for each stage of fibrosis are as follows: F≥2 (AUROC: 0.72-0.81), F≥3 (AUROC: 0.75-0.89) and 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.

5.4.10 Variability in histological assessment of fibrosis staging

Our study highlights interobserver variability that can occur in liver histology assessment. Using a second histological assessment as the reference, the first assessment classified almost half (44.4%) differently. The intra-class correlation coefficient was found to only be at a moderate level (k=0.457, P< 0.001).

Inter-observer variability has been reported widely in the literature and is estimated to account for 15-33% of the variation (41, 42, 45). Higher levels of variability (49.9%) have been demonstrated when comparing fibrosis stage assessment between non specialist community pathologist and academic hepatopathologists in chronic hepatitis C patients (46). We report similar findings in this hepatitis B cohort.

The primary issue of misclassification in our study was under-staging disease especially at the advanced stages F3 and F4. There were 6 true cases of F3, but 5 assessments made an underestimation. There were 3 true cases of F4, with 1 assessment making an underestimation. Despite our overall small number of cases, and the lack of data on biopsy quality, it does highlight the limitation in using liver biopsy as a gold standard.

5.4.11 Problems with comparing the LSM against liver biopsy

The diagnostic performances of non-invasive markers are compared against the reference standard of liver biopsy. However, its inaccuracy subsequently limits the accuracy of any non-invasive markers that is being compared against biopsy for the staging of fibrosis. Our study has only highlighted the degree of inter-observer variability that may occur. There can also be inaccuracy stemming from inadequate biopsy length (43) and patchy disease (41). Because of these limitations, some experts feel that an AUROC >0.90 cannot be truly achieved (200), and that that non-invasive fibrosis tests with an AUROC of 0.85-0.90 may be as good as liver biopsy for staging liver fibrosis (323).

Another issue is that the LSM, which is a scale variable, awkwardly compares with fibrosis stage, which is a categorical variable. While this comparison is commonly performed, some claim this to be flawed (324). Fibrosis staging is a qualitative morphological assessment of the distribution of fibrosis. The mix of features in the qualitative description does not include any component that assesses the quantitative amount of fibrosis in each stage. Although the quantity of fibrosis increases, the criteria for the diagnosis of increasing stages depend primarily on where fibrosis is located and distributed within the hepatic lobules. The fibrosis quantity for progressive fibrosis stages do not increase linearly or in a proportionate manner. For instance, stage 4 fibrosis does not imply twice the fibrosis quantity of stage 2 fibrosis.

This can be well illustrated by the relationship between Ishak fibrosis stage and the collagen proportionate area (CPA) which is a measure of the quantity of fibrosis. The CPA and Ishak score clearly do not have a linear relationship (see ***Figure 33***). They are related, but in the end are different evaluations.

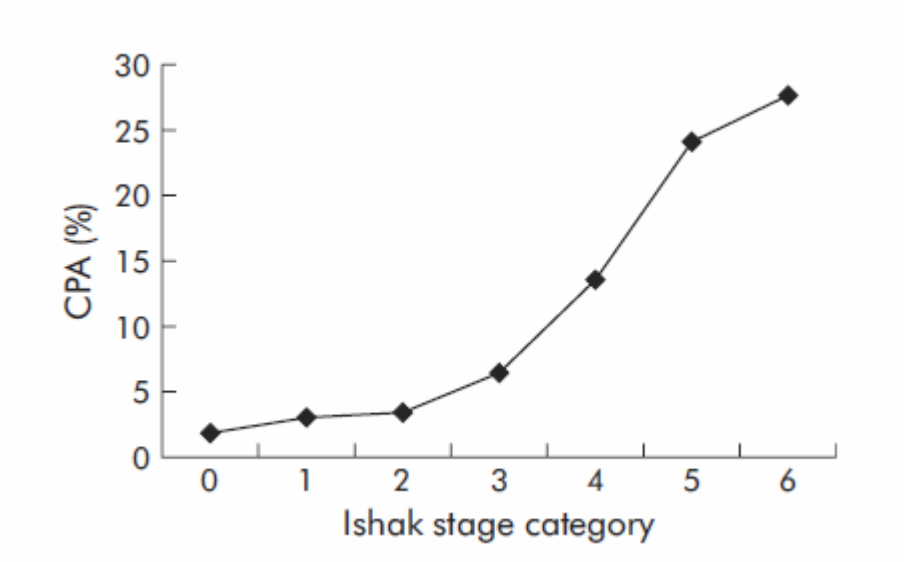


Figure 33: Collagen Proportionate Area and Ishak stage

From Standish 2006 (324)

As mentioned in 5.4.2, LSM correlates better with latter stages of fibrosis rather than earlier stages of fibrosis. Wong et al's (280) study that shows LSM correlates better with pericellular fibrosis is only partly the answer. The over-arching reason for the poor correlation is that the relationship is not linear or proportional between each stage. In the figure, fibrosis quantity is observed to increase minimally in the early stages (Ishak 0 to 2), before the steepness of the slope increases between Ishak 2 and 3. The slope becomes increasingly steeper from Ishak 3 to 4 and again from Ishak 4 to 5. This explains why the LSM interval for the early stages of fibrosis are close together, but then become much wider in the latter stages. This observation is found consistently in all TE studies that aim to define cut-offs for each fibrosis stage. Our study also illustrates this observation very well. The cut-offs for F1, 2, 3 and 4 are 6.5 kPa, 7.5 kPa, 9.6 kPa and 11.9 kPa respectively. The corresponding interval between each cut-off is 1 kPa, 2.1 kPa and 2.3 kPa. As the gap between cut-offs are much smaller for earlier stages of fibrosis, the margin in which the LSM must fall within to reflect the fibrosis stage also becomes much smaller. Thus, small differences in the LSM at these intervals become more significant. Hence the LSM is not as accurate for diagnosing earlier stages compared to latter stages of fibrosis.

It follows that a much better comparison of LSM can be made with a quantitative measure of fibrosis rather than the fibrosis stage. 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).

Early studies show that LSM correlates better with CPA rather than fibrosis stage. Isgro et al. has demonstrated that LSM is better correlated with CPA in CHB patients (r^2=0.61) than to Ishak staging (r^2=0.52) (325).

Despite being an awkward comparison, prominent journals still publish studies that compare TE and other non-invasive markers with the histological fibrosis stage. CPA and similar measures of fibrosis volume are not commonly performed. At its core, fibrosis stage is still the most established and validated method of assessing fibrosis severity. Comparing non-invasive measures against the fibrosis stage is entrenched because the fibrosis stage has prognostic significance. While quantitative fibrosis measures such as CPA are a more scientifically sound method of comparison; it is largely an unknown quantity to most clinicians. Few data exists that allow us to make prognostic assumptions based on the CPA. Thus LSM being compared with fibrosis stage is still the most pragmatic way to determine its clinical significance.

This begs the question upon whether LSM is better being considered as stand-alone marker. Rather than being used as a marker for fibrosis stage, can certain LSM cut-off values be applied to imply clinically important outcomes? Longitudinal studies analyzing the relationship between LSM and clinical end points over long term follow up are required to determine the usefulness of the LSM being considered a stand-alone marker. The utility of TE being used in a longitudinal fashion in chronic hepatitis B patients is explored in the next chapter.

5.5 CONCLUSION

Fibroscan is a reliable and accurate non-invasive tool for diagnosing fibrosis stage. It has a superior or equal diagnostic accuracy compared to FIB-4, APRI, API, AAR and FI for each stage of fibrosis. It has a particularly high diagnostic accuracy for F≥3 and F4 and reduced the need for liver biopsies in the majority of chronic hepatitis B patients. Liver histology was observed to have significant inter-observer variability. The comparison of non-invasive markers against the imperfect "gold standard" of liver biopsy may underestimate the true accuracy of Fibroscan and other non-invasive markers. Correlation with clinically important are needed to determine the utility of the LSM beyond a marker of liver fibrosis.

This page has been left intentionally blank

CHAPTER 6: A LONGITUDINAL STUDY OF TRANSIENT ELASTOGRAPHY IN CHRONIC HEPATITIS B PATIENTS

6.0 CHAPTER SUMMARY

Introduction: Monitoring of liver fibrosis in chronic hepatitis B (CHB) allows for optimal timing for initiating antiviral therapy, assessing the response to treatment and for surveillance for cirrhotic complications. There is little longitudinal data on the utility of Fibroscan LSM measurements in the long term follow up of CHB patients. We sought to identify which patients experienced a LSM decline, and determine clinical parameters associated with this change.

Methods: Chronic hepatitis B patients who had a Fibroscan in the last 4 years were invited to have a follow up scan. Fibroscan was performed as previously described in chapter 1.7. Relevant clinical data were collected from medical records.

Results: One hundred and twenty four patients were recruited and included in the analysis. The mean follow up period was 31.2 months (SD 13.1). LSM decline compared to baseline was observed in the following: patients who had antiviral therapy initiated 7.4 kPa (SD 3.0) vs 5.9 kPa (SD 2.5), p=0.009; patients who, with or without antiviral therapy, experienced HBeAg seroconversion: 8.6 (SD 3.7) kPa vs 5.4 (SD 2.8) kPa and or experienced viral suppression: 6.2 (SD 2.5) kPa vs 4.8 (SD 2.2) kPa, p<0.001. A mild LSM decline occurred in patients who did not receive antiviral therapy: 5.2 (SD 2.0) kPa vs 4.6 (SD 1.4) kPa, p=0.006. In those with persistent normal ALT (PNALT), a numerical fall in the mean LSM occurred, which was not statistically significant: 5.3 kPa (SD 1.8) vs 5.1 kPa (SD 1.7), p=0.353. However, in more than 1/3rd of patients the decline was > 1 kPa. Parameters that correlated with the change in LSM in those who began antiviral therapy were HBeAg seroconversion (r=0.668, p<0.001) and change in level of ALT (r=0.489, p<0.001). On the other hand, duration of monitoring was the parameter most correlated with LSM change in those who had PNALT (r=0.381, p=0.008) or were monitored and not given antivirals (r=0.321, p=0.007).

Conclusions: CHB patients, who have active disease and subsequently treated with antivirals, have the greatest decline in LSM. This is associated with ALT normalization, HBeAg seroconversion and viral suppression. CHB patients who had quiescent disease and did not require antiviral treatment had a mild decline in LSM, while those with persistently normal ALT irrespective of treatment had a numeric but not significant decline in the mean LSM. The LSM change is correlated with the duration of monitoring in these patients. The LSM may not be an ideal tool to monitor fibrosis regression. Future studies on the use of the LSM should focus on its potential in being a prognostic tool for liver related morbidity and mortality.

6.1 Background

6.1.1 Assessing liver fibrosis in chronic hepatitis B

Chronic hepatitis B (CHB) affects 240 million people worldwide and causes over 780 000 deaths per year from cirrhosis and hepatocellular carcinoma (HCC) (93, 94). The development of liver fibrosis in CHB patients is an adverse prognostic factor (326).

Treatment with antivirals results in fibrosis regression (15, 327-329) and reduces the incidence of long term complications of chronic hepatitis B such as HCC (330, 331). However, not all patients universally benefit with treatment. In up to 44%, fibrosis can remain static or may worsen (329), and in some patients already with advanced disease, treatment may not reduce the risk of HCC (330). Thus regular monitoring of fibrosis in those on treatment conveys important information about response to therapy, prognosis, and may be useful in guiding continued treatment and surveillance for liver related complications (259, 332).

In patients who are not treated, regular fibrosis monitoring will help accurately identify patients who develop worsening fibrosis which may prompt antiviral therapy and surveillance for liver related complications (259, 332).

Liver biopsy is the reference standard for the assessment of liver fibrosis. However, it has several disadvantages that have been highlighted in detail earlier (see Section 1.5 and 5.4.10). Liver biopsy is not an ideal method for the assessment of liver fibrosis, especially if repeated monitoring is desired. Transient Elastography (TE) performed with Fibroscan ® is a rapid, safe and pain free noninvasive method for assessing liver fibrosis and has become well-accepted by patients. In chapter 5, we demonstrated that Fibroscan is accurate and reliable for diagnosing fibrosis stage. It has a high diagnostic accuracy for F≥3 (AUROC 0.874 95% CI 0.775-0.973, p<0.001) and F4 (AUROC 0.945, 95% CI 0.867-1.000, p=0.001), and is superior to FIB-4, APRI, API, AAR and FI. Fibroscan has one of the highest accuracies amongst non-invasive measures, while also being the most widely validated (303).

6.1.2 The interpretation of liver stiffness decline in chronic hepatitis B patients is contentious

There is limited data on the role of Fibroscan in long term follow up of patients with CHB. The focus has been primarily on the changes in LSM in response to antiviral therapy (333-336), and the role of LSM in prognostication of liver related mortality and morbidity (326, 337) referred to as liver related events (LREs).

Reduction in LSM on treatment could be presumed to be due to fibrosis regression. But studies examining the changes in LSM in response to antiviral treatment have so far reported conflicting results. In a paired biopsy study, the reduction in LSM was found to correlate with ALT normalization and was unreliable in reflecting fibrosis regression (336). In contrast, 2 other studies (which excluded patients with elevated ALT levels) reported a reduction in LSM with antiviral therapy, concluding that it was due

to fibrosis regression. The latter 2 studies confirmed fibrosis regression by performing liver biopsies in a limited sample of patients in their study population (334, 335).

Further conflicting data was reported in a study of 426 CHB patients over 3 years (333). LSM decline was observed in patients treated with antivirals who had an elevated baseline ALT with subsequent normalization. However, LSM decline was not seen in those with persistently normal ALT (PNALT) who were treated. On the other hand, untreated patients with raised baseline ALT with subsequent normalization did not experience an LSM decline, but an LSM decline was seen in untreated patients with PNALT. The data is inconsistent and further studies to evaluate the role of Fibroscan are needed.

We seek to clarify these issues by analyzing the change in LSM over time in our local cohort of CHB patients, specifically attempting to address the effect of ALT and antiviral therapy on the LSM.

Lastly, the few studies have also examined the prognostic value of LSM only included a total of 64 cases of HCC and 40 other LREs (variceal bleeding, ascites, hepatic encephalopathy) (326, 337). Thus we sought to describe any LREs in relation to the LSM at baseline and follow up.

6.1.3 Hypothesis and specific objectives

The main hypothesis is that LSM decline occurs in patients who have been treated with antivirals and who achieve PNALT. We also predict the LSM to decline in CHB patients who experience a decline in ALT, a decrease in HBV viral load and HBeAg seroconversion.

Objective 1: To evaluate the LSM change between follow up and baseline, and to correlate this change with clinical parameters - in patients according to the status of their antiviral treatment

Objective 2: To evaluate the LSM change between follow up and baseline, and to correlate this change with clinical parameters - in patients with persistently normal ALT (PNALT)

Objective3: To evaluate the LSM change between follow up and baseline according to

- Type of antiviral therapy (for those who were treated)
- HBeAg status and
- HBV viral load.

Objective 4: To describe liver related events in relation to the LSM at baseline and follow up.

6.2 METHODS

6.2.1 Patient selection, recruitment, Fibroscan assessment and data collection

All Fibroscans that were performed from August 2008 to January 2012 at Concord Repatriation General Hospital, Concord Sydney NSW Australia were reviewed. There were 2157 scans, of which 304 were performed for chronic hepatitis B patients (as defined by surface antigen positivity > 6 months). From January 2012 to December 2013, 166/304 patients returned for follow-up during this period of time and had a repeat Fibroscan. All patients were older than 18yrs and gave written informed consent.

The performance of the Fibroscan has been described in detail earlier in this thesis (refer to section 1.7).

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, and any other documented chronic liver disease, imaging results, INR, liver function tests [bilirubin (BR), alkaline phosphatase (ALP), gamma glutamyl-transferase (GGT), alanine aminotransferase (ALT), aspartate aminotransferase (AST) and albumin (ALB). hepatitis B surface antigen status, hepatitis B E antigen status, hepatitis B DNA viral load, details of antiviral therapy, platelet count, and any liver disease related morbidity.

For variable data (eg. hepatitis B DNA viral load), the value recorded was closest to the date of Fibroscan and not exceeding 3 months. If the parameter of interest was not evaluated within 3 months of the date of the Fibroscan, it was not recorded and data was considered incomplete for that subject. Clinical parameters corresponding to both the baseline and follow-up Fibroscan were recorded. Patients with complete data totaled 124/166 and were included in final analysis.

6.2.3 Patient groups

To determine the effect of antiviral treatment on LSM change over time, patients were categorized into four mutually exclusive groups. Group 1 was patients who were on no antiviral treatment. These patients were clinically assessed to not require antiviral treatment and were monitored without antivirals throughout the follow-up duration. Group 2 were patients who started treatment at the time of monitoring (defined as within 6 months of the baseline TE). Group 3 had prior antiviral therapy (last dose was greater than 6 months before the baseline TE). Lastly, group 4 was patients who were already on long-term antivirals (defined as being on at least 6 months of therapy prior to baseline TE) which were continued during the monitoring period. These groups are described in ***Figure 34***.

6.2.3 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 and Fisher's exact test were used for categorical variables when appropriate. Multivariate analysis was performed with stepwise multiple linear regression. Categorical variables were assigned dummy continuous variables for the purposes of statistically analysis. A p-value of <0.05 was considered statistically significant.

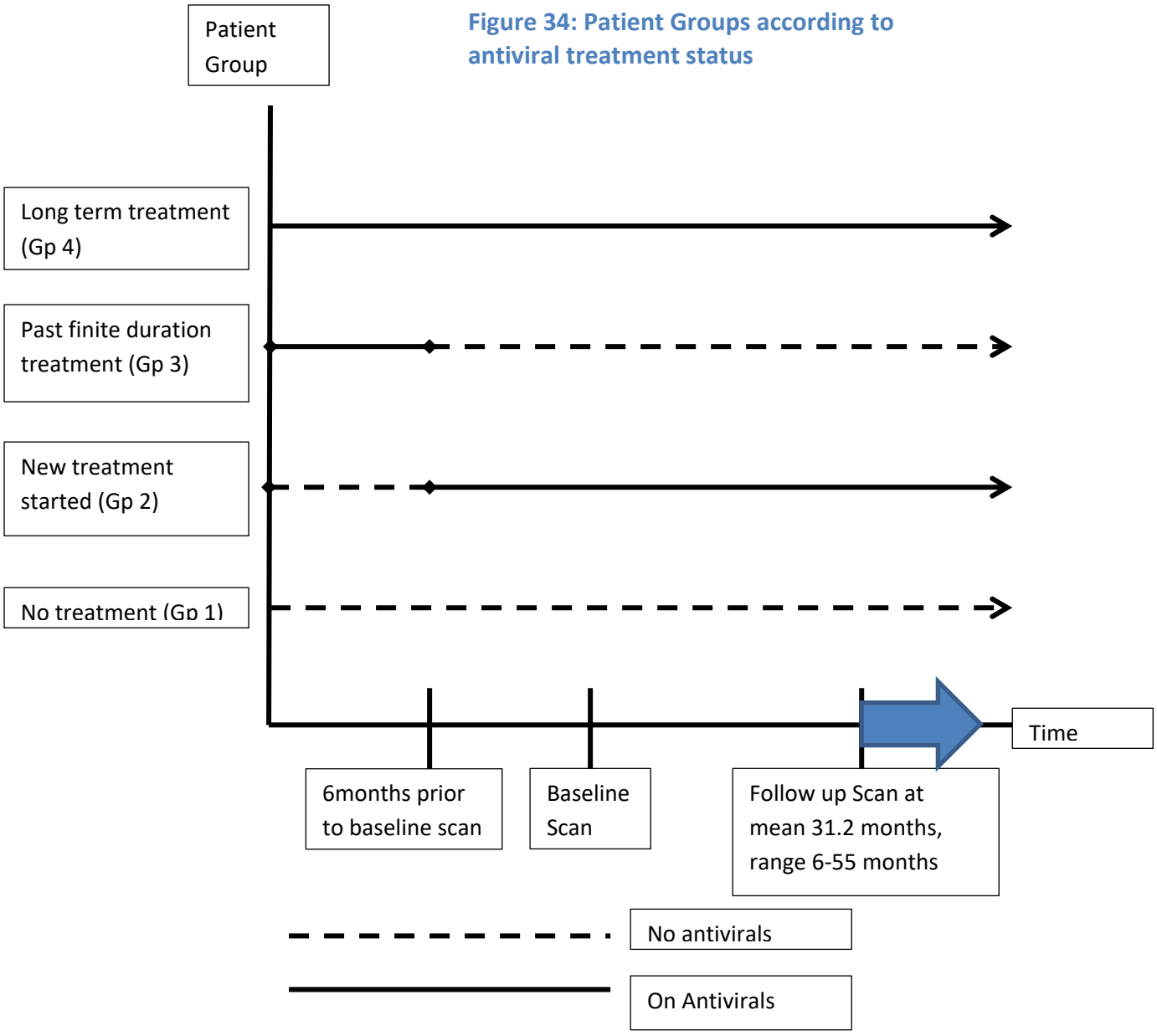


Figure 34: Patient Groups according to antiviral treatment status

6.3 RESULTS

6.3.1 Liver Stiffness and clinical characteristics at baseline and follow up for the entire study cohort

The study population had 68/124 (54.8%) patients who were male. The mean age was 47.6 (SD 10.6) yrs. At baseline, the mean level of liver biochemistry and platelet count were within reference limits. The mean baseline LSM was 5.8 kPa (SD 2.4) and the mean IQR/M ratio 19.0% (SD 11.9). Reliable LSM measurements were taken in 118/124 (91.2%) patients. The mean success rate was 91% (SD 15.6). HBeAg was positive in 36 patients (29%). The total number of patients on antiviral treatment was 42 (33.9%) and the mean $\log_{10}$ HBV viral load was 3.6 IU/ml (SD 2.0). The patient characteristics at baseline for the entire study cohort and for groups 1,2,3 and 4 are shown in ***Table 38***: Clinical Characteristics of the entire study cohort, and subdivided into groups 1,2,3 and 4..

The mean time interval between the baseline and follow-up fibroscan was 31.2 months (SD 13.1, range 6-55). There was a significant LSM decline at follow up: 5.8 (SD 2.5) kPa vs 5.1 (SD 1.9) kPa, p<0.001. The IQR/M ratio was higher at baseline compared to follow up, but was below the accepted reliability criteria threshold of ≤ 30%: 19.0 (SD 11.9) vs 15.5 (10.0), p=0.017. Furthermore, the number of reliable scans was not significantly different: 118/124 (95.2%) vs 122/124 (98.4%), p=0.281; nor was the success rate of scans 91.0% (SD 15.6%) vs 88.8% (SD19.0%), p=0.386.

Other clinical parameters that changed at follow-up compared to baseline include the following: number of HBeAg positive patients decreased from 36/124 (29%) to 27/124 (21.8%), p<0.001; number of patients on antiviral treatment had increased from 42/124 (33.9%) to 67/124 (54%), p=0.0021; and $\log_{10}$HBV viral load decreased from 3.6 (SD 2.0) IU/ml to 2.6 (SD 1.7) IU/ml, p<0.001. There were also small differences that were statistically significant, but not clinically significant in the following: bilirubin 13 (SD 6) umol/L vs 11 (SD 5) umol/L; albumin 45.4 (SD 3) g/L vs 44.6 (SD 3) g/L; and AST 35(SD 32) U/L vs 30 (SD 22) U/L. There was no difference between baseline and follow up levels of ALP, GGT, and ALT. The patient characteristics at baseline and follow up are summarised in ***Table 39***.

Table 38: Clinical Characteristics of the entire study cohort, and subdivided into groups 1,2,3 and 4.

Patient Characteristic	Entire Cohort	Group 1	Group 2	Group 3	Group 4
Male[b]	68/124 (54.8)	26/57 (46.0%)	13/25 (52.0%)	3/5	26/37 (70.3)
Age[a] (years)	47.6 (10.8)	50.3 (9.9)	41.6 (12.3)	37.8 (5.8)	49.2 (9.5)
LSM (kPa)[a]	5.8 (2.5)	5.2 (2.0)	7.4 (3.0)	5.6 (2.3)	5.8 (2.1)
IQR/M ratio[a]	19.0 (11.9)	0.20 (0.11)	0.17 (0.08)	0.26 (0.17)	0.17 (0.14)
Reliable scans[b]	118/124 (95.2)	54/57 (94.7)	25/25 (100)	5/5 (100)	36/37 (97.2)
Success rate (%)[a]	91.0 (15.6)	90.1 (13.4)	94.5 (14.6)	76.1 (22)	90.6 (17.5)
HBeAg +[b]	36/124 (29.0)	4/57 (7.0%)	11/25 (44%)	0/5 (0%)	20/37 (54%)
Antiviral treatment[b] (%)	42/124 (33.9)	0/57	25/25	0/5	37/37
BR (umol/L)[a]	13 (6)	12 (6)	10 (5)	16 (11)	14 (7)
ALB (g/L)[a]	45 (3)	45 (2)	46 (6)	46 (4)	45 (2)
ALP (IU/L)[a]	70 (19)	67 (17)	72 (23)	66 (18)	74 (20)
GGT (IU/L)[a]	26 (16)	22 (15)	32 (17)	19 (14)	27 (14)
ALT (IU/L)[a]	49 (73)	31 (19)	114 (142)	43 (37)	35 (21)
AST (IU/L)[a]	35 (32)	27 (9)	64 (62)	27 (7)	28 (9)
Log_{10}HBV viral load (IU/ml)[a]	3.6 (2.0)	3.7 (1.6)	5.2 (2.5)	3.5 (1.4)	2.5 (1.5)

a. Continuous variables: The mean and (standard deviation) are shown.
b. Categorical variables: The numbers and (proportions of total patients) are shown.

Table 39: Clinical characteristics of the entire study cohort at baseline and follow up

Patient characteristic	Baseline	Follow up	P value
Male[b]	68/124 (54.8%)	n/a	n/a
Age[a]	47.6 (10.8)	n/a	n/a
LSM (kPa)[a]	5.8 (2.5)	5.1 (1.9)	< 0.001
IQR/M ratio[a]	19.0 (11.9)	15.5 (10.0)	0.017
Reliable scans[a] (%)	118/124 (95.2%)	122/124 (98.4%)	0.281
Success rate (%)[a]	91.0 (15.6)	88.8 (19.0)	0.386
HBeAg +[b] (%)	36/124 (29.0%)	27/124 (21.8%)	<0.001
Antiviral treatment [b] (%)	42/124 (33.9%)	67/124 (54.0%)	0.0021
BR (umol/L)[a]	13 (6)	11 (5)	0.005
ALB (g/L)[a]	45 (3)	45 (3)	0.012
ALP (IU/L)[a]	70 (19)	70 (19)	0.952
GGT (IU/L)[a]	26 (16)	23 (14)	0.057
ALT (IU/L)[a]	49 (73)	37 (40)	0.068
AST (IU/L)[a]	35 (32)	30 (22)	0.028
Log_{10}HBV viral load (IU/ml)[a]	3.6 (2.0)	2.6 (1.7)	< 0.001

a. Continuous variables: The mean and (standard deviation) are shown. Paired samples T-test was performed where applicable
b. Categorical variables: The proportions and percentages of total patients are shown. Fisher's exact test and Chi-square test was performed where applicable

6.3.2 Change in liver stiffness measurement from baseline compared with the change in other clinical parameters for the entire study cohort

The change in LSM at follow up was found to have a significant association with the change in several biochemical and serological variables in univariate analysis. These include: ALB, ALP, GGT, ALT, AST, log_{10}HBV DNA viral load; HBe antigen seroconversion; and the duration of follow-up. But in the multivariate analysis, the only significant correlation were the change in ALT: -12 IU/L (SD 74), r=0.489, p<0.001; and the HBeAg seroconversion: 9 cases (27 HBeAg positive cases at follow up compared to 36 at baseline) r=-0.377, p<0.001. No other correlations were found. The analysis is summarised in ***Table 40***.

Table 40: Correlation of the change in clinical parameters with the change in liver stiffness for the entire study cohort

Patient characteristic	Change at follow up	Correlation to change in LSM (r)	P value
Male[c]	n/a	-0.105	0.105
Duration between LSM[a] (months)	31.2 (13.1)	-0.153	0.045
LSM (kPa)[a]	0.7 (1.8)		
HBeAg +[b] (# cases)	-9	-0.377	<0.001[d]
LogHBV viral load (IU/ml)[a]	1.0 (2.0)	0.350	<0.001
Antiviral treatment [b] (# cases)	+25	0.035	0.355
BR (umol/L)[a]	-1 (5)	-0.055	0.274
ALB (g/L)[a]	-1 (4)	-0.208	0.010
ALP (IU/L)[a]	0 (13)	0.179	0.023
GGT (IU/L)[a]	-2 (14)	0.318	<0.001
ALT (IU/L)[a]	-12 (74)	0.489	<0.001[d]
AST (IU/L)[a]	-5 (25)	0.430	<0.001

a. Continuous variables: The mean (standard deviation) is shown.
b. Categorical variables: A dummy continuous variable was assigned for multiple regression
c. Baseline was compared to LSM change
d. Significant in multivariate analysis

6.3.3 Patients not treated with antivirals (Group 1)

Fifty seven out of 124 patients were treatment naïve through the study period (group 1 - see 6.2.3). These patients were predominantly HBeAg negative 53/57 (92.9%), and had persistently normal ALT (PNALT) in 36/57 (61.2%) of cases. Males comprised 26/57 (46.0%) of total subjects. The mean age was 50.3 yrs (SD 9.9), and the mean duration between baseline and follow up Fibroscans was 26.4 months (SD 11.3). The mean baseline ALT, AST and log_{10}HBV viral load were: 31 IU/L (SD 19), 27 IU/L (SD 9) and 3.7 IU/ml (SD 1.6) respectively. The rest of the liver biochemistry at baseline was within normal limits.

At follow up, there was a significant decline in LSM: 5.2 kPa (SD 2.0) vs 4.6 kPa (SD 1.4), p=0.006. The LSM declined in 38/57 (67%) of subjects, and by > 1 kPa in 22/57 (39%) of subjects. There were small statistically significant, but clinically unimportant changes in the following parameters: : BR 12 umol/L (SD 6) vs 11 umol/L (SD 5), p<0.001; ALB 45 g/L (SD 2) vs 44 g/L (SD 3), p<0.001; ALP 67 IU/L (SD 17) vs 66 IU/L (SD 16), p<0.001; GGT 22 IU/L (SD 15) vs 23 IU/L (SD 14), p<0.001; and log_{10}HBV viral load 3.7 IU/ml (SD 1.6) vs 3.5 IU/ml (SD 1.7), p<0.001. The ALT and AST were not statistically different at follow up: 31 IU/L (SD 19) vs 35 IU/L (SD 27), p=0.070 and 27 IU/L (SD 9) vs 30 IU/L (SD 10), p=0.197 respectively. The difference between the mean baseline and follow-up LSM and other clinical variables for group 1 is summarised in ***Table 41***

The change in LSM at follow up was found to have a significant association with male gender, duration between scans, and the change in ALT. But in multivariate analysis, the only associations were the duration between scans (r=-0.321, p=0.007) and the change in ALT (r=0.256, p=0.027). These relationships are shown in scatterplots below (see ***Figure 35 and Figure 36***). A summary of the analysis is also shown in ***Table 42***.

Table 41: Clinical characteristics of group 1 patients at baseline and follow up

Patient Characteristic	Baseline	Follow up	P value
Male[b]	26/57 (46.0%)	n/a	n/a
Age[a] (years)	50.3 (9.9)	n/a	n/a
LSM (kPa)[a]	5.2 (2.0)	4.6 (1.4)	<0.001
IQR/M ratio[a]	0.20 (0.11)	0.15 (0.08)	0.231
Reliable scans[b]	54/57 (94.7)	57/57 (100)	0.243
Success rate (%)[a]	90.1 (13.4)	92.1 (17.3)	0.861
HBeAg +[b]	4/57 (7.0%)	4/57 (7.0%)	1.000
BR (umol/L)[a]	12 (6)	11 (5)	<0.001
ALB (g/L)[a]	45 (2)	44 (3)	<0.001
ALP (IU/L)[a]	67 (17)	66 (16)	<0.001
GGT (IU/L)[a]	22 (15)	23 (14)	<0.001
ALT (IU/L)[a]	31 (19)	35 (27)	0.070
AST (IU/L)[a]	27 (9)	30 (10)	0.197
Log_{10}HBV viral load (IU/ml)[a]	3.7 (1.6)	3.5 (1.7)	<0.001

a. Continuous variables: The mean and (standard deviation) are shown. Paired samples T-test was performed where applicable
b. Categorical variables: The proportions of total patients are shown. Fisher's exact test and Chi-square test was performed where applicable

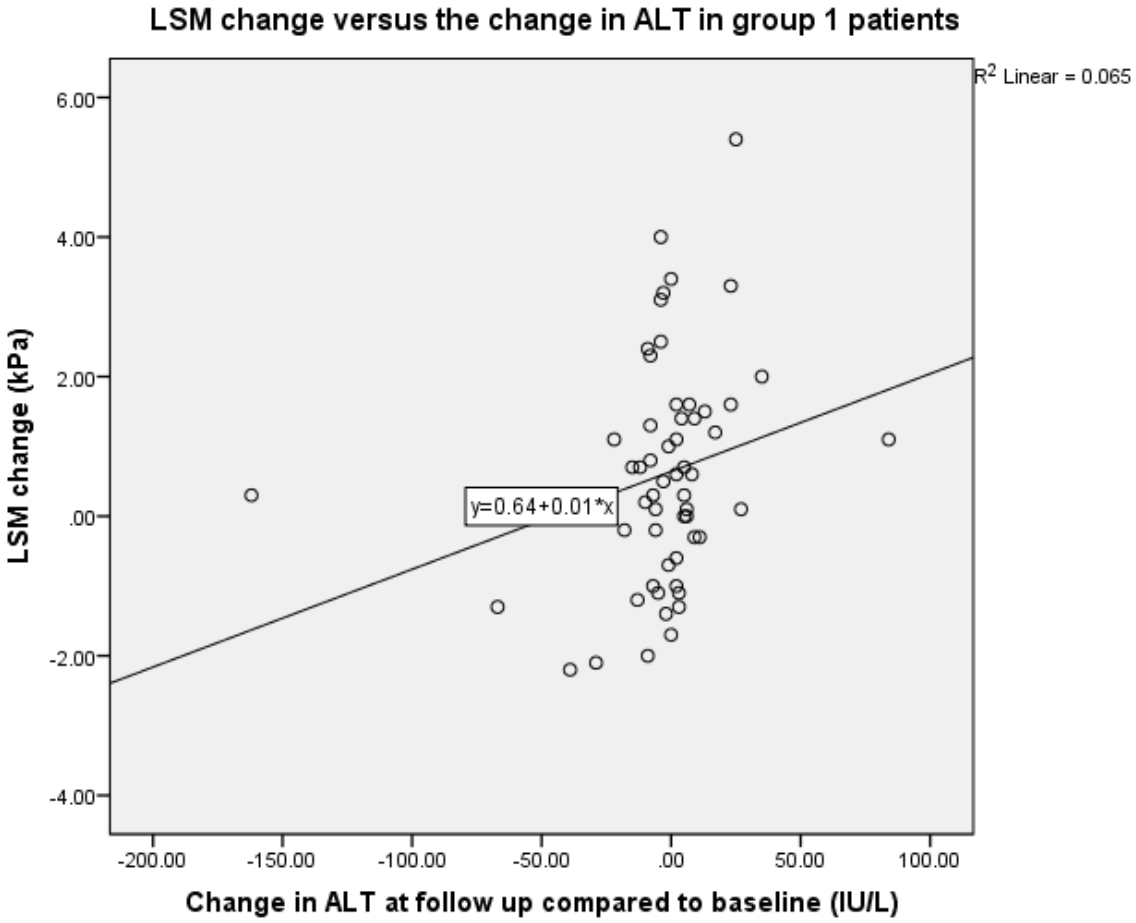


Figure 35: Scatterplot of LSM change and ALT change in group 1 patients

Y-axis is the LSM change (kPa), while x-axis is the change in ALT between TE assessments. A line of best fit is drawn.

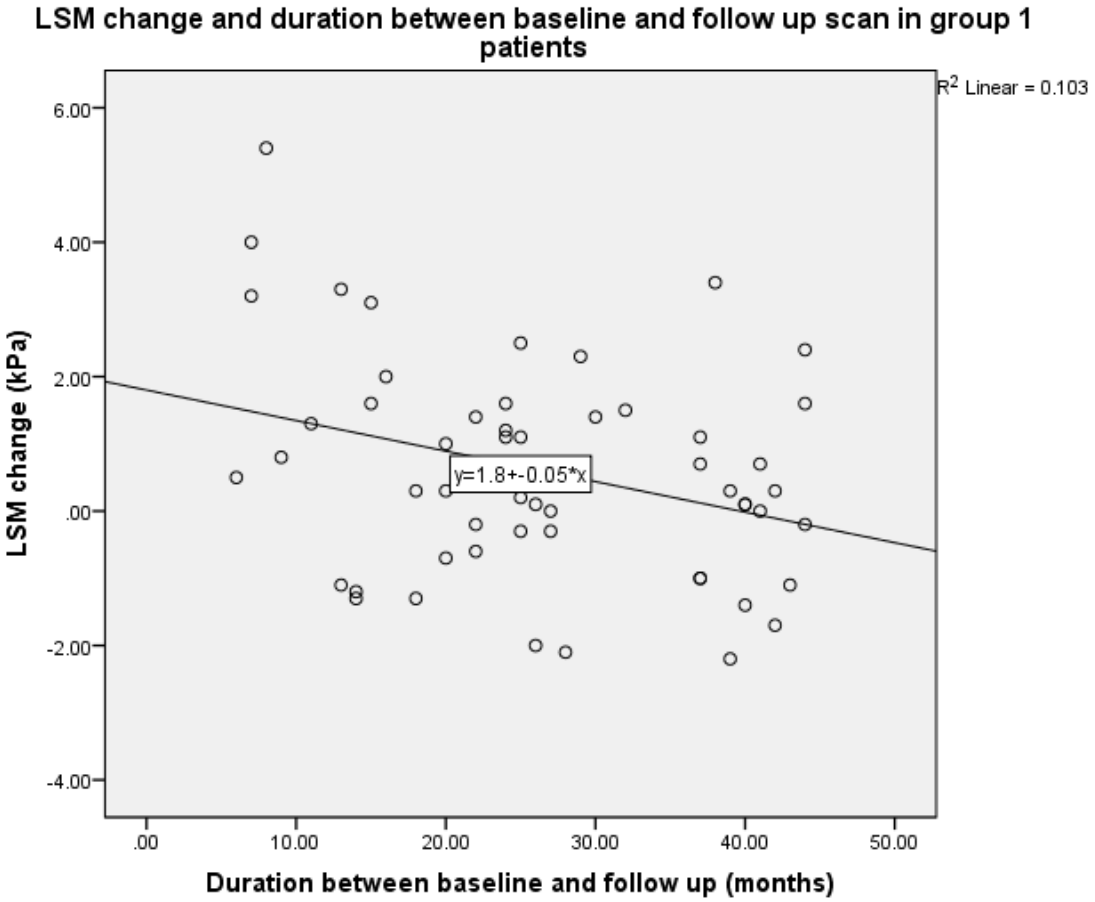


Figure 36: Scatterplot of LSM change and duration between scans in group 1 patients

Y-axis is the LSM change (kPa), while x-axis is the duration (months) between TE assessments. A line of best fit is drawn.

Table 42: Correlation of the change clinical parameters with the change in baseline liver stiffness in group 1 patients

Patient Characteristic	Change at follow up	Correlation to change in LSM (r)	P value
Male[b]	n/a	-0.293	0.013
Duration between LSM[a] (months)	26.4 (11.3)	-0.321	0.007[c]
HBeAg +[b] (cases)	0	-0.085	0.265
Log_{10}HBV viral load (IU/ml)[a]	0.2 (1.4)	0.067	0.309
BR (umol/L)[a]	-1 (4)	0.028	0.419
ALB (g/L)[a]	-1 (3)	-0.046	0.367
ALP (IU/L)[a]	-1 (12)	0.133	0.163
GGT (IU/L)[a]	1 (13)	0.032	0.407
ALT (IU/L)[a]	3 (29)	0.256	0.027[c]
AST (IU/L)[a]	3 (12)	0.212	0.057

a. Continuous variables: The mean (standard deviation) is shown.
b. Categorical variables: A dummy continuous variable was assigned for the multiple regression
c. Significant in multivariate analysis

6.3.4 Patients started on antiviral therapy (Group 2)

There were 25 out of 124 patients who were newly treated with antivirals (group 2 – see 6.2.3). There were 5/25 subjects treated with 48 weeks of pegylated interferon, 16/25 subjects were treated with entecavir and 4/25 subjects who were treated with tenofovir. Most of these patients had elevated ALT (23/25 subjects - 92.0%) and almost half were HBeAg positive (11/25 - 44.0%). There were 13/25 subjects who were male (41.6%), with a mean age of 41.6 yrs (SD 12.3). The mean duration between baseline and follow up was 31.4 months (SD 14.0). The mean baseline ALT, AST and log_{10}HBV viral load were: 114 IU/L (SD 142); 64 IU/L (SD 62) and 5.2 IU/ml (SD 2.5) respectively. The other parameters were within normal limits.

There was a significant decline at follow up in the LSM: 7.4 kPa (SD 3.0) vs 5.9 kPa (SD 2.5), p=0.009. There was also a significant reduction in: GGT 32 IU/L (SD 17) vs 21 IU/L (SD 9), p=0.006; ALT 114 IU/L (SD 142) vs 50 IU/L (SD 74), p=0.005; AST 64 IU/L (SD 62) vs 37 IU/L (SD 45), p=0.005; and log_{10}HBV viral load 5.2 IU/ml (SD 2.5) vs 2.2 IU/ml (SD 1.8), p<0.001. No other differences were found at follow-up in the other parameters. This is summarised in ***Table 43***

The change in LSM at follow up was found to be associated with HBeAg seroconversion, and with the change at follow up in ALB, GGT, ALT and AST on univariate analysis. But in multivariate analysis, the only association was HBeAg seroconversion (r=-0.668, p<0.001). An analysis of the clinical variables that were associated with the change in LSM from baseline compared follow-up for group 2 is shown in ***Table 44***.

Table 43: **Clinical** characteristics of patients of group 2 patients at baseline and follow up

Patient Characteristic	Baseline	Follow up	P value
Male[b]	13/25 (52.0%)	n/a	n/a
Age[a] (years)	41.6 (12.3)	n/a	n/a
LSM (kPa)[a]	7.4 (3.0)	5.9 (2.5)	0.009
IQR/M ratio[a]	0.17 (0.08)	0.14 (0.09)	0.751
Reliable scans[b]	25/25 (100)	24/25 (96)	1.000
Success rate (%)[a]	94.5 (14.6)	85.3 (20.0)	0.198
HBeAg +[b]	11/25 (44%)	7/25 (28%)	0.377
BR (umol/L)[a]	10 (5)	10 (4)	0.501
ALB (g/L)[a]	46 (6)	44 (3)	0.306
ALP (IU/L)[a]	72 (23)	68 (20)	0.301
GGT (IU/L)[a]	32 (17)	21 (9)	0.006
ALT (IU/L)[a]	114 (142)	50 (74)	0.005
AST (IU/L)[a]	64 (62)	37 (45)	0.005
LogHBV viral load (IU/ml)[a]	5.2 (2.5)	2.2 (1.8)	<0.001

a. Continuous variables: The mean and (standard deviation) are shown. Paired samples T-test was performed where applicable
b. Categorical variables: The proportions of total patients are shown. Fisher's exact test and Chi-square test was performed where applicable

Table 44: Correlation of the change in clinical parameters with the change in liver stiffness for group 2 patients

Patient Characteristic	Change at follow up	Correlation to change in LSM (r)	P value
Male[b]	n/a	0.130	0.268
Duration between LSM[a] (months)	31.4 (14.0)	0.027	0.450
HBeAg +[b] (# cases)	-4	0.668	<0.001[c]
LogHBV viral load (IU/ml)[a]	-3.0 (2.7)	0.565	0.002
BR (umol/L)[a]	-1 (4)	0.145	0.245
ALB (g/L)[a]	-1 (6)	0.457	0.011
ALP (IU/L)[a]	-4 (15)	0.272	0.094
GGT (IU/L)[a]	-11 (17)	0.532	0.003
ALT (IU/L)[a]	-63 (147)	0.605	0.001
AST (IU/L)[a]	-28 (45)	0.498	0.006

a. Continuous variables: The mean (standard deviation) is shown.
b. Categorical variables: A dummy continuous variable was assigned for the multiple regression
c. Significant in multivariate analysis

6.3.5 Patients with previous antiviral therapy for a finite duration (group 3)

There was only a small group of (5/124) of subjects who previously received antiviral therapy for a finite duration (group 3 – see 6.2.3). There were 3 patients who were treated with 48 weeks of pegylated interferon, 1 patient treated with both lamivudine and pegylated interferon and 1 patient treated with lamivudine only. Treatment had ceased for at least 6 months prior to the baseline Fibroscan. There were 3 males (60%), and the mean age was 37.8 yrs (SD 5.8). The mean duration between baseline and follow up Fibroscans was 29.4 months (SD 7.2). The mean baseline ALT, AST and $\log_{10}$HBV viral load was 43 IU/L (SD 37), 27 IU/L (SD 7) and 3.5 IU/ml (SD 1.4) respectively. The rest of the liver biochemistry was within normal limits.

Group 3 patients showed a decline in LSM at follow up that approached significance: 5.6 kPa (SD 2.3) vs 4.4 kPa (SD 1.5), p=0.055. None of the other parameters were showed a significant decline. An analysis of the clinical variables that were correlated with the change in LSM was not performed due to small numbers. The analysis for group 3 is summarised in ***Table 45.***

Table 45: Clinical characteristics of patients at baseline and follow up of Group 3 patients

Patient Characteristic	Baseline	Follow up	P value
Male[b]	3/5	n/a	n/a
Age[a] (years)	37.8 (5.8)	n/a	n/a
LSM (kPa)[a]	5.6 (2.3)	4.4 (1.5)	0.055
IQR/M ratio[a]	0.26 (0.17)	0.14 (0.06)	0.205
Reliable scans[b]	5/5 (100)	5/5 (100)	1.000
Success rate (%)[a]	76.1 (22)	97 (6.4)	0.052
HBeAg +[b] (# cases)	0/5 (0%)	0/5 (0%)	1.000
BR (umol/L)[a]	16 (11)	10 (4)	0.160
ALB (g/L)[a]	46 (4)	46 (4)	0.573
ALP (IU/L)[a]	66 (18)	71 (14)	0.191
GGT (IU/L)[a]	19 (14)	20 (11)	0.849
ALT (IU/L)[a]	43 (37)	28 (16)	0.341
AST (IU/L)[a]	27 (7)	22 (6)	0.310
Log_{10}HBV viral load (IU/ml)[a]	3.5 (1.4)	3.3 (1.3)	0.570

a. Continuous variables: The mean and (standard deviation) are shown. Paired samples T-test was performed where applicable
b. Categorical variables: The proportions of total patients are shown. Fisher's exact test and Chi-square test was performed where applicable

6.3.6 Patients on long term antiviral therapy started prior to baseline (group 4):

There were 37/124 subjects on long term with antivirals, started before the baseline scan (group 4-see 6.2.3). Twelve patients were maintained with entecavir monotherapy and 11 patients with tenofovir monotherapy. Five patients were treated with combination of entecavir and tenofovir. Lamivudine monotherapy was given in 4 patients and adefovir monotherapy given in 2 patients. Three patients were treated with combination of lamivudine and adefovir.

There were 26 males (70.3%), a mean age of 49.2 yrs (SD 9.5), and a mean duration between baseline and follow up of 38.6 months (SD 13.3). Elevated ALT was present in 18/37 (48.6%) and the number of HBeAg positive was 20/37 (54.1%). The mean baseline ALT, AST and $\log_{10}$HBV viral load was 35 IU/L (SD 21); 28 IU/L (SD 9) and 2.5 IU/ml (SD 1.5) respectively. The rest of the liver biochemistry at baseline was within normal limits.

At follow up, the mean LSM did not significantly decline: 5.8 kPa (SD 2.1) vs 5.5 kPa (SD 2.1), p=0.156. However, the LSM decreased in 14/37 patients (38%), and in 9/37 (24%), the mean liver stiffness decreased by > 1 kPa. At follow up, only the $\log_{10}$HBV viral load showed a significant decline: 2.5 IU/ml (SD 1.5) vs 1.4 IU/ml (SD 0.5), p<0.001. Other parameters did not change significantly from baseline. A summary is shown in ***Table 46.***

In univariate analysis, the change in LSM was correlated with HBeAg seroconversion, and the change in GGT, ALT and AST. But in multivariate analysis, the only association with the change in LSM was the change in ALT levels (r=0.455, p<0.002). This relationship is shown in the scatterplot below (see ***Figure 37***). The analysis is shown in ***Table 47***.

Table 46: Clinical characteristics of patients at baseline and follow up of Group 4 patients

Patient Characteristic	Baseline	Follow up	P value
Male[b]	26/37 (70.3)	n/a	n/a
Age[a] (years)	49.2 (9.5)	n/a	n/a
LSM (kPa)[a]	5.8 (2.1)	5.5 (2.1)	0.156
IQR/M ratio[a]	0.17 (0.14)	0.18 (0.13)	0.754
Reliable scans[a]	36/37 (97.2)	36/37 (97.2)	1.000
Success rate (%)[a]	90.6 (17.5)	84.9 (20.9)	0.200
HBeAg +[b]	20/37 (54%)	15/37 (40%)	0.352
BR (umol/L)[a]	14 (7)	13 (6)	0.298
ALB (g/L)[a]	45 (2)	45 (2)	0.556
ALP (IU/L)[a]	74 (20)	75 (22)	0.642
GGT (IU/L)[a]	27 (14)	25 (17)	0.339
ALT (IU/L)[a]	35 (21)	34 (20)	0.812
AST (IU/L)[a]	28 (9)	26 (8)	0.377
LogHBV viral load (IU/ml)[a]	2.5 (1.5)	1.4 (0.5)	<0.001

a. Continuous variables: The mean and (standard deviation) are shown. Paired samples T-test was performed where applicable
b. Categorical variables: The proportions of total patients are shown. Fisher's exact test and Chi-square test was performed where applicable

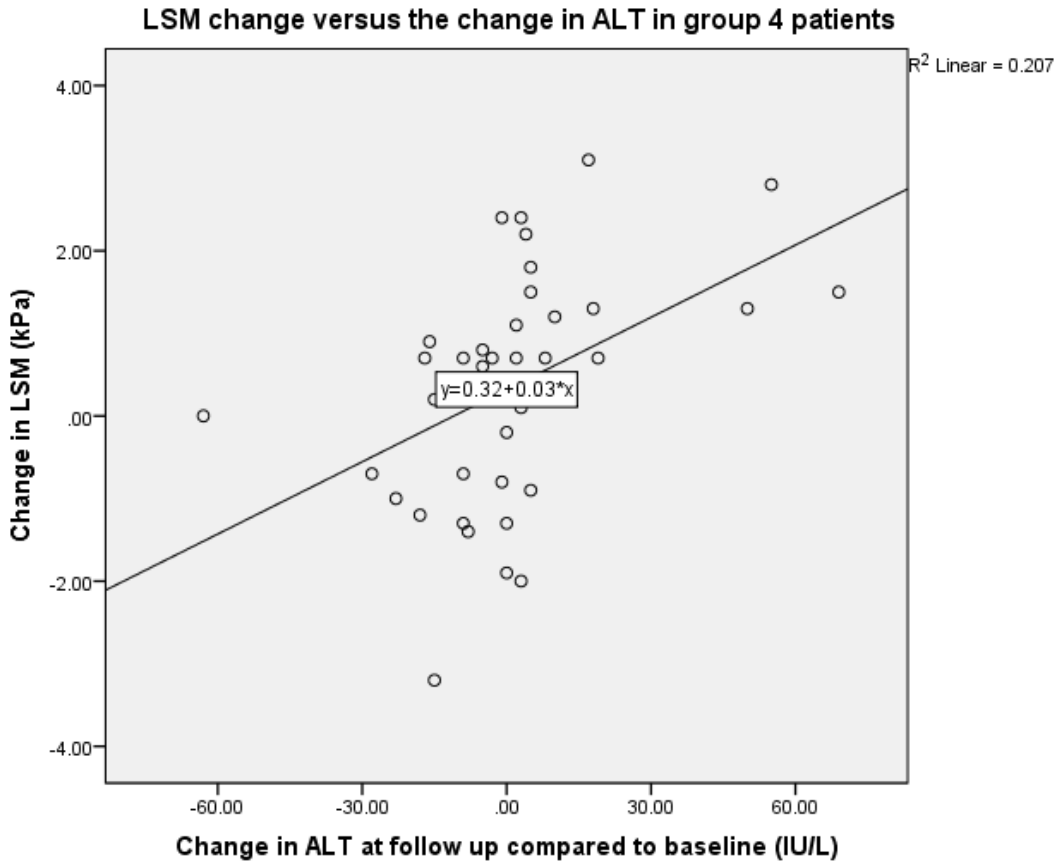


Figure 37: Scatterplot of the change in LSM and change in ALT in group 4 patients

Y-axis is the LSM change (kPa), while x-axis is the change in ALT between TE assessments. A line of best fit is drawn.

Table 47: Correlation of the change in clinical parameters with the change in liver stiffness for group 4 patients

Patient Characteristic	Change at follow up	Correlation to Change in LSM (r)	P value
Male[b]	n/a	0.144	0.197
Duration between LSM[a] (months)	38.6 (13.3)	-0.104	0.270
HBeAg +[b] (# cases)	-5	-0.244	0.073
LogHBV viral load (IU/ml)[a]	1.03 (1.43)	0.181	0.142
BR (umol/L)[a]	-1 (6)	-0.070	0.340
ALB (g/L)[a]	0 (3)	0.024	0.443
ALP (IU/L)[a]	-1 (14)	0.077	0.325
GGT (IU/L)[a]	-2 (10)	0.304	0.034
ALT (IU/L)[a]	-1 (23)	0.455	0.002[c]
AST (IU/L)[a]	-1 (10)	0.453	0.002

a. Continuous variables: The mean (standard deviation) is shown.
b. Categorical variables: A dummy continuous variable was assigned for the multiple regression
c. Significant in multivariate analysis

6.3.7 Patients in study cohort with persistently normal ALT

To clarify whether the LSM decline is influenced by other factors other than the ALT, patients who have persistently normal ALT (PNALT) were identified from the entire study cohort, irrespective of whether they were treated or not treated with antivirals. The level of ALT taken as the upper limits of normal was 30 IU/L for males and 19 IU/L for females. A total of 43 subjects were identified as having PNALT. Of these, 21/43 (49%) patients did not require therapy and remained treatment naïve (group 1 – see 6.2.3), 2/43 were started on antiviral therapy (group 2), 2/43 had past antiviral treatment for a finite duration (group 3) and 18/43 were treated with long term antiviral therapy started before baseline (group 4).

There was no difference in the LSM at baseline compared to the LSM at follow up: 5.3 kPa (SD 1.8) vs 5.1 kPa (SD 1.7), p=0.353. There were small clinically insignificant differences in the following parameters: ALT 21 U/L (SD 5) vs 27 U/L (SD 12), p = 0.002; AST 21 U/L (SD 4) vs 25 U/L (SD 4), p < 0.001; and $\log_{10}$HBV DNA viral load 2.7 IU/ml (SD 1.5) vs 2.4 IU/ml (SD 1.4), p = 0.028. No other factors showed any significant change. The findings are presented in ***Table 48***.

Although, the overall mean liver stiffness between baseline and follow up was not significantly different (mean difference LSM = 0.3 kPa, SD 1.6, p=0.353), the mean liver stiffness decreased in 19/43 cases, and by > 1 kPa in 15/43 patients.

The only variable found to be associated with the change in the LSM was the duration of monitoring (r=-0.381, p = 0.008). The scatterplot diagram with the line of best fit to demonstrate this relationship is shown in

Figure 38.

No significant association was found with the change in any of the other clinical parameters. A summary of the analysis is shown in ***Table 49.***

Table 48: Clinical characteristics of normal ALT patients at baseline and follow up and analysis of the difference

Patient characteristic	Baseline	Follow up	P value
Male[b]	33/43	n/a	n/a
Age (yrs)	49.4 (10.6)	n/a	n/a
LSM (kPa)[a]	5.3 (1.8)	5.1 (1.7)	0.353
IQR	1.1 (0.9)	0.8 (0.9)	0.068
Reliable scans[a]			
Success rate (%)[a]	88.6 (15.2)	93.3 (10.4)	0.120
HBeAg +[b]	10/43 (23%)	9/43 (21%)	1.000
Antiviral treatment [b]	21/43	23/43	0.829
BR (umol/L)[a]	13 (6)	12 (6)	0.202
ALB (g/L)[a]	45 (3)	45 (2)	0.232
ALP (IU/L)[a]	70 (20)	70 (17)	0.682
GGT (IU/L)[a]	20 (10)	21 (14)	0.269
ALT (IU/L)[a]	21 (5)	27 (12)	0.002
AST (IU/L)[a]	21 (4)	25 (4)	<0.001
LogHBV viral load (IU/ml)[a]	2.7 (1.5)	2.4 (1.4)	0.028

a. Continuous variables: The mean and (standard deviation) are shown. Independent samples T-test was performed where applicable
b. Categorical variables: The proportions of total patients are shown. Chi-square test was performed where applicable

Table 49: Correlation of the change in clinical parameters with the change in liver stiffness in persistently normal ALT patients

Patient characteristic	Change at follow up	Correlation to change in LSM (r)	P value
Male[c]	n/a	-0.121	0.228
Duration between LSM[a] (months)	30.7 (14.1)	-0.381	0.008
HBeAg +[b] (#cases)	-1	-0.191	0.119
Log_{10}HBV viral load (IU/ml)[a]	0.4 (1.0)	-0.184	0.128
Antiviral treatment [b] (# cases)	+2	-0.161	0.143
BR (umol/L)[a]	-1 (5)	0.084	0.303
ALB (g/L)[a]	-1 (3)	0.045	0.392
ALP (U/L)[a]	-1(12)	-0.067	0.341
GGT (U/L)[a]	1 (7)	0.248	0.061
ALT (U/L)[a]	6 (12)	0.093	0.284
AST (U/L)[a]	3 (5)	0.139	0.197

a. Continuous variables: The mean (standard deviation) is shown.
b. Categorical variables: A dummy continuous variable was assigned for multiple regression
c. Baseline was compared to LSM change

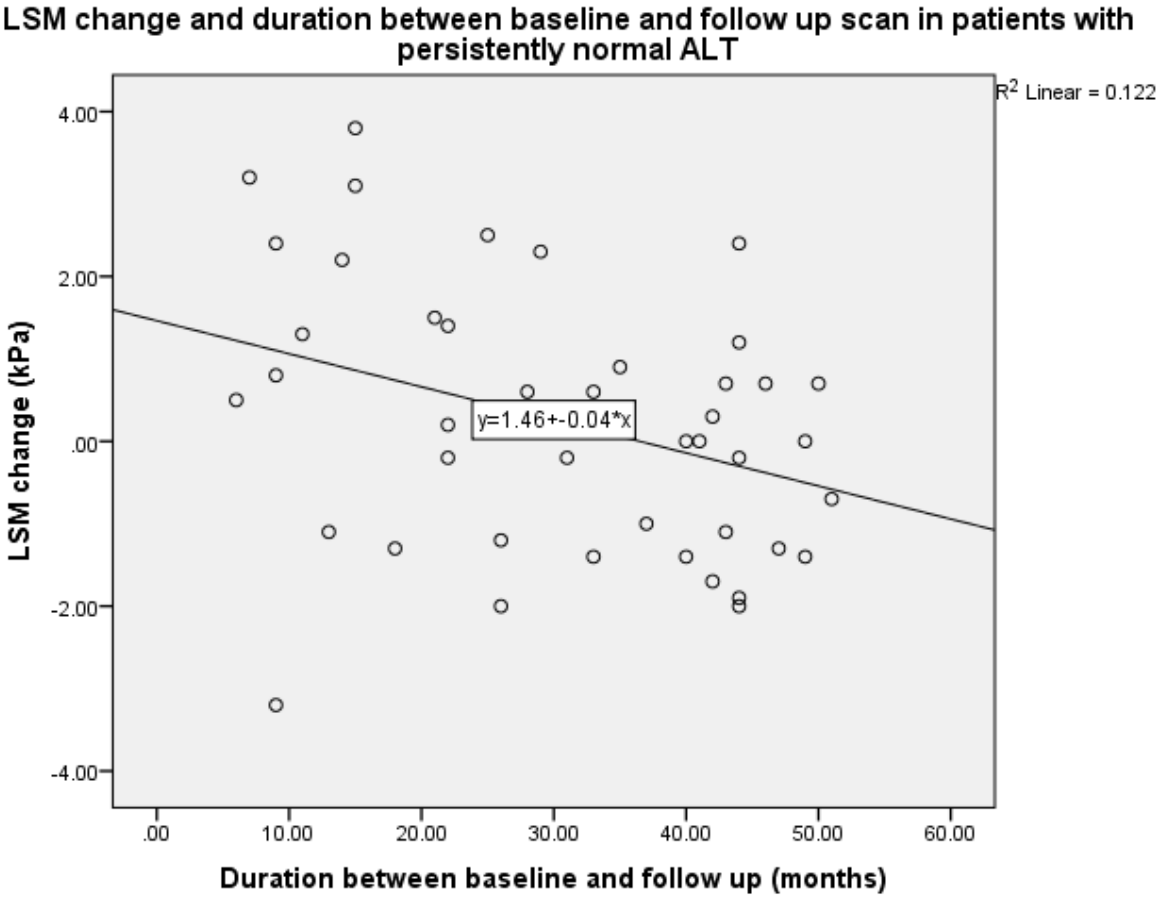


Figure 38: Scatterplot of LSM change and duration for PNALT patients

Y-axis is the LSM change (kPa), while x-axis is the duration (months) between TE assessments. A line of best fit is drawn.

6.3.8 Change in liver stiffness at follow up according to specific antiviral therapy

The mean LSM at baseline and follow-up according to the type of antiviral therapy is shown in ***Table 50.*** Only entecavir use was associated with a statistically significant decline in LSM from 6.6 (2.7) kPa to 5.7 (1.6) kPa (p=0.030). Pegylated interferon and tenofovir treated patients had a numerical decline in LSM that approached statistical significance.

Table 50: Antiviral therapy and LSM change between baseline and follow up

Antiviral	N*	Baseline LSM (kPa)	Follow up LSM (kPa)	P value
pegylated interferon	14	6.9 (2.7)	5.7 (2.2)	0.081
lamivudine	28	5.7 (2.2)	5.5 (2.1)	0.369
adefovir	17	5.7 (2.4)	5.4 (2.4)	0.538
entecavir	33	6.6 (2.7)	5.7 (2.3)	0.030
tenofovir	22	5.6 (1.7)	5.1 (1.6)	0.084
Paired samples t-test was performed. The standard deviation is shown in the parentheses. * Some patients had received more than 1 form of antiviral therapy				

6.3.9 Change in liver stiffness at follow up according to hepatitis B e antigen for entire study cohort

The mean LSM at baseline and follow up according to HBeAg status is shown in ***Table 51***. HBeAg positive patients who underwent seroconversion had a significant decline in LSM at follow up: 8.6 (SD 3.7) kPa vs 5.4 (SD 2.8) kPa, p=0.007. Patients who were persistently HBeAg negative also had a significant decrease: 5.6 (SD 2.3) kPa to 5.1 (SD 2.0) kPa, p=0.004. Patients who were persistently HBeAg positive had a decline in LSM that approached statistical significance: 5.6 (SD 1.8) kPa vs 5.0 (SD 2.0) kPa, p=0.054).

Table 51: Hepatitis B e antigen status and Liver Stiffness at baseline and follow up

HBe Antigen Status	Cases	Baseline LSM (kPa)	Follow up (kPa)	P value
HBe Antigen Positive at baseline, seroconversion to negative by follow up	9	8.6 (3.7)	5.4 (2.8)	0.007
HBe Antigen Positive at baseline and follow up	27	5.6 (1.8)	5.0 (1.7)	0.054
HBe Antigen Negative at baseline and follow up	88	5.6 (2.3)	5.1 (2.0)	0.004
Paired samples t-test was performed. The standard deviation is shown in the parentheses.				

6.3.10 Change in liver stiffness at follow up according to hepatitis B viral load for the entire study cohort

The change in LSM was analysed according to the whether HBV DNA viral load was: undetectable at both baseline and follow up; detectable at baseline, but undetectable at follow up; detectable at baseline and follow up but had either decreased, unchanged or increased; or was undetectable at baseline, but detectable at follow up .

The LSM significantly declined at follow up in patients where the viral load was detectable at baseline and subsequently became undetectable: 6.2 (SD 2.5) kPa vs 4.8 (SD 2.2) kPa, p<0.001; or if the viral load decreased by at least $1\log_{10}$ IU/ml: 6.1 (SD 3.1) kPa vs 5.0 (SD 2.1) kPa, p=0.020. No significant change in the LSM was observed if the viral load was already undetectable at baseline and remained so at follow-up, or if detectable at baseline, had unchanged or increased. The findings are shown in ***Table 52***.

Table 52: Hepatitis B viral load and Liver Stiffness at baseline and follow up

Baseline viral load detectable? [a]	Follow up viral load detectable? [a]	Number of patients	Baseline LSM (kPa)	Follow up LSM (kPa)	P value
N	N	29	6.1 (2.4)	5.8 (1.9)	0.326
Y	N	38	6.2 (2.5)	4.8 (2.2)	<0.001
Y	Y: viral load decreased [b]	13	6.1 (3.1)	5.0 (2.1)	0.020
Y	Y: viral load unchanged [c]	37	5.2 (2.0)	4.8 (1.7)	0.206
Y	Y: viral load increased [b]	5	5.9 (2.8)	5.5 (0.9)	0.714
N	Y	2	5.8 (0.0)	6.1 (2.4)	0.889

a. Detectable viral load defined as > 100 IU/ml
b. Viral load decrease and increase defined as < or > $1\log_{10}$ from baseline respectively
c. Viral load unchanged defined as <$1\log_{10}$ change from baseline

6.3.11 Cases of interest: Hepatitis B surface antigen seroconversion (2 cases) and HCC (1 case)

Only two patients during the monitoring period underwent seroconversion from surface antigen to surface antibody.

Patient 1 was a 61-year-old male who at baseline was HBeAg negative, HBV DNA undetectable, ALT 19 IU/L and LSM 9.4 kPa. He was treated with entecavir, which started four years prior. At follow up 15 months later, his LSM was 5.6 kPa, ALT 23 IU/L and HBV DNA undetectable.

Patient 2 was a 63-year-old male who at baseline was HBeAg negative, HBV DNA undetectable, ALT 17 IU/L and LSM 5.7 kPa. He was treated with pegylated interferon 3 years prior to the start of the study period. At 51 month follow up, his LSM was 2.6 kPa, ALT 26 IU/L and HBV DNA undetectable.

There was only 1 patient who developed HCC during the study period. No other patients developed any other form of liver related morbidity (eg. varices, ascites, and hepatic encephalopathy). The patient was 52-year-old female with a baseline LSM 8.8 kPa. A liver biopsy was performed six years prior to her baseline LSM which showed Metavir fibrosis stage 3. Her HBeAg was negative, ALT 32 IU/l and an undetectable viral load. Entecavir treatment was started at the time of her liver biopsy 6 years prior to the study period. During the monitoring period, she was diagnosed with locally advanced hepatocellular carcinoma and received transarterial chemo embolization. At 49 months follow up, her LSM was 6.4 kPa, ALT 33 IU/L and an undetectable HBV DNA viral load.

6.3.12 Summary of Results

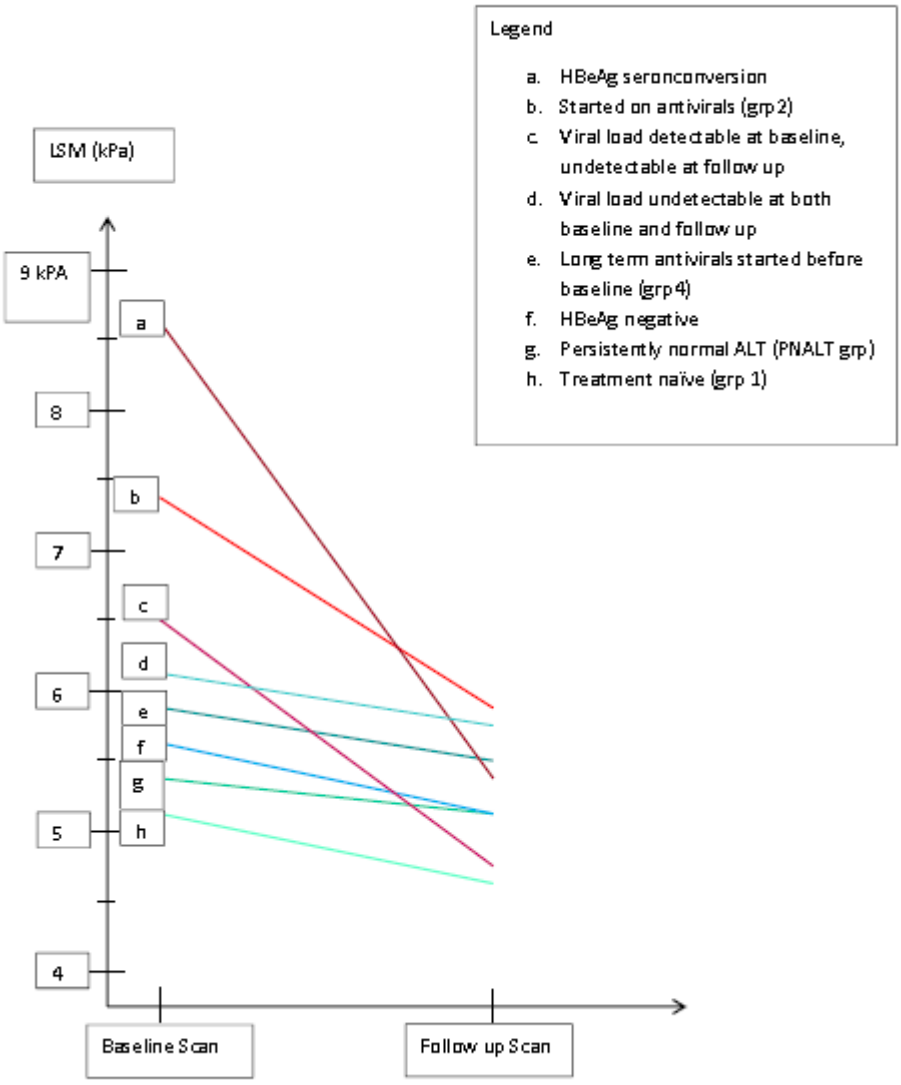
A graphical representation of the LSM decline in patients according to their clinical characteristic is shown in ***Figure 39***

- The mean LSM decline was the greatest in the following (a, b and c in figure 36)
    - Group 2 patients (newly treated with antivirals): 7.4 (SD 3.0) kPa vs 5.9 (SD 2.5) kPa, p=0.009
    - Patients who experienced HBeAg seroconversion: 8.6 (SD 3.7) kPa vs 5.4 (SD 2.8) kPa, p=0.007.
    - Patients who had a detectable viral load at baseline which became undetectable at follow up: 6.2 (SD 2.5) kPa vs 4.8 (SD 2.2) kPa, p<0.001
- The mean LSM declined minimally or not significantly at follow up for (d, e, f, g and h in figure 36)
    - Group 1 patients (treatment naïve): 5.2 (SD 2.0) kPa vs 4.6 kPa (SD 1.4) kPa, p=0.006. There were 67% of subjects who had an LSM decline, and 39% who had a decline > 1 kPa.
    - Patients with PNALT did not have a significant change in LSM: 5.3 kPa (SD 1.8) vs 5.1 kPa (SD 1.7), p=0.353. However 44% experienced a decline in LSM and 35% had a decline of > 1 kPa.
    - Group 4 patients (long term treatment started before study), however 38% experienced a decline, and 24% had a decline of > 1 kPa.
    - Patients whose viral load did not change, even for those who viral load remained undetectable at baseline and follow up, and for those whose viral load increased (are you saying that if viral load went up, LSM did not change??)
    - Patients with persistently HBeAg negative: 5.6 (SD 2.3) kPa to 5.1 (SD 2.0) kPa, p=0.004.

The parameters correlated strongest with the change in LSM at follow up were

- HbeAg seroconversion in the overall cohort (r=0.377, p<0.001) and in the sub-cohort of group 2 patients (r=0.668, p<0.001)
- Change in ALT in the overall cohort (r=0.489, p<0.001) and in the sub-cohorts of group 1 (r=0.256, p=0.027) and group 4 (r=0.455, P=0.002).
    - Duration between baseline and follow up in those who were PNALT (r=0.381, p=0.008), and in. group 1 (r=0.321, p=0.007)

Figure 39: LSM decline between baseline and follow up in patients according to clinical features



6.4 DISCUSSION

6.4.1 Decline in LSM occurs mostly in patients with active disease who are newly treated with antivirals (group 2)

In our study cohort, group 2 patients had the greatest reduction in LSM (1.5 kPa). This cohort of patients had a high baseline ALT (mean 114 IU/L) and viral load (5.2 log IU/L), and had the greatest decline in the mean ALT (63 IU/L decline), viral load (3 log IU/L decline) and HBeAg seroconversion rate (16% patients). Patients with active disease and who were treated with antiviral therapy experienced the largest decrease in LSM. This is associated with ALT decline, viral load suppression and HBeAg seroconversion.

Our results differ from findings in 2 studies (334, 335). In a study of 41 patients initiated on antivirals, Kim et al. reported significant decline in LSM values and DNA levels, while ALT decline occurred but did not reach statistical significance after year 12 months (p=0.063) or 24 months (p=0.086) (334). This study was also limited by small numbers. Enomoto et al. also concluded there was no correlation between the LSM with histological inflammatory activity or rate of ALT decline in 20 CHB patients receiving entecavir (335). However, this was a small study, and paired biopsies were not performed and so strong conclusions cannot be drawn.

Although our study did not include liver biopsy, it has the advantage of directly comparing and statistically correlating the change in LSM with the change in clinical parameters at follow up, which allowed for a rigorous analysis for whether the change in ALT was associated. Two other studies have shown a relation between LSM change and ALT levels. In a study of 71 newly treated CHB patients with liver biopsy performed at baseline and at 48 week follow up, the main cause of LSM decline was found to be due to the degree of change in histological necroinflammation and ALT (336). In a study of 426 CHB patients, a significant LSM decline was observed in patients who received antiviral therapy and had a decline in ALT over a 3 year follow up (333).

The ALT level has been consistently found to affect the LSM in CHB patients. Earlier in this thesis, we reported that elevated ALT increases the LSM (see 5.4.3). Higher ALT tends to reflect more severe hepatic necroinflammation. There is an increase in inflammatory cell infiltrate and oedema in the liver, which causes liver stiffness to increase. This effect has been observed in other studies and shown to be independent of the stages of liver fibrosis (81, 280).

6.4.2 Mild or not significant decline in LSM occurs in patients with quiescent disease, correlating strongest with time

Patients who did not require antiviral treatment and were monitored (group 1 -treatment naïve), and patients with persistently normal ALT (PNALT) irrespective of treatment (49% were treatment naïve, 51% on past or current treatment) predominantly had quiescent disease with minimal liver inflammation. These patients were mostly HBeAg negative (79-93%), had minimal decline in ALT (0- 4 IU/L) and minimal decline in viral load (0.2-0.4 $\log_{10}$) over the follow up period. These cohorts had a mean LSM decline that was very mild (0.2-0.6 kPa), and statistically significant only for the treatment

naïve (group 1) and not for those with PNALT. Although, the mean LSM decline for these subgroups was minimal, some individual patients experienced a marked decline, with 39% and 35% of group 1 and PNALT patients respectively had a marked LSM decline of > 1 kPa.

In the PNALT group, duration between scans was the only correlating factor with the change in LSM at follow up. In group 1, the duration between scans had greater correlation with decline in LSM compared with ALT (Group 1: LSM decline correlation with duration r=0.321, p=0.007 vs correlation with ALT r=0.256, p=0.027). Thus the duration between scans was the most important factor for these patients.

Fibrosis is a bidirectional dynamic process that changes with the function of time. The degree of fibrotic change depends not only on the degree of inflammation and tissue damage to the liver, but also on the duration of injury. Conversely, a longer duration with the absence of, or lesser degrees of injury and inflammation (as represented by ALT normalization), would expect to result in the regenerative capabilities of the liver enabling fibrosis regression over time.

Fibrosis regression over time has been well demonstrated by a study that examined liver histology at baseline, year 1 and year 5 in CHB patients treated with tenofovir (15). At baseline, 38% of participants had bridging fibrosis or worse, but this proportion declined to 28% at year 1 and 12% at year 5 – see **Figure 40**. Our findings are consistent with the current understanding that liver fibrosis regresses as a function of time after removal of the injurious agent.

We were able to elucidate the correlation between LSM and duration of monitoring because the follow up interval was not fixed unlike in other studies (333-336). It allowed us to observe that the time between scans was in itself a critical factor in LSM decline. Prospective studies involving serial measurements of the LSM at regular fixed time intervals for each patient would be useful in confirming the findings.

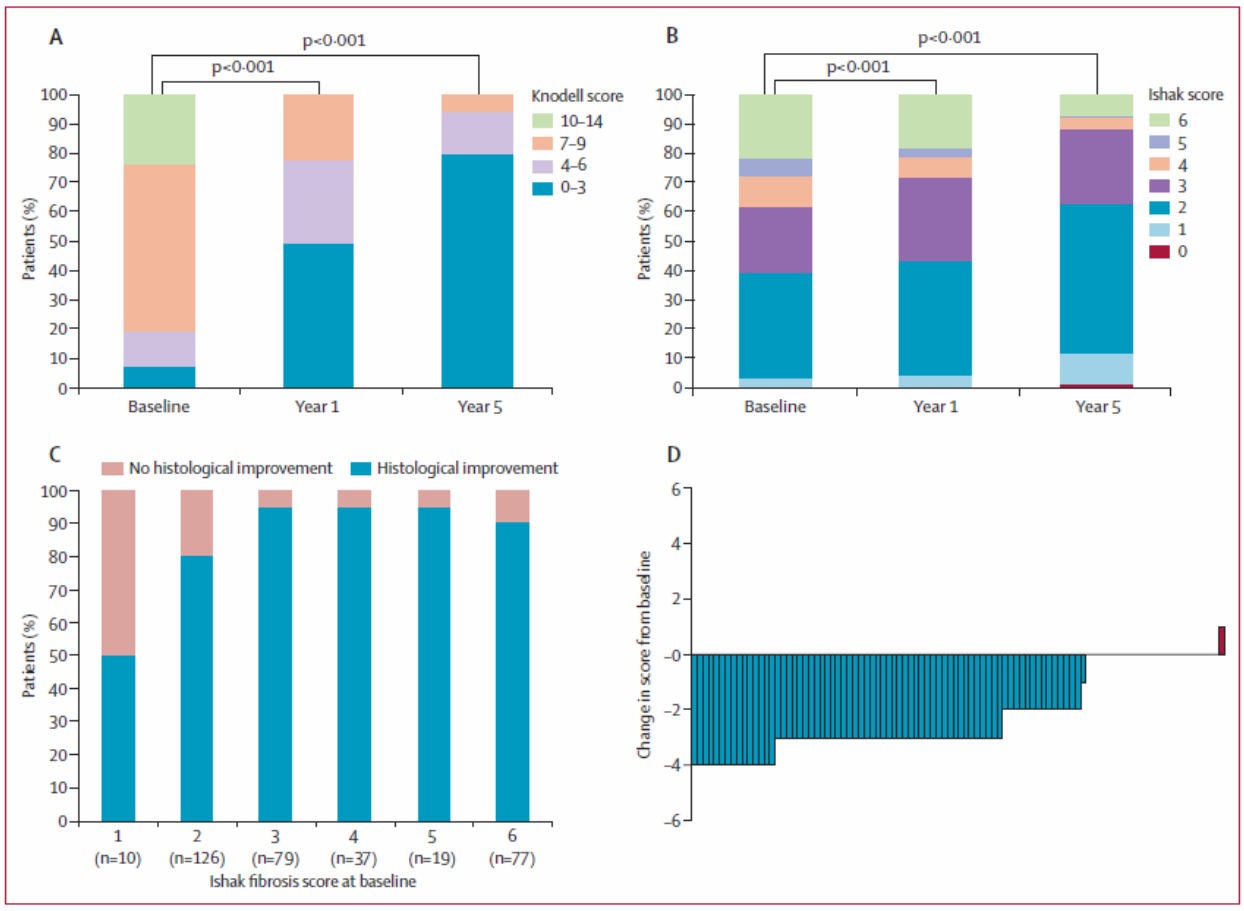


Figure 40: Histology results over 5-year treatment phase

(A) Distribution of Knodell necroinflammatory scores in 348 patients with results available at each time point. (B) Distribution of Ishak fibrosis scores in 348 patients with baseline and year 5 data, and 344 with data for all three time points. (C) Histological response at year 5 according to baseline Ishak fibrosis scores for 348 patients with data available at baseline and year 5. (D) Change from baseline to year 5 in Ishak fibrosis scores for the subset of 96 patients with cirrhosis (Ishak score ≥5) at baseline; each cell represents an individual patients' response. For 24 of 96 patients no changes were noted in Ishak score from baseline. Figure adopted from Marcellin (15)

6.4.3 Limitations of the study

This study is limited by the lack of comparison to liver biopsy. LSM decline in patients with persistently normal ALT was presumed to reflect fibrosis regression. This is a reasonable assumption, but without liver histology, the findings should be interpreted with care.

The study was also limited by the relatively mild liver disease of the study cohort, with low baseline LSM in the study population so that there is very little room for further decline. This may be the reason why a significant decline in LSM was not observed for patients who were on long-term NAs (group 4). The greatest reduction in LSM occurred in those with active disease requiring antiviral therapy (group 2). Presumably, group 4 patients have already experienced biochemical resolution and viral suppression at some stage prior to the study, similar to what group 2 patients experienced during the study period. Hence they may have already undergone most of the potential decline in LSM, thus resulting in the minimal improvement observed during the study.

Another consequence of the study population having low baseline LSM meant that LREs were less likely to occur. Subsequently, there was only one incident case of HCC, and no other liver related morbidities in the patients during the study period. Conclusions about the prognostic significance of the severity of the LSM could not be drawn.

There is a paucity of literature regarding the LSM for patients who obtain favorable clinical end points. Unfortunately, small study numbers prevented any generalisations about the change of the LSM in patients who experienced HBsAg loss (only 2 cases) and patients who achieved a durable response on finite antiviral therapy (group 3 – only 5 cases).

The study examines a retrospective cohort vulnerable to selection bias. Not all chronic hepatitis B patients seen in the outpatient clinic had Fibroscan routinely performed. At the time of the study, the role of Fibroscan had not been completely defined for chronic hepatitis B, and variation in the clinical application of Fibroscan existed between clinicians. One of the perceived uses of Fibroscan is that it is more sensitive in diagnosing compensated cirrhosis that is otherwise not clinically apparent. Patients with obvious clinical features of decompensation would therefore not be scanned, based on the perception that clinical management would not be any different since cirrhosis is already established. This perception lead to cases of overt cirrhosis being excluded, causing a selection bias and probably contributed to our study population overall having low baseline LSM scores. Other issues related to the retrospective aspects of the study include patients who received an initial Fibroscan were often not available for a follow up scan, and clinical data often being incomplete.

This study is the first to describe the correlation between LSM change and length of time in CHB patients with quiescent disease. There is very little data regarding longitudinal LSM monitoring in the treatment of chronic hepatitis B patients. Fibroscans were performed with a high rate of reliability and validity. This also is the first longitudinal study of Fibroscan in Australian CHB patients.

6.4.4 Future directions and the case for LSM to be a stand-alone marker

Our results demonstrate the difficulty in determining whether LSM changes reflected more of fibrosis or inflammation. LSM is influenced by the combination of these factors over the function of time. Our data is unable to objectively quantify to what degree changes in the LSM reflect change in fibrosis or change in the level of inflammatory activity.

The importance of showing LSM change to be reflective of fibrosis is because the fibrosis stage is traditionally perceived as being one of the most important prognostic indicators in CHB (259, 332). Further prospective studies with serial liver biopsies compared with LSM may be able to shed more light on this subject. However, large studies involving liver biopsies are difficult to perform. The poor correlation of LSM decline with fibrosis regression demonstrated by Wong et al (336) from the 48 week study of paired liver biopsies in 71 CHB patients also suggests larger long term studies may not be worthwhile.

A possible better use of the LSM is to perceive it as a stand-alone variable that indicates disease severity. We have demonstrated that the decline in LSM is correlated with decline in ALT, HBeAg seroconversion, HBeAg negativity and viral suppression of which each has been shown to be important in prognosis (260). Perhaps rather than being strictly thought of as a marker for fibrosis, the LSM may be used more meaningfully if that by itself has prognostic implications?

There is a strong rationale for attempting to demonstrate the LSM has use as a standalone variable when considering the problems with using it as an indicator for fibrosis stage as highlighted in chapter 5.4. In brief, histological staging of fibrosis is subject to error from sampling and interobserver variability. Furthermore, histological stage is not a proportionate correlation to fibrosis quantity as each stage represents a morphologic description of the distribution of fibrosis. The correlation with liver stiffness, which presumably measures extracellular matrix components that determine fibrosis quantity, is therefore an imperfect comparison. LSM probably correlates better to a quantifiable marker of liver fibrosis, such as the collagen proportion area (CPA) – see discussion in 5.4.11.

Other studies have already shifted the focus of TE towards prediction of LREs such as HCC (326, 337), and thus there is a growing recognition that TE actually has additional use beyond being limited to measuring the classical end point of liver fibrosis, in which the role is to simply lessen the need for liver biopsy. We attempted to further evaluate the roles of TE in the prediction of long-term disease prognosis. However, the low incidence of HBsAg seroconversion, low numbers of patients who experienced a durable response after finite antiviral treatment and low numbers of LREs (discussed in 6.4.3) prevented meaningful generalisations to be made from our findings on this aspect. Further longitudinal studies that correlate long term outcomes with the LSM cut-offs, while comparing it against currently used clinical parameters and prognostic scoring systems (eg Childs Pugh Score, MELD score) will help determine the usefulness of LSM as a standalone marker that has prognostic significance.

6.5 CONCLUSION

CHB patients, who have active disease and subsequently treated with antivirals, have the greatest decline in LSM. This is associated with ALT normalization, HBeAg seroconversion and viral suppression.

CHB patients who do not require antiviral treatment and/or who have persistently normal ALT irrespective of treatment, have a mild or non-significant decline in LSM. The LSM change is most strongly correlated with length of time and may suggest fibrosis regression.

It is difficult to attribute to what extent the degree of LSM decline is reflective of a reduction in inflammation versus fibrosis regression. The LSM may not be an ideal tool to monitor fibrosis regression. Future studies on the use of the LSM should focus on its potential in being a prognostic tool for liver related morbidity and mortality.

This page has been left intentionally blank

CHAPTER 7: ASSESSMENT OF LIVER FIBROSIS USING TRANSIENT ELASTOGRAPHY AND ASPARTATE AMINOTRANSFERASE PLATELET RATIO IN PATIENTS TREATED WITH METHOTREXATE FOR CHRONIC INFLAMMATORY DISEASE

7.0 CHAPTER SUMMARY

Background: Long term use of methotrexate is considered a risk factor for liver fibrosis. Most of this observation is based on small biopsy series. Guidelines suggest close monitoring and liver biopsy when cumulative doses of 1.0-4.0g are reached (338, 339). The emergence of non-invasive modalities in the evaluation of liver fibrosis have allowed further assessment, and so far do not demonstrate correlation with cumulative methotrexate dose. Using Transient Elastography (TE), we sought that to determine the relationship between liver fibrosis and methotrexate dose, and to identify any other factors associated with fibrosis. We also compared to liver stiffness measurement (LSM) between patients on methotrexate and healthy controls who were not taking methotrexate.

Methods: Patients who had been on methotrexate therapy for at least 6 months were recruited consecutively from outpatient clinics. Total cumulative dose of methotrexate and other relevant data was recorded. Transient Elastography was performed as per the manufacturer's instructions. At least 10 successful measurements were required for a valid Liver Stiffness Measurement (LSM). A reliable scan was defined by having an interquartile range to median ratio of ≤ 30% if the LSM≥7.1 kPa. The LSM cutoff of ≥ 7.1 kPa was used to define Metavir F≥2 from the landmark study by Foucher (see 1.9) (84).

Results: Thirty nine patients on long term methotrexate therapy were recruited. LSM identified 7/39 cases of F≥2 (17.9%). No correlation was found between LSM and methotrexate cumulative dose (r=0.044, p=0.394) or duration of treatment (r=0.018, p=0.457). There was no difference in the mean LSM of patients who had taken a cumulative dose of methotrexate < 1.5g compared with ≥ 1.5g: 5.0 (SD 1.5) kPa vs 6.4 (SD 4.5) kPa, p=0.214; or <4.0g compared with ≥4.0g: 5.9 (SD 2.3) kPa vs 6.7 (SD 5.9) kPa, p=0.618. Independent predictors of LSM included albumin (r = -0.350, p=0.014) and platelet count (r= -0.357, p=0.013). The LSM cutoff indicative F≥2 was significantly associated with BMI≥30: LR 4.442, p=0.029. No difference was found in the mean LSM of methotrexate subjects and a matched population: 6.3 (SD 4.2) kPa vs 4.8 (SD 1.3) kPa, p=0.090.

Conclusions: In this cohort of 39 patients on at least 6 months of methotrexate therapy, dose and duration were not associated with liver stiffness. The only independent predictors of LSM were lower albumin, lower platelet counts and BMI>30. LSM was not increased in methotrexate subjects when compared to controls. Fibroscan appears to be a useful tool in monitoring long term methotrexate induced liver fibrosis. Longitudinal studies with in larger numbers that can account for intercurrent chronic liver disease are needed to further evaluate the risk of fibrosis and the exact role of TE.

7.1 BACKGROUND and AIMS

7.1.1 Introduction

Methotrexate is a folate antimetabolite that inhibits DNA synthesis, repair, and cellular replication. Methotrexate irreversibly binds to dihydrofolate reductase, inhibiting the formation of reduced folates, and thymidylate synthetase. Purine and thymidylic acid synthesis are blocked, resulting in cellular replication halting in the S phase (340). This medication has been used for over 40 years as an oncologic therapy, and for the treatment of other pathologies, mainly in chronic inflammatory diseases such as rheumatoid arthritis, juvenile polyarthritis, corticoid-dependent asthma, severe cases of psoriasis, and inflammatory bowel disease (IBD).

7.1.2 Evidence of methotrexate-induced hepatotoxicity and guidelines for monitoring

Hepatotoxicity is well recognized adverse effect of methotrexate. Liver histological changes that occur include steatosis, cellular hypertrophy, anisonucleosis, and liver fibrosis (341). The exact mechanism of injury remains poorly understood (342). Concomitant folate supplementation has been shown to reduce the incidence of abnormal transaminases, but folate depletion has never been directly demonstrated in hepatotoxicity (343).

Studies from the 1970s reported high rates of liver fibrosis (up to 50%) and cirrhosis (up to 25.6%) in psoriasis patients on long term methotrexate, which appeared to be associated with higher cumulative doses (2, 3). Hence, early dermatology guidelines recommended a liver biopsy at a cumulative dose of 1.5 g and then for every additional 1g (2). Although methotrexate had been observed to cause ALT and AST abnormalities when used in high doses for oncologic conditions (344), liver biochemistry was not recommended in these early guidelines. These early dermatological guidelines however, were drawn from studies that were limited by a lack of longitudinal data and proper controls for hepatotoxins such as alcohol, vitamin A and arsenic (5).

Longitudinal data on methotrexate hepatotoxicity became available when its use became popularized for treating inflammatory arthropathies. AST levels were reported to be predictive of liver fibrosis in rheumatoid arthritis (345, 346) and subsequent rheumatology guidelines suggested dose reduction according to AST levels. A liver biopsy was recommended if there was persistent elevation or if there were other risk factors for liver disease (347). In contrast, dermatological society guidelines recommended biopsies without monitoring of liver biochemistry. However, the data in which these recommendations were based had limitations as there were only small studies that did not account for alcohol consumption.

More recent data suggests that methotrexate hepatotoxicity is less prevalent than was previously thought. A retrospective study of 125 patients who had a liver biopsy and who had taken a median cumulative dose of 2.1g found only 4% of subjects developed moderate fibrosis or cirrhosis (4). Fibrosis was not related to cumulative dose, but was associated with diabetes and obesity.

In response to new data and the growing recognition of the limitations of early studies (5), the latest guidelines from the National (US) Psoriasis foundation (339) have recommended adoption of some aspects of the monitoring protocol of the American College of Rheumatology (338), which is traditionally less stringent and do not advocate liver biopsies as routine practice.

Patients who are at high risk (ie. pre-existing liver disease), the National Psoriasis foundation still recommends liver biopsy performed at the beginning of methotrexate therapy, and then a repeat biopsy for every additional cumulative dose of 1.0-1.5 g (339).

For low risk patients, the National Psoriasis foundation suggests monitoring as per the American College of Rheumatology guidelines. Monitoring is recommended every 1 month during the first year of treatment, and then every 1-3 months. A liver biopsy is recommended when there is a persistent elevation in transaminases or if there are any clinical or biochemical signs of decompensated liver disease. When a cumulative dose of 3.5 - 4.0g is reached and patients remain asymptomatic with normal biochemistry, the following 3 options are recommended: (a) return to monthly clinical monitoring, (b) liver biopsy or (c) switch to alternate therapy (338).

7.1.3 Recent studies of methotrexate-induced hepatotoxicity using transient elastography

There has been increased research in methotrexate induced liver fibrosis with the availability of Transient Elastography (TE), as previously studies were limited by the need for liver biopsy to accurately assess liver fibrosis. In studies using TE, long-term methotrexate use has been reported to have a prevalence of advanced fibrosis of 4.4 - 8.5% (89, 90, 348-350), which is much lower than early biopsy studies and more in line with recent biopsy studies (4). The other common finding is that liver stiffness was not correlated with the cumulative dose of methotrexate, but instead associated with risk factors for metabolic syndrome, suggesting liver injury is caused by concomitant nonalcoholic fatty liver disease (NAFLD).

7.1.4 Hypothesis and specific objectives

The main hypothesis is that the moderate fibrosis as determined by the LSM is not associated with the cumulative dose of methotrexate consumed, but with risk factors for NAFLD such as type II diabetes and obesity. The specific objectives are:

1. To determine whether the relationship between LSM and methotrexate dose and duration
2. To identify those with moderate fibrosis F≥2 and determine the factors associated, specifically examining the relationship with cumulative doses of methotrexate greater than 1.5g and 4.0g that have been implicated in guidelines.
3. To compare the LSM of patients on methotrexate with a matched control group

7.2 METHODS

7.2.1 Patient Selection and data collection

Patients treated with methotrexate long term and attending the Rheumatology outpatient clinic at Concord Repatriation Hospital (Concord Sydney, Australia) and private specialist rooms were identified. Criteria for recruitment included: age older than 18yrs, minimum 6 months of methotrexate treatment, and no previously documented history of liver disease. All patients gave informed written consent. A total of 46 patients were enrolled from July 2009 to October 2009.

Clinical Assessment was performed using a standardized questionnaire. Age, gender, height, weight, BMI, methotrexate cumulative dose, indication for methotrexate, other anti-inflammatory and immune-modulating therapy (steroids, leflunomide, sulphasalazine, cyclosporine, and abatacept), alcohol intake, presence of co-morbid liver disease and other co-morbid medical history was recorded.

Recent laboratory data (within 2 weeks) in relation to the date of the patient having a Fibroscan was recorded. These were: bilirubin (BR), alkaline phosphatase (ALP), gamma-glutamyl transpeptidase (GGT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), albumin (ALB), and platelet count.

Controls were recruited from healthy volunteers (staff and accompanying persons with patients) who were not taking methotrexate. After matching for age, sex, height, weight, body mass index (BMI) and weekly alcohol intake, 28 controls were included in the analysis.

7.2.2 Liver Stiffness assessment

The performance of the Fibroscan has been described in detail earlier in this thesis (refer to section 1.7). In brief, 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 is ≥ 7.1 kPa (55). All scans were performed by trained operators.

The LSM cut-off values used for moderate fibrosis (F≥2) was ≥ 7.1 kPa (sensitivity 67%, specificity 89%), and for cirrhosis (F=4) ≥ 17.6 kPa (sensitivity 87%, specificity 91%). These were derived from the landmark study by Foucher et al (84), which has been described previously in section 1.9.

7.2.3 Sample size determination

The null hypothesis (H_0) is that the mean LSM of healthy controls (μ_c) is lower compared to the mean LSM of patients on long term methotrexate (μ_m) by a clinically significant amount (d).

That is:

H_0: $\mu_c + d \leq \mu_m$

The mean LSM in healthy controls in our study was found to be 4.8 kPa and standard deviation 1.3 kPa (see ***Table 58***). The LSM cut-off for F1 was 6.5 kPa in chronic hepatitis B patients from our earlier studies (see ***Table 23***). Extrapolating LSM cut-offs is unavoidable as there is no literature for LSM cut-offs for methotrexate induced liver fibrosis. Thus, the mean LSM of methotrexate patients would need to be at least 6.5 kPa to be a clinically significant difference. The difference (d) would therefore need to be 1.7 kPa.

Based on these assumptions, if there truly is no difference between the control group and methotrexate group, a total sample size of 16 subjects (8 in each cohort) will be required to achieve 80% power with a 95% confidence interval.

7.2.4 Statistical analysis

Statistical analysis was performed using the Statistical package for Social Sciences (SPSS) version 21.0. Continuous variables were expressed as the mean (standard deviation) where appropriate. Scale variables between groups were compared using Independent samples T-test. Categorical variables were compared using the chi-square test.

7.3 RESULTS

7.3.1 Transient Elastography assessment in the methotrexate cohort

Forty six subjects were recruited into the study. Clinical records were incomplete for 3 subjects and so 43 subjects had TE attempted. A valid TE was unable to be obtained in 4 subjects which left 39/43 (90.7%) for the final analysis. Reliable scans were obtained in 37/39 (94.8%). The mean liver stiffness for the methotrexate group was 6.3 kPa (SD 4.2). T There were 7/39 (17.9%) and 1/39 (2.6%) subjects with LSM cut-offs predictive of moderate fibrosis and cirrhosis respectively. The results are summarised in ***Table 53.***

Table 53: Transient Elastrography results of methotrexate subjects

Valid Liver Stiffness Measurement	39/43 subjects (90.7%)
Reliable Liver Stiffness Measurement	37/39 subjects (94.8%)
Liver Stiffness (kPa)	6.3 (4.3)
Interquartile range/Median ratio	0.27 (0.15)
Success rate (%)	76.4 (24.0)
Subjects (n) LSM ≥ 7.1 kPa	7/39 subjects (17.9%)
Subjects (n) LSM ≥ 17.6 kPa	1/39 subjects (2.6%)
Values represent the mean with standard deviation in parentheses unless otherwise indicated	

7.3.2 Clinical and anthropometric characteristics of the methotrexate study population

Females comprised 27/39 (69.2%) of total methotrexate group subjects. The mean age was 59.0 (SD 10.5) yrs, mean alcohol intake 30.8 (SD 55.6) g/week and mean BMI 27.9 (SD 6.2) kg/m^2. The mean methotrexate treatment duration was 5.9 (SD 3.4) yrs, while the mean methotrexate cumulative dose was 5.3 (SD 3.5) g. The indication for methotrexate was rheumatoid arthritis for 30/39 (76.9%), psoriasis 5/39 (12.8%) and other inflammatory arthropathies in remaining 4/39 (10.3%) subjects. Rheumatoid factor was found in 21/39 (53.8%). Use of other disease modifying anti-rheumatoid drugs and NSAIDs are described in ***Table 54***. The means of the liver biochemistry and platelet count were within normal range. Diabetes was present in 3/39 (7.7%) of subjects. The results are summarised in ***Table 54***.

Table 54: Anthropometric and clinical characteristics of study subjects

Anthropometric and clinical characteristics of study subjects		
Female		27/39 (69.2%)
Age		59.0 (10.5)
Alcohol (grams/week)		30.9 (51.6)
Height (m)		1.65(0.10)
Weight (kg)		75.8 (19.4)
BMI (kgm^{-2})		27.85 (6.17)
Methotrexate Duration (yrs)		5.95 (3.39)
Methotrexate cumulative dose (g)		5.29 (3.48)
Indication for Methotrexate	Rheumatoid Arthritis	30/39 (76.9%)
	Psoriasis	5/39 (12.8%)
	Other inflammatory arthropathy	4/39 (10.3%)
Rheumatoid Factor		21/39 (53.8%)
Other DMARDS/anti-inflammatory agents	Prednisone	15/39 (38.5%)
	Leflonomide	5/39 (12.8%)
	NSAIDS	17/39 (43.6%)
	Anti-TNF alpha	4/39 (10.3%)
	Hydroxychloroquine	6/39 (15.4%)
	Sulphasalazine	4/35 (10.3%)
	Cyclosporine	1/39 (2.6%)
	Intramuscular Gold	1/39 (2.6%)
	Abatacept	2/39 (5.1%)
Bilirubin (umol/L)		9 (4)
Albumin (g/L)		44 (3)
Alkaline Phosphatase (IU/L)		83 (21)
Gamma Glutamyl Transpeptidase (IU/L)		36 (39)
Alanine Aminotransferase (IU/L)		30 (14)
Aspartate Aminotransferase (IU/L)		26 (9)
Platelet count (x10^9/L)		262 (59)
Type II Diabetes		3/39 (7.7%)

Values represent the mean with standard deviation in parentheses except where there is a "/" which represents number of cases/total cases with percentage in parentheses.

7.3.3 Relationship between methotrexate and liver stiffness

A scatterplot diagram was plotted for liver stiffness compared to the cumulative dose of methotrexate *(Figure 41),* and the duration of methotrexate treatment *(Figure 42).* Logical regression performed for liver stiffness found no significant correlation to methotrexate cumulative dose (r=0.044, p=0.394) and duration of treatment (r=0.018, p=0.457). There was no significant difference in the mean liver stiffness of those who had taken a cumulative dose of methotrexate < 1.5g compared with ≥ 1.5g: 5.0 (SD 1.5) kPa vs 6.4 (SD 4.5) kPa, p=0.214; or < 4.0g compared with ≥ 4.0g: 5.9 (SD 2.3) kPa vs 6.7 (SD 5.9) kPa, p=0.618.

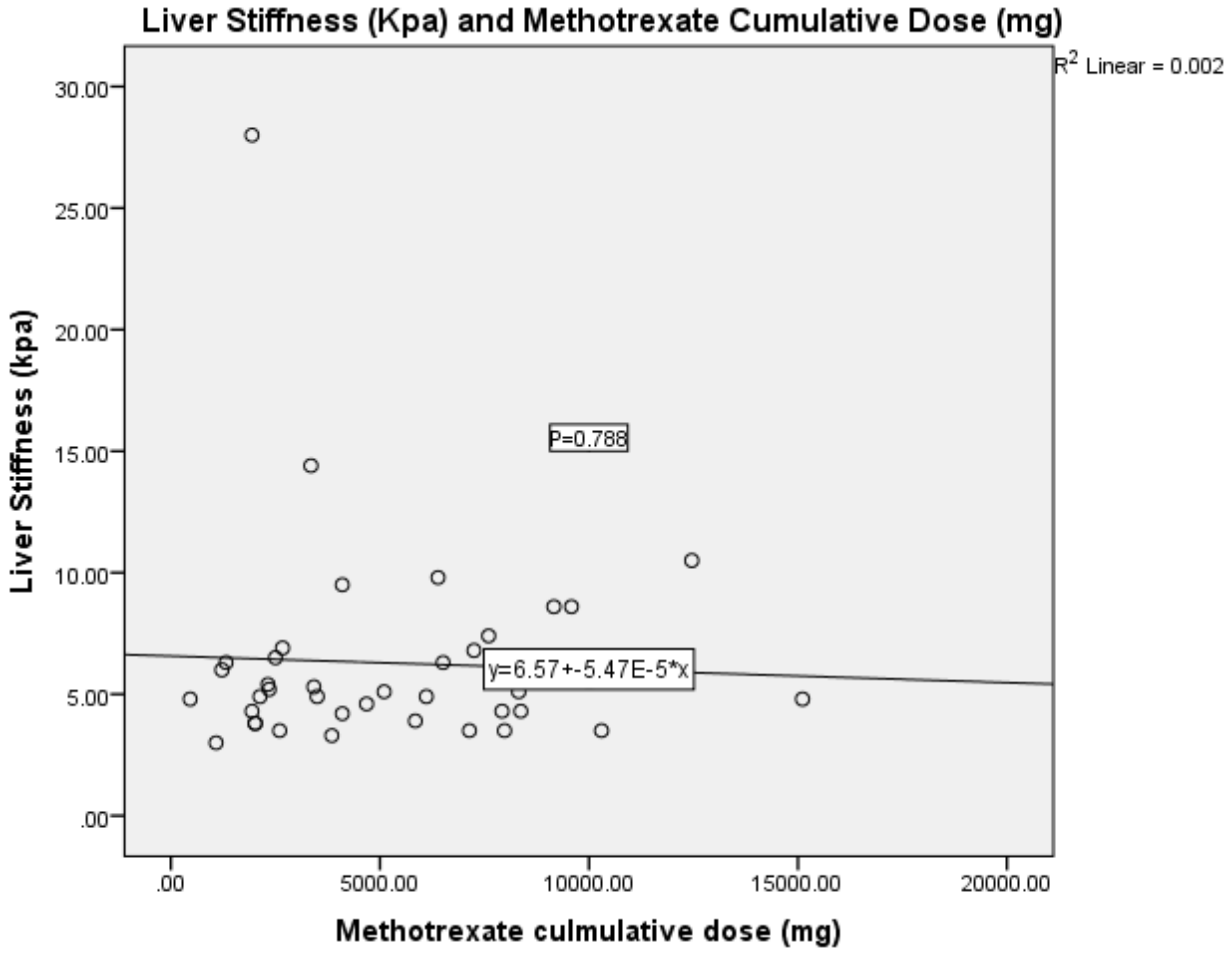


Figure 41: Scatterplot of methotrexate cumulative dose and liver stiffness with a line of best fit

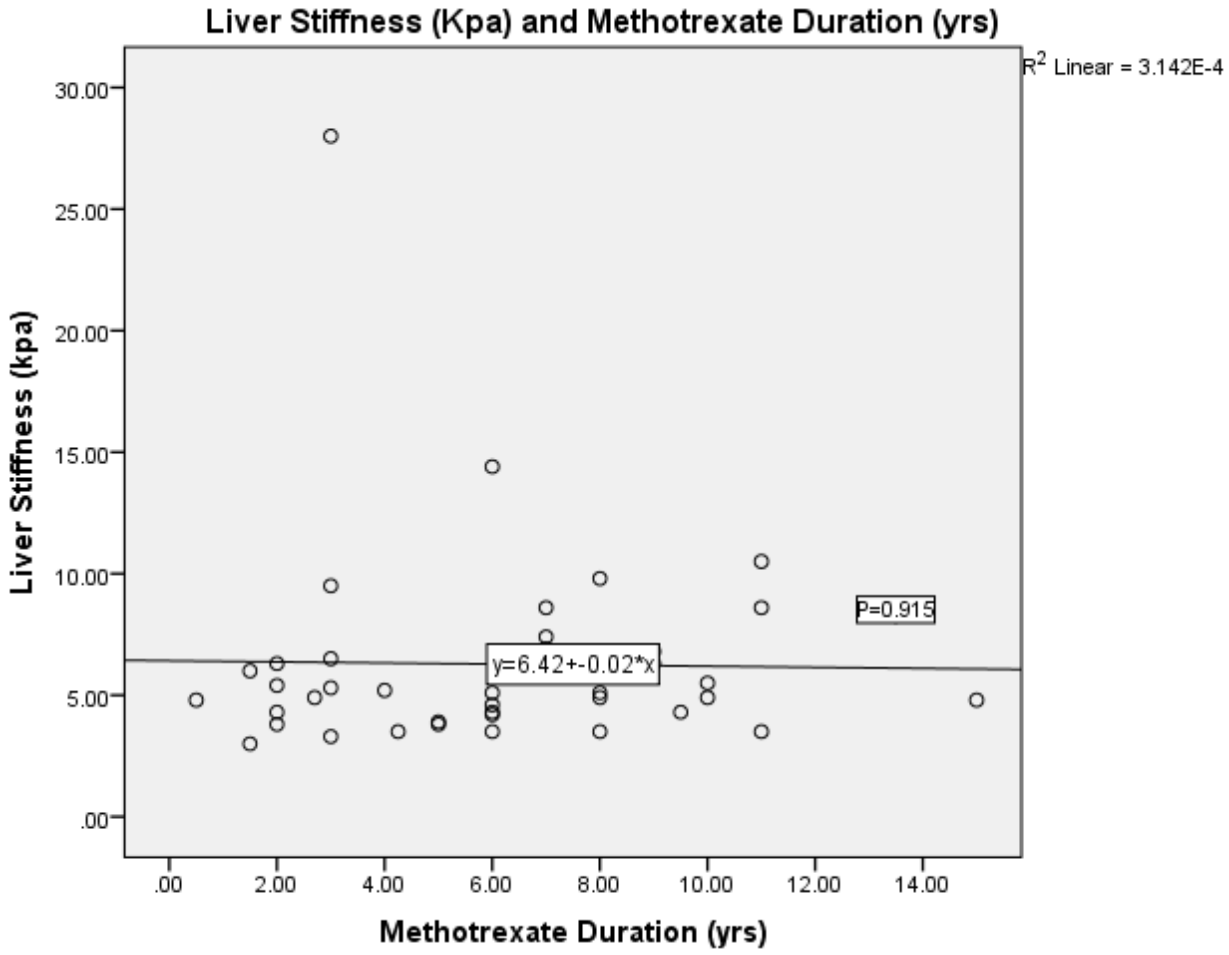


Figure 42: Scatterplot of methotrexate treatment duration and liver stiffness with a line of best fit

7.3.4 Relationship between Liver Stiffness and other clinical variables

The relationship between liver stiffness and other clinical variables apart from methotrexate was examined. Continuous variables were correlated against the LSM. Categorical variables; the mean LSM of each level within the category was compared. Multivariate analysis of the continuous variables showed significant correlations with albumin (r = -0.350, p=0.014) and platelet count (r= -0.357, p=0.013), see ***Table 55***. Age was significantly associated in univariate analysis, but not in multivariate analysis. No significant differences were found in the mean LSM between the levels for each categorical factor that was examined – see ***Table 56.***

Table 55: Correlation of continuous variables with liver stiffness

	Correlation coefficient (r)	P-value
Age	0.283	0.041[b]
Alcohol (grams/week)	-0.073	0.330
BMI (kgm^{-2})	0.064	0.349
Methotrexate Duration (yrs)	-0.018	0.457
Methotrexate cumulative dose (g)	-0.044	0.394
Bilirubin (umol/L)	0.007	0.484
Albumin (g/L)	-0.350	0.014[a]
Alkaline Phosphatase (IU/L)	-0.061	0.357
Gamma Glutamyl Transpeptidase (IU/L)	0.053	0.374
Alanine Aminotransferase (IU/L)	0.044	0.396
Aspartate Aminotransferase (IU/L)	0.076	0.323
Platelet count (x10^9/L)	-0.357	0.013[a]

a. ALB and PLT significant after multivariate analysis
b. Age was not significant in multivariate analysis

Table 56: Comparison of Categorical Variables and Liver Stiffness

Variable		N	Liver Stiffness (kpa) [a]	p-value
Gender	Male	12	7.9 (6.6)	0.262
	Female	27	5.6 (2.5)	
Methotrexate Cumulative Dose (g)	< 1.5	4	5.0 (1.5)	0.214
	> 1.5	35	6.4 (4.5)	
Methotrexate Cumulative Dose (g)	> 4.0	21	5.9 (2.3)	0.618
	< 4.0	18	6.7 (5.9)	
Rheumatoid Factor	Positive	21	5.4 (1.8)	0.174
	Negative	18	7.3 (5.9)	
Type II Diabetes?	Yes	3	12.5 (13.5)	0.478
	No	36	5.8 (2.4)	
Known Liver Disease?	Yes	4	11.2 (11.2)	0.401
	No	35	5.7 (2.5)	
Prednisone	Yes	15	5.4 (2.1)	0.236
	No	24	6.8 (5.2)	
Leflonomide	Yes	5	5.2 (2.0)	0.325
	No	34	6.4 (4.5)	
NSAIDs	Yes	17	6.2 (3.0)	0.900
	No	22	6.4 (5.1)	
Anti-TNF alpha agent	Yes	4	8.0 (4.3)	0.404[b]
	No	35	6.1 (4.3)	
Hydroxychloroquine	Yes	6	4.7 (1.2)	0.052
	No	33	6.6 (4.6)	
Sulphasalazine	Yes	4	6.2 (2.0)	0.959
	No	35	6.3 (4.5)	
Cyclosporine	Yes	1	9.8	0.412
	No	38	6.2 (4.3)	
Intramuscular Gold	Yes	1	6.8	0.904
	No	38	6.3 (4.3)	
Abatacept	Yes	2	5.1 (0.4)	0.124
	No	37	6.3 (4.3)	
BMI (kgm^2)	<25	14/39	6.9 (6.5)	N/A[c]
	25-30	13/39	4.6 (1.0)	
	>30	12/39	7.3 (2.8)	
Methotrexate Indication	Psoriasis	5/39	11.4 (10.2)	N/A[c]
	Rheumatoid Arthritis	30/39	5.6 (2.0)	
	Other Inflammatory Arthropathy	4/39	5.2 (1.6)	

a. Mean LSM with standard deviation in parentheses.
b. Independent sample T-tests performed for all variables with 2 levels. Equal variances were not assumed except anti-tnf alpha agent
c. One way ANOVA attempted for BMI and Methotrexate indication. Tests of homogeneity failed and so p-value cannot be interpreted.

7.3.5 Characteristics of subjects with LSM ≥ 7.1 kPa (indicative of moderate fibrosis)

The characteristics of subjects with LSM ≥ 7.1 kPa were compared in those with LSM < 7.1 kPa. This cut-off was chosen to be indicative of moderate fibrosis as discussed in section 7.2.2. BMI ≥ 30 was the only significantly associated variable in subjects with an LSM ≥ 7.1 kPa. The corresponding likelihood ratio (LR) of having an LSM ≥ 7.1 kPa when the BMI≥ 30 compared to BMI < 30 was: 4.442 (p=0.029). No other characteristic, including the methotrexate cumulative dose and duration, was found to increase the likelihood of having LSM ≥ 7.1 kPa – see ***Table 57.***

Table 57: Comparison of variables for LSM <7.1kpa and LSM ≥7.1kpa

	LSM <7.1 kPa Mean (SD)	LSM≥7.1 kPa Mean (SD)	P value
Scale Variables[ab]			
Age	58.3 (10.7)	61.8 (9.9)	0.413
Methotrexate Duration (yrs)	5.7 (3.5)	7.0 (3.1)	0.331
Methotrexate cumulative dose (g)	4.842 (3.396)	6.822 (3.565)	0.154
Alcohol (grams/week)	33.4 (56.6)	21.3 (24.2)	0.560
BMI (kgm^{-2})	27.7 (6.2)	28.5 (6.3)	0.762
Bilirubin (µmol/L)	8 (4)	10 (5)	0.268
Albumin (g/L)	45 (3)	43 (4)	0.207
Alkaline Phosphatase (IU/L)	83 (23)	84 (14)	0.938
Gamma Glutamyl Transpeptidase (IU/L)	37 (44)	30 (13)	0.638
Alanine Aminotransferase (IU/L)	29 (14)	34 (15)	0.346
Aspartate Aminotransferase (IU/L)	25 (9)	28 (6)	0.384
Platelet count (x10^9/L)	271 (57)	227 (57)	0.063
Categorical Variables[ab]	LSM <7.1 kPa N cases (%)	LSM ≥ 7.1 kPa N cases (%)	P value
Gender (male)	8/31 (25.8%)	4/8 (50%)	0.186
Type II Diabetes	2/31 (6.5%)	1/8 (12.5%)	0.567
BMI ≥25	20/31 (64.5%)	5/8 (62.5%)	0.916
BMI ≥30	7/31 (22.6%)	5/8 (62.5%)	0.029
			LR[d] 4.442
BMI ≥35	4/31 (12.9%)	1/8 (12.5%)	0.976
Rheumatoid factor	18/31 (58.1%)	3/8 (37.5%)	0.298
Methotrexate Indication[c]			
Rheumatoid Arthritis	25/31 (80.6%)	5/8 (62.5%)	0.476
Psoriasis	3/31 (9.7%)	2/8 (25.0%)	
Other Inflammatory Arthropathy	3/31 (9.7%)	1/8 (12.5%)	
Prednisone	13/31 (41.9%)	2/8 (25.0%)	0.380
Leflonomide	4/31 (12.9%)	1/8 (12.5%)	0.976
NSAIDS	12/31 (38.7%)	5/8 (62.5%)	0.226
Anti-TNF alpha agent	3/31 (9.7%)	1/8 (12.5%)	0.815
Hydroxychloroquine	6/31 (19.4%)	0/8 (0%)	0.176
Sulphasalazine	3/31 (9.7%)	1/8 (12.5%)	0.815
Cyclosporine	0/31 (0%)	1/8 (12.5%)	0.205
Intramuscular Gold	1/31 (3.2%)	0/8 (0%)	0.795
Abatacept	2/31 (6.5%)	0/8 (0%)	0.628

a. Independent T-test applied for scale variables, while Chi-square and Fisher's tests applied to categorical variables
b. Values reported for scale variables represent the mean and standard deviation in parentheses. Values reported for categorical variables are proportions with percentages in parentheses.
c. Chi Square statistic was applied across all levels within the categorical variable of Methotrexate indication and BMI. Null hypothesis being no difference in proportion of moderate fibrosis across all levels.
d. Likelihood ratio of having Moderate Fibrosis if BMI > 30.

7.3.6 Comparison of LSM between methotrexate subjects and matched controls

Methotrexate were compared with controls matched for gender, age, alcohol use and BMI (see ***Table 58***). There was no significant difference between methotrexate subjects and controls for mean LSM: 6.3 (SD 4.2) kPa vs 4.8 (SD 1.3) kPa, p=0.090; or frequency of LSM ≥ 7.1 kPa (7/39 subjects vs 1/28, p=0.126). There was also no difference in the mean IQR/median ratio, proportion of reliable scans, proportion of valid scans, number of valid scans obtained or the mean scan success rate ***(see Table 58).***

Table 58: Comparison of between methotrexate and control subjects

Compared variables	Methotrexate	Control	P- value
Subjects	39	28	
Women	27 (69.2%)	16 (57.1%)	0.448
Age (yrs)	59.0 (10.5)	55.0 (12.5)	0.165
ETOH(g/wk)	30.9 (51.6)	24.5 (57.6)	0.633
BMI kgm^{-2}	27.9 (6.2)	27.6 (6.4)	0.897
LSM (kpa)	6.3 (4.3)	4.8 (1.3)	0.090
LSM≥7.1kpa (F≥2)	7/39 (17.9%)	1/28 (3.6%)	0.126
IQR/M ratio	0.27 (0.15)	0.25 (0.15)	0.589
Reliable scans	37/39 (94.5%)	28/28 (100%)	0.506
Total Valid scans	10.9 (3.0)	11.3 (4.2)	0.589
Success Rate (%)	76.4 (24.0)	69.1 (30.6)	0.277

Independent T tests applied to scale variables while Fisher's exact test applied to categorical variables.

Values represent the mean for scale variables with standard deviation in parentheses and proportions for categorical variables with percentages In parentheses.

7.3.7 Summary of main findings

- No correlation was found in subject's LSM and methotrexate cumulative dose (r=0.044, p=0.394) or duration of methotrexate treatment (r=0.018, p=0.457)
- No difference was found in the mean LSM of patients who had taken a cumulative dose methotrexate < 1.5g vs ≥ 1.5g: 5.0 (SD 1.5) kPa vs 6.4 (SD 4.5) kPa, p=0.214
- No difference was found in the mean LSM of patients who had taken a cumulative dose methotrexate <4.0g vs ≥ 4.0g: 5.9 (SD 2.3) kPa vs 6.7 (SD 5.9) kPa, p=0.618
- LSM identified 7/39 cases (17.9%) of LSM ≥7.1 kPa – the chosen cut-off to indicate moderate fibrosis
- An LSM≥7.1 kPa was significantly associated with BMI≥30: LR 4.442, p=0.029. Albumin (r = -0.350, p=0.014) and platelet count (r= -0.357, p=0.013) were also independent predictors of LSM
- No difference in the mean LSM of methotrexate subjects and matched controls: 6.3 (SD 4.2) kPa vs 4.8 (SD 1.3) kPa, p=0.090.

7.4 DISCUSSION

7.4.1 BMI ≥ 30 kgm^{-2} is associated with LSM ≥7.1 kPa, but not methotrexate duration or dose.

In our study there was no evidence of any association between the LSM and the methotrexate cumulative dose or treatment duration. Instead an LSM ≥7.1 kPa (cut-off chosen for Metavir F2) was associated with BMI ≥30 kgm^{-2} (LR 4.44, p=0.029). The LSM being associated with obesity rather than methotrexate suggests the predominant aetiology for liver stiffness is more likely to be NAFLD in this particular study population.

Similar findings have been reported in recent studies that have assessed the LSM in methotrexate subjects. The largest study to date investigated 518 patients treated with methotrexate for a variety of chronic inflammatory conditions. Advanced liver fibrosis (determined by Fibroscan or Fibrotest) was not associated with methotrexate dose or length of treatment duration, but was predicted by BMI > 28 and excessive alcohol consumption (350). Three smaller studies of 46, 53 and 54 patients also found no association with methotrexate dose and liver stiffness (90, 348, 349). In these cohorts, a total of 7 subjects were found to have significant fibrosis, in which 4 were attributed to morbid obesity and alcohol abuse (348, 349).

Furthermore, a liver biopsy study of 125 patients who had taken a median cumulative dose of 2.1g methotrexate reported that fibrosis was not related to cumulative dose, but diabetes and obesity (4). Our data and recent literature consistently indicate that fibrosis is unrelated to the methotrexate dose, and that metabolic syndrome risk factors are more important.

In our study, liver stiffness was also correlated by lower albumin levels and platelet count. The capability for diagnosing LSM≥7.1 kPa was further explored in a side analysis (not shown in results). The platelet count AUROC= 0.750 (95% CI 0.549-0.951, p=0.031) while for albumin, results were not statistically significant. The best sensitivity and specificity for predicting LSM≥7.1 kPa was 74.2% and 87.5%, at a platelet level of 253 x 10^9/L. This moderate level of accuracy is unsurprising, as it is well recognized that lower platelet levels can indicate progressive liver damage, but not at a high enough accuracy as a stand-alone marker. Hence the frequent inclusion of platelet count in combination with other markers as part of fibrosis predicting panels, such as the FIB4-I, APRI and API which have been discussed in earlier chapters.

7.4.2 The prevalence of fibrosis in methotrexate patients is lower in recent studies

Earlier studies had suggested rates of fibrosis up to 50% and cirrhosis up to 26% (2). Our study results show that LSM ≥ 7.1 kPa (cut-off for Metavir F2) occurred in 17.9%, while LSM ≥ 17.6 kPa (cut-off for Metavir F4) occurred in 2.6% of the study population. Other studies using Fibroscan or Fibrotest have shown similar prevalence rates. The rate of moderate fibrosis or worse (Metavir F≥2), advanced fibrosis or worse (Metavir F≥3), and cirrhosis (Metavir F4) have been reported to be 17.4%, 6.5-8.5% and 0-2.1% respectively (90, 348-350). In studies that assessed liver biopsy, moderate fibrosis and cirrhosis were each described in only 2% of patients (4).

Furthermore there have also been studies that have compared the prevalence of fibrosis amongst patients with chronic inflammatory disorders who are on methotrexate treatment versus methotrexate naïve. The largest study which comprised of 518 patients with a variety of chronic inflammatory conditions, reported no difference in the median LSM between cases and controls (350). Another study of 54 patients with Crohn's disease also found no difference in the median LSM (90).

Overall, these results suggest that methotrexate is associated with a low risk for developing fibrosis. It would appear that early studies probably overestimated the prevalence of liver fibrosis in methotrexate users. Experts have increasingly recognized that early studies lacked proper controls for hepatotoxins such as alcohol and pre-existing liver disease (5, 351), lending support to the findings in more recent studies.

We reported the comparison of the LSM between methotrexate subjects and population controls who were not on methotrexate. By showing that the mean LSM and the rate of Metavir F≥2 (as determined by LSM ≥ 7.1 kPa) is no different to healthy controls, this raises the question of whether methotrexate really is a significant risk factor for liver fibrosis at all. Possibly, a more appropriate comparison would be to assess the prevalence in methotrexate users against the general population. While the rate of fibrosis in the general population is unknown, the rate of cirrhosis is estimated to be at 0.8% (352). Our data and other recent studies so far indicate the rate of cirrhosis in methotrexate users is around 2%, which is in contrast with early studies reporting rates of up to 26% (2).

The pathological features of methotrexate induced liver toxicity have been described to closely resemble NASH (353). While it is suggested that methotrexate aggravates pre-existing NASH (338, 354), this is based on expert opinion and has not been substantiated with data. There is a possibility that the observation of liver fibrosis in methotrexate is coincidental, and is actually mostly due to NAFLD because it is so common (95, 133). Further prospective studies would be helpful in evaluating this area.

7.4.3 Cautious interpretation of results, implications for guidelines and use of Fibroscan in monitoring methotrexate patients

The findings in our study should be interpreted with caution due to the small size of the study population limiting generalizability. In addition, although in population controls the age, gender, BMI and alcohol intake were adequately controlled, NALFD was not able to be directly assessed and controlled for. The Fibroscan model at the time of the study had yet to feature the CAP score for diagnosing steatosis, and other forms of imaging such as ultrasound was available for only a few subjects. The exclusion of chronic viral hepatitis and other liver disease was based on clinical history and medical records alone, as the relevant confirmatory laboratory tests were unavailable. However, it is unlikely this last issue led to any significant impact on the findings as the prevalence of these conditions combined is estimated be only 2.5% in the local population (113-115). Another criticism may be that Fibroscan has yet to be widely validated by liver biopsy for its use in methotrexate induced liver fibrosis. One study assessed the accuracy of Fibroscan against liver biopsy in 24 methotrexate patients, and reported that the ≥ 7.1 kPa cut-off had 88% specificity for identifying F≥2 (89). This suggests extrapolating the cut-offs from the landmark study by Foucher et al (84) is a reasonable strategy.

Current guidelines maintain very careful and strict recommendations regarding the monitoring of hepatotoxicity. Until there are rigorous longitudinal studies that assess liver fibrosis pre and post methotrexate treatment, it is remains prudent to adopt a vigilant approach. Such studies must account comprehensively for intercurrent liver diseases.

Lastly, in this study Fibroscan was demonstrated to be feasible for the assessment of liver stiffness in methotrexate patients, with valid measurements obtained in 90.7%, and reliable measurements in 94.8% of subjects. The reliability, reasonable accuracy, high acceptance and noninvasive nature of Fibroscan suggest that it can be used in place of liver biopsy when monitoring for hepatic fibrosis in long term users of methotrexate. This may also allow for more frequent monitoring of hepatic fibrosis where repeated liver biopsy is undesirable. In cases where liver biopsy is contraindicated, Fibroscan is likely to be a sufficient alternative.

7.5 CONCLUSION

In long-term users of methotrexate, dose and duration are not associated with liver stiffness. Cumulative doses of 1.5g and 4.0g, where current guidelines suggest liver biopsy, was not associated with a high LSM indicative of F≥2. Instead obese range BMI was found to be a risk factor for high LSM. The rate of F≥2 and F4 (as assessed with Fibroscan) are lower than reports from early studies, and consistent with more recent data. The mean LSM in methotrexate treated patients was no different to healthy population controls who were not on methotrexate. Further studies that adequately account for NAFLD and other liver diseases are required to accurately establish the risk of developing methotrexate induced fibrosis.

GENERAL CONCLUSION

The goals of research were to further explore the clinical utility of transient elastography, targeting the liver disease populations which had a paucity of literature at the time. These were patients with NAFLD, chronic hepatitis B and long term users of methotrexate.

In a pilot study, we established that a significant proportion of diabetes patients had high liver stiffness suggestive of advanced fibrosis. However, the usefulness of TE was somewhat limited because of the high rate of invalid or unreliable scans in obese patients. Our review and meta-analysis confirmed the accuracy and widespread validation of using TE to assess fibrosis in NAFLD, and that the XL probe ameliorates concerns over its usefulness in obese patients. Thus we proceeded to examine a much larger group of diabetes patients. TE was performed validly and reliably in 98.2% of 1918 diabetes patients using the M or XL probes. NAFLD was diagnosed in 72.7% using the CAP, and advanced fibrosis was found in 17.7% using indicative LSM cut-offs. Diabetic patients with high BMI and dyslipidaemia were at particularly high risk. Our data supports liver assessment to be incorporated as part of routine screening for complications in diabetes patients.

We demonstrated that TE is an accurate and reliable noninvasive tool for assessing fibrosis stage in chronic hepatitis B. TE especially has excellent accuracy for advanced fibrosis and cirrhosis, and is superior to FIB 4, APRI, AAR, API and FI. Our results are consistent with more recent validation studies. Different LSM cut-offs have been reported because of variability in study populations. This can make deciding which LSM cut-offs to apply in the local setting difficult. Our study on local CHB patients was thus able to provide LSM cut-offs that would most accurately reflect and can be applied in our local patient population. TE is most rationally applied by determining high and low cut-offs that can “rule in” or “rule out” fibrosis with very high degree of specificity or sensitivity respectively. Higher LSM cut-offs for high ALT patients are needed to maintain the accuracy of TE. Our study showed that by incorporating TE into the assessment of CHB patients, liver biopsy can be avoided in the majority of patients.

Two distinct patterns of LSM decline in CHB patients were described in our 31 month longitudinal study. Patients with active disease and subsequently treated with antivirals have the largest decline in LSM, while those with quiescent disease have only a mild or negligible decline in LSM. ALT normalization, HBeAg seroconversion and viral suppression are important factors that are associated with LSM decline in those with active disease. Hence the level of inflammatory activity has a strong influence on the LSM in those with active disease. In those with quiescent disease, length of time is the most important influence to LSM decline. A longer period of time allows for the regenerative ability of the liver to regress fibrosis to a greater capacity. Hence, LSM decline may be more reflective of fibrosis regression in patients with quiescent CHB infection.

Performing TE in long term methotrexate users revealed that cumulative dose and duration are not associated with liver fibrosis. The dose thresholds of 1.5g and 4.0g whereby liver biopsy is recommended were not found to be associated, but instead, obese BMI was the only risk factor. The rate of moderate fibrosis and cirrhosis is low, which contrasts with early studies, but consistent with recent data. The mean LSM is not significantly different in methotrexate users compared population controls. The risk of liver fibrosis in methotrexate users is likely to be low and may not be much higher

than non-methotrexate users from the general population. More studies that assess the pre and post methotrexate liver fibrosis are required to clarify this issue further.

One of the recurrent discussion themes was the potential use of the LSM as a standalone marker. There are a number of problems with using the LSM to assess fibrosis stage. Liver biopsy has several limitations, one of which is interobserver variability. This was demonstrated by our findings of 44% biopsies being staged differently upon a second assessment. The variability in histological assessment subsequently limits evaluation of the diagnostic accuracy of TE and any other non-invasive markers that are compared against liver biopsy. Furthermore, fibrosis stage is a morphological description of the distribution of fibrosis. Evaluating it against the liver stiffness, which best reflects fibrosis quantity, is an awkward comparison.

Thus LSM may be better considered as a standalone variable, but it needs to have prognostic value. We attempted to correlate the LSM in chronic hepatitis B patients with important clinical outcomes such as HbsAg loss and cirrhotic complications. Due to small study numbers and relative infrequency of these events, we were not able to make any useful generalisations. The future of Fibroscan will be to determine the prognostic significance of the LSM values. Our large study of diabetes patients in chapter 4 was only the report of the baseline assessment. Prospective follow-up is planned for another 10 years. The prognostic implications of liver stiffness measurement will be unraveled in the near future.

In conclusion, transient elastography is a valuable tool for the noninvasive diagnosis of fibrosis in nonalcoholic fatty liver disease, chronic hepatitis B and liver fibrosis in patients who use methotrexate. We showed that TE has a high level of diagnostic accuracy for CHB in our validation study and for NAFLD in our meta-analysis. CAP and TE are useful for screening NAFLD and advanced fibrosis in diabetes patients. TE is useful for establishing the baseline level of fibrosis and may also be a useful marker in the monitoring the progress of CHB patients. As long-term study results become available, the LSM may be considered a standalone variable that has prognostic implications rather than simply being an estimation of the fibrosis stage.

This page has been left intentionally blank

8.0 REFERENCES

1. Yoneda M, Yoneda M, Mawatari H, Fujita K, Endo H, Iida H, et al. Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with nonalcoholic fatty liver disease (NAFLD). Dig Liver Dis. 2008;40(5):371-8.
2. Roenigk HH, Jr., Auerbach R, Maibach HI, Weinstein GD. Methotrexate guidelines--revised. Journal of the American Academy of Dermatology. 1982;6(2):145-55.
3. Lewis JH, Schiff E. Methotrexate-induced chronic liver injury: guidelines for detection and prevention. The ACG Committee on FDA-related matters. American College of Gastroenterology. Am J Gastroenterol. 1988;83(12):1337-45.
4. Berends MA, Snoek J, de Jong EM, van de Kerkhof PC, van Oijen MG, van Krieken JH, et al. Liver injury in long-term methotrexate treatment in psoriasis is relatively infrequent. Aliment Pharmacol Ther. 2006;24(5):805-11.
5. Zachariae H. Liver biopsies and methotrexate: a time for reconsideration? Journal of the American Academy of Dermatology. 2000;42(3):531-4.
6. Heidelbaugh JJ, Bruderly M. Cirrhosis and chronic liver failure: part I. Diagnosis and evaluation. Am Fam Physician. 2006;74(5):756-62.
7. Cowie BC, MacLachlan JH. The global burden of liver disease attributable to hepatitis B, hepatitis C, and alcohol: Increasing mortality, differing causes. Hepatology. 2013;1):218A-9A.
8. WorldHealthOrganisation 2014;Pages. Accessed at The World Health Organisation at http://www.who.int/mediacentre/factsheets/fs310/en/ on 6th June 2015 2015.
9. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet. 2014;384(9945):766-81.
10. Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Aliment Pharmacol Ther. 2011;34(3):274-85.
11. LaBrecque DR, Abbas Z, Anania F, Ferenci P, Khan AG, Goh K-L, et al. World Gastroenterology Organisation Global Guidelines: Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis. Journal of Clinical Gastroenterology. 2014;48(6):467-73.
12. Statistics ABo 2012;Pageswww.abs.gov.au on 14th June 2015.
13. GESA DAEa. The economic cost and health burden of liver diseases in Australia. GESA; 2013.
14. Poynard T, Mathurin P, Lai CL, Guyader D, Poupon R, Tainturier MH, et al. A comparison of fibrosis progression in chronic liver diseases. J Hepatol. 2003;38(3):257-65.
15. Marcellin P, Gane E, Buti M, Afdhal N, Sievert W, Jacobson IM, et al. Regression of cirrhosis during treatment with tenofovir disoproxil fumarate for chronic hepatitis B: a 5-year open-label follow-up study. The Lancet. 2013;381(9865):468-75.
16. D'Ambrosio R, Aghemo A, Rumi MG, Ronchi G, Donato MF, Paradis V, et al. A morphometric and immunohistochemical study to assess the benefit of a sustained virological response in hepatitis C virus patients with cirrhosis. Hepatology. 2012;56(2):532-43.
17. Wong VW, Wong GL, Choi PC, Chan AW, Li MK, Chan HY, et al. Disease progression of non-alcoholic fatty liver disease: a prospective study with paired liver biopsies at 3 years. Gut. 2010;59(7):969-74.
18. Friedman SL. Hepatic stellate cells: protean, multifunctional, and enigmatic cells of the liver. Physiol Rev. 2008;88(1):125-72.
19. Pinzani M, Macias-Barragan J. Update on the pathophysiology of liver fibrosis. Expert Rev Gastroenterol Hepatol. 2010;4(4):459-72.
20. Trautwein C, Friedman SL, Schuppan D, Pinzani M. Hepatic fibrosis: Concept to treatment. J Hepatol. 2015;62(1 Suppl):S15-24.

21. Liu C, Chen X, Yang L, Kisseleva T, Brenner DA, Seki E. Transcriptional repression of the transforming growth factor beta (TGF-beta) Pseudoreceptor BMP and activin membrane-bound inhibitor (BAMBI) by Nuclear Factor kappaB (NF-kappaB) p50 enhances TGF-beta signaling in hepatic stellate cells. J Biol Chem. 2014;289(10):7082-91.
22. Marra F, Tacke F. Roles for chemokines in liver disease. Gastroenterology. 2014;147(3):577-94.e1.
23. Tacke F, Zimmermann HW. Macrophage heterogeneity in liver injury and fibrosis. J Hepatol. 2014;60(5):1090-6.
24. Wasmuth HE, Zaldivar MM, Berres ML, Werth A, Scholten D, Hillebrandt S, et al. The fractalkine receptor CX3CR1 is involved in liver fibrosis due to chronic hepatitis C infection. J Hepatol. 2008;48(2):208-15.
25. de Alwis NM, Day CP. Non-alcoholic fatty liver disease: the mist gradually clears. J Hepatol. 2008;48 Suppl 1:S104-12.
26. Novo E, Parola M. Redox mechanisms in hepatic chronic wound healing and fibrogenesis. Fibrogenesis Tissue Repair. 2008;1(1):5.
27. Schnabl B, Brenner DA. Interactions between the intestinal microbiome and liver diseases. Gastroenterology. 2014;146(6):1513-24.
28. Novo E, Cannito S, Zamara E, Valfre di Bonzo L, Caligiuri A, Cravanzola C, et al. Proangiogenic cytokines as hypoxia-dependent factors stimulating migration of human hepatic stellate cells. Am J Pathol. 2007;170(6):1942-53.
29. Mann DA. Epigenetics in liver disease. Hepatology. 2014;60(4):1418-25.
30. Olsen AL, Bloomer SA, Chan EP, Gaca MD, Georges PC, Sackey B, et al. Hepatic stellate cells require a stiff environment for myofibroblastic differentiation. Am J Physiol Gastrointest Liver Physiol. 2011;301(1):G110-8.
31. Knodell RG, Ishak KG, Black WC, Chen TS, Craig R, Kaplowitz N, et al. Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. Hepatology. 1981;1(5):431-5.
32. GROUP TFMCS. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. The French METAVIR Cooperative Study Group. Hepatology. 1994;20(1 Pt 1):15-20.
33. Ishak K, Baptista A, Bianchi L, Callea F, De Groote J, Gudat F, et al. Histological grading and staging of chronic hepatitis. J Hepatol. 1995;22(6):696-9.
34. Scheuer PJ. Classification of chronic viral hepatitis: a need for reassessment. J Hepatol. 1991;13(3):372-4.
35. Kim SU, Oh HJ, Wanless IR, Lee S, Han KH, Park YN. The Laennec staging system for histological sub-classification of cirrhosis is useful for stratification of prognosis in patients with liver cirrhosis. J Hepatol. 2012;57(3):556-63.
36. Asselah T, Bieche I, Sabbagh A, Bedossa P, Moreau R, Valla D, et al. Gene expression and hepatitis C virus infection. Gut. 2009;58(6):846-58.
37. Gilmore IT, Burroughs A, Murray-Lyon IM, Williams R, Jenkins D, Hopkins A. 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. 1995;36(3):437-41.
38. McGill DB, Rakela J, Zinsmeister AR, Ott BJ. A 21-year experience with major hemorrhage after percutaneous liver biopsy. Gastroenterology. 1990;99(5):1396-400.
39. Piccinino F, Sagnelli E, Pasquale G, Giusti G. Complications following percutaneous liver biopsy. A multicentre retrospective study on 68,276 biopsies. J Hepatol. 1986;2(2):165-73.
40. Eisenberg E, Konopniki M, Veitsman E, Kramskay R, Gaitini D, Baruch Y. Prevalence and characteristics of pain induced by percutaneous liver biopsy. Anesth Analg. 2003;96(5):1392-6, table of contents.

41. Regev A, Berho M, Jeffers LJ, Milikowski C, Molina EG, Pyrsopoulos NT, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. American Journal of Gastroenterology. 2002;97(10):2614-8.
42. Bedossa P, Dargere D, Paradis V**.** Sampling variability of liver fibrosis in chronic hepatitis C. Hepatology. 2003;38(6):1449-57.
43. Colloredo G, Guido M, Sonzogni A, Leandro G**.** Impact of liver biopsy size on histological evaluation of chronic viral hepatitis: the smaller the sample, the milder the disease. Journal of Hepatology. 2003;39(2):239-44.
44. Cholongitas E, Senzolo M, Standish R, Marelli L, Quaglia A, Patch D, et al. A systematic review of the quality of liver biopsy specimens. American Journal of Clinical Pathology. 2006;125(5):710-21.
45. Westin J, Lagging LM, Wejstal R, Norkrans G, Dhillon AP**.** Interobserver study of liver histopathology using the Ishak score in patients with chronic hepatitis C virus infection. Liver. 1999;19(3):183-7.
46. Robert M, Sofair AN, Thomas A, Bell B, Bialek S, Corless C, et al. A comparison of hepatopathologists' and community pathologists' review of liver biopsy specimens from patients with hepatitis C. Clin Gastroenterol Hepatol. 2009;7(3):335-8.
47. Sandrin L, Fourquet B, Hasquenoph JM, Yon S, Fournier C, Mal F, et al. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. Ultrasound in Medicine & Biology. 2003;29(12):1705-13.
48. Askeland DR, Phulé PP**.** The science and engineering of materials. 5th ed: Cengage Learning; 2006.
49. Berzigotti A, De Gottardi A, Vukotic R, Siramolpiwat S, Abraldes JG, Garcia-Pagan JC, et al. Effect of meal ingestion on liver stiffness in patients with cirrhosis and portal hypertension. PLoS ONE [Electronic Resource]. 2013;8(3):e58742.
50. Mederacke I, Wursthorn K, Kirschner J, Rifai K, Manns MP, Wedemeyer H, et al. Food intake increases liver stiffness in patients with chronic or resolved hepatitis C virus infection. Liver International. 2009;29(10):1500-6.
51. Arena U, Lupsor Platon M, Stasi C, Moscarella S, Assarat A, Bedogni G, et al. Liver stiffness is influenced by a standardized meal in patients with chronic hepatitis C virus at different stages of fibrotic evolution. Hepatology. 2013;58(1):65-72.
52. Echosens**.** Fibroscan Proven Sources. 2007.
53. Fraquelli M, Rigamonti C, Casazza G, Conte D, Donato MF, Ronchi G, et al. Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. Gut. 2007;56(7):968-73.
54. Lucidarme D, Foucher J, Le Bail B, Vergniol J, Castera L, Duburque C, et al. Factors of accuracy of transient elastography (fibroscan) for the diagnosis of liver fibrosis in chronic hepatitis C. Hepatology. 2009;49(4):1083-9.
55. Boursier J, Zarski JP, de Ledinghen V, Rousselet MC, Sturm N, Lebail B, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. Hepatology. 2013;57(3):1182-91.
56. Myers RP, Crotty P, Pomier-Layrargues G, Ma M, Urbanski SJ, Elkashab M**.** Prevalence, risk factors and causes of discordance in fibrosis staging by transient elastography and liver biopsy. Liver International. 2010;30(10):1471-80.
57. Echosens**.** Criteria for a reliable stiffness examination. 2013.
58. Ascha MS, Hanouneh IA, Lopez R, Tamimi TA, Feldstein AF, Zein NN**.** The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. Hepatology. 2010;51(6):1972-8.
59. Bhala N, Angulo P, van der Poorten D, Lee E, Hui JM, Saracco G, et al. The natural history of nonalcoholic fatty liver disease with advanced fibrosis or cirrhosis: an international collaborative study. Hepatology. 2011;54(4):1208-16.

60. Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR**.** Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. Am J Gastroenterol. 1999;94(9):2467-74.
61. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology. 2005;41(6):1313-21.
62. Brunt EM, Kleiner DE, Wilson LA, Belt P, Neuschwander-Tetri BA, Network NCR**.** Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. Hepatology. 2011;53(3):810-20.
63. Bedossa P, Poitou C, Veyrie N, Bouillot JL, Basdevant A, Paradis V, et al. Histopathological algorithm and scoring system for evaluation of liver lesions in morbidly obese patients. Hepatology. 2012;56(5):1751-9.
64. Bedossa P, Consortium FP**.** Utility and appropriateness of the fatty liver inhibition of progression (FLIP) algorithm and steatosis, activity, and fibrosis (SAF) score in the evaluation of biopsies of nonalcoholic fatty liver disease. Hepatology. 2014;60(2):565-75.
65. Loomba R, Chalasani N**.** The Hierarchical Model of NAFLD: Prognostic Significance of Histologic Features in NASH. Gastroenterology. 2015;149(2):278-81.
66. EchoSens**.** Publications and Communications for Fibroscan (R); 2007.
67. Vergara S, Macias J, Rivero A, Gutierrez-Valencia A, Gonzalez-Serrano M, Merino D, et al. The use of transient elastometry for assessing liver fibrosis in patients with HIV and hepatitis C virus coinfection. Clinical Infectious Diseases. 2007;45(8):969-74.
68. Vizzutti F, Arena U, Romanelli RG, Rega L, Foschi M, Colagrande S, et al. Liver stiffness measurement predicts severe portal hypertension in patients with HCV-related cirrhosis. Hepatology. 2007;45(5):1290-7.
69. Posthouwer D, Mauser-Bunschoten EP, Fischer K, KJ VANE, RJ DEK**.** Significant liver damage in patients with bleeding disorders and chronic hepatitis C: non-invasive assessment of liver fibrosis using transient elastography. Journal of Thrombosis & Haemostasis. 2007;5(1):25-30.
70. de Ledinghen V, Trimoulet P, Mannant PR, Dumas F, Champbenoit P, Baldit C, et al. Outbreak of hepatitis C virus infection during sclerotherapy of varicose veins: long-term follow-up of 196 patients (4535 patient-years). Journal of Hepatology. 2007;46(1):19-25.
71. Barreiro P, Labarga P, Martin-Carbonero L, Amor A, Ruiz-Sancho A, Castellares C, et al. Sustained virological response following HCV therapy is associated with non-progression of liver fibrosis in HCV/HIV-coinfected patients. Antiviral Therapy. 2006;11(7):869-77.
72. de Ledinghen V, Douvin C, Kettaneh A, Ziol M, Roulot D, Marcellin P, et al. Diagnosis of hepatic fibrosis and cirrhosis by transient elastography in HIV/hepatitis C virus-coinfected patients. Journal of Acquired Immune Deficiency Syndromes: JAIDS. 2006;41(2):175-9.
73. Takeda T, Yasuda T, Nakayama Y, Nakaya M, Kimura M, Yamashita M, et al. Usefulness of noninvasive transient elastography for assessment of liver fibrosis stage in chronic hepatitis C. World Journal of Gastroenterology. 2006;12(48):7768-73.
74. Nahon P, Thabut G, Ziol M, Htar MT, Cesaro F, Barget N, et al. Liver stiffness measurement versus clinicians' prediction or both for the assessment of liver fibrosis in patients with chronic hepatitis C. American Journal of Gastroenterology. 2006;101(12):2744-51.
75. Barreiro P, Martin-Carbonero L, Nunez M, Rivas P, Morente A, Simarro N, et al. Predictors of liver fibrosis in HIV-infected patients with chronic hepatitis C virus (HCV) infection: assessment using transient elastometry and the role of HCV genotype 3. Clinical Infectious Diseases. 2006;42(7):1032-9.
76. Colletta C, Smirne C, Fabris C, Toniutto P, Rapetti R, Minisini R, et al. Value of two noninvasive methods to detect progression of fibrosis among HCV carriers with normal aminotransferases. Hepatology. 2005;42(4):838-45.

77. Ziol M, Handra-Luca A, Kettaneh A, Christidis C, Mal F, Kazemi F, et al. Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with chronic hepatitis C. Hepatology. 2005;41(1):48-54.
78. Castéra L, Vergniol J, Foucher J, Le Bail B, Chanteloup E, Haaser M, et al. Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. Gastroenterology. 2005;128(2):343-50.
79. Carrion JA, Navasa M, Bosch J, Bruguera M, Gilabert R, Forns X**.** Transient elastography for diagnosis of advanced fibrosis and portal hypertension in patients with hepatitis C recurrence after liver transplantation. Liver Transplantation. 2006;12(12):1791-8.
80. Kim KM, Choi WB, Park SH, Yu E, Lee SG, Lim YS, et al. Diagnosis of hepatic steatosis and fibrosis by transient elastography in asymptomatic healthy individuals: a prospective study of living related potential liver donors. Journal of Gastroenterology. 2007;42(5):382-8.
81. Coco B, Oliveri F, Maina AM, Ciccorossi P, Sacco R, Colombatto P, et al. Transient elastography: a new surrogate marker of liver fibrosis influenced by major changes of transaminases. Journal of Viral Hepatitis. 2007;14(5):360-9.
82. Kawamoto M, Mizuguchi T, Katsuramaki T, Nagayama M, Oshima H, Kawasaki H, et al. Assessment of liver fibrosis by a noninvasive method of transient elastography and biochemical markers. World Journal of Gastroenterology. 2006;12(27):4325-30.
83. Kazemi F, Kettaneh A, N'Kontchou G, Pinto E, Ganne-Carrie N, Trinchet JC, et al. Liver stiffness measurement selects patients with cirrhosis at risk of bearing large oesophageal varices. Journal of Hepatology. 2006;45(2):230-5.
84. Foucher J, Chanteloup E, Vergniol J, Castera L, Le Bail B, Adhoute X, et al. Diagnosis of cirrhosis by transient elastography (FibroScan): a prospective study. Gut. 2006;55(3):403-8.
85. Kettaneh A, Marcellin P, Douvin C, Poupon R, Ziol M, Beaugrand M, 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. Journal of Hepatology. 2007;46(4):628-34.
86. Foucher J, Castera L, Bernard PH, Adhoute X, Laharie D, Bertet J, et al. Prevalence and factors associated with failure of liver stiffness measurement using FibroScan in a prospective study of 2114 examinations. European Journal of Gastroenterology & Hepatology. 2006;18(4):411-2.
87. Marcellin P, Ziol M, Bedossa P, Douvin C, Poupon R, de Ledinghen V, et al. Non-invasive assessment of liver fibrosis by stiffness measurement in patients with chronic hepatitis B. Liver Int. 2009;29(2):242-7.
88. Chan HL, Wong GL, Choi PC, Chan AW, Chim AM, Yiu KK, et al. Alanine aminotransferase-based algorithms of liver stiffness measurement by transient elastography (Fibroscan) for liver fibrosis in chronic hepatitis B. Journal of Viral Hepatitis. 2009;16(1):36-44.
89. Berends MA, Snoek J, de Jong EM, Van Krieken JH, de Knegt RJ, van Oijen MG, et al. Biochemical and biophysical assessment of MTX-induced liver fibrosis in psoriasis patients: Fibrotest predicts the presence and Fibroscan predicts the absence of significant liver fibrosis. Liver Int. 2007;27(5):639-45.
90. Laharie D, Zerbib F, Adhoute X, Boue-Lahorgue X, Foucher J, Castera L, et al. Diagnosis of liver fibrosis by transient elastography (FibroScan) and non-invasive methods in Crohn's disease patients treated with methotrexate. Alimentary Pharmacology & Therapeutics. 2006;23(11):1621-8.
91. Maida I, Soriano V, Castellares C, Ramos B, Sotgiu G, Martin-Carbonero L, et al. Liver fibrosis in HIV-infected patients with chronic hepatitis B extensively exposed to antiretroviral therapy with anti-HBV activity. HIV Clinical Trials. 2006;7(5):246-50.
92. Corpechot C, El Naggar A, Poujol-Robert A, Ziol M, Wendum D, Chazouilleres O, et al. Assessment of biliary fibrosis by transient elastography in patients with PBC and PSC. Hepatology. 2006;43(5):1118-24.
93. <Lavanchy-2004-Journal_of_Viral_Hepatitis.pdf>.

94. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. 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.[Erratum appears in Lancet. 2013 Feb 23;381(9867):628 Note: AlMazroa, Mohammad A [added]; Memish, Ziad A [added]]. Lancet. 2012;380(9859):2095-128.
95. Williams CD, Stengel J, Asike MI, Torres DM, Shaw J, Contreras M, et al. Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. Gastroenterology. 2011;140(1):124-31.
96. Wong VW-S, Chu WC-W, Wong GL-H, Chan RS-M, Chim AM-L, Ong A, et al. Prevalence of non-alcoholic fatty liver disease and advanced fibrosis in Hong Kong Chinese: a population study using proton-magnetic resonance spectroscopy and transient elastography. Gut. 2012;61(3):409-15.
97. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. Hepatology. 2012;55(6):2005-23.
98. Neuschwander-Tetri BA, Caldwell SH**.** Nonalcoholic steatohepatitis: summary of an AASLD Single Topic Conference. Hepatology. 2003;37(5):1202-19.
99. Brunt EM**.** Histopathology of nonalcoholic fatty liver disease. World Journal of Gastroenterology. 2010;16(42):5286.
100. Brunt EM**.** Nonalcoholic steatohepatitis: definition and pathology. Seminars in Liver Disease. 2001;21(1):3-16.
101. Brunt EM**.** Pathology of fatty liver disease. Modern Pathology. 2007;20 Suppl 1:S40-8.
102. Adams LA, Sanderson S, Lindor KD, Angulo P**.** The histological course of nonalcoholic fatty liver disease: a longitudinal study of 103 patients with sequential liver biopsies. J Hepatol. 2005;42(1):132-8.
103. Ong JP, Pitts A, Younossi ZM**.** Increased overall mortality and liver-related mortality in non-alcoholic fatty liver disease. J Hepatol. 2008;49(4):608-12.
104. Ekstedt M, Franzen LE, Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. Hepatology. 2006;44(4):865-73.
105. Matteoni CA, Younossi ZM, Gramlich T, Boparai N, Liu YC, McCullough AJ**.** Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. Gastroenterology. 1999;116(6):1413-9.
106. Rafiq N, Bai C, Fang Y, Srishord M, McCullough A, Gramlich T, et al. Long-term follow-up of patients with nonalcoholic fatty liver. Clin Gastroenterol Hepatol. 2009;7(2):234-8.
107. Soderberg C, Stal P, Askling J, Glaumann H, Lindberg G, Marmur J, et al. Decreased survival of subjects with elevated liver function tests during a 28-year follow-up. Hepatology. 2010;51(2):595-602.
108. Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P, et al. Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology. 2015;149(2):389-97 e10.
109. Kwok R, Tse YK, Wong GL, Ha Y, Lee AU, Ngu MC, et al. Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease - the role of transient elastography and plasma cytokeratin-18 fragments. Aliment Pharmacol Ther. 2014;39(3):254-69.
110. Leite NC, Salles GF, Araujo AL, Villela-Nogueira CA, Cardoso CR**.** Prevalence and associated factors of non-alcoholic fatty liver disease in patients with type-2 diabetes mellitus. Liver International. 2009;29(1):113-9.
111. Kelley DE, McKolanis TM, Hegazi RA, Kuller LH, Kalhan SC**.** Fatty liver in type 2 diabetes mellitus: relation to regional adiposity, fatty acids, and insulin resistance. Am J Physiol Endocrinol Metab. 2003;285(4):E906-16.

112. Targher G, Bertolini L, Padovani R, Rodella S, Tessari R, Zenari L, et al. Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. Diabetes Care. 2007;30(5):1212-8.
113. MacLachlan JH, Allard N, Towell V, Cowie BC**.** The burden of chronic hepatitis B virus infection in Australia, 2011. Aust N Z J Public Health. 2013;37(5):416-22.
114. Sievert W, Razavi H, Estes C, Thompson AJ, Zekry A, Roberts SK, et al. Enhanced antiviral treatment efficacy and uptake in preventing the rising burden of hepatitis C-related liver disease and costs in Australia. J Gastroenterol Hepatol. 2014;29 Suppl 1:1-9.
115. Edwards CQ, Griffen LM, Goldgar D, Drummond C, Skolnick MH, Kushner JP**.** Prevalence of hemochromatosis among 11,065 presumably healthy blood donors. N Engl J Med. 1988;318(21):1355-62.
116. Dowman JK, Tomlinson JW, Newsome PN**.** Pathogenesis of non-alcoholic fatty liver disease. QJM. 2010;103(2):71-83.
117. Petta S, Di Marco V, Camma C, Butera G, Cabibi D, Craxi A**.** Reliability of liver stiffness measurement in non-alcoholic fatty liver disease: the effects of body mass index. Aliment Pharmacol Ther. 2011;33(12):1350-60.
118. Kumar R, Rastogi A, Sharma MK, Bhatia V, Tyagi P, Sharma P, et al. Liver stiffness measurements in patients with different stages of nonalcoholic fatty liver disease: diagnostic performance and clinicopathological correlation. Dig Dis Sci. 2013;58(1):265-74.
119. Machado M, Marques-Vidal P, Cortez-Pinto H**.** Hepatic histology in obese patients undergoing bariatric surgery. J Hepatol. 2006;45(4):600-6.
120. Colicchio P, Tarantino G, del Genio F, Sorrentino P, Saldalamacchia G, Finelli C, et al. Non-alcoholic fatty liver disease in young adult severely obese non-diabetic patients in South Italy. Annals of Nutrition & Metabolism. 2005;49(5):289-95.
121. Beymer C, Kowdley KV, Larson A, Edmonson P, Dellinger EP, Flum DR**.** Prevalence and predictors of asymptomatic liver disease in patients undergoing gastric bypass surgery. Archives of Surgery. 2003;138(11):1240-4.
122. Hossain N, Afendy A, Stepanova M, Nader F, Srishord M, Rafiq N, et al. Independent predictors of fibrosis in patients with nonalcoholic fatty liver disease. Clinical Gastroenterology & Hepatology. 2009;7(11):1224-9, 9.e1-2.
123. Ratziu V, Giral P, Charlotte F, Bruckert E, Thibault V, Theodorou I, et al. Liver fibrosis in overweight patients. Gastroenterology. 2000;118(6):1117-23.
124. Gaia S, Carenzi S, Barilli AL, Bugianesi E, Smedile A, Brunello F, et al. Reliability of transient elastography for the detection of fibrosis in non-alcoholic fatty liver disease and chronic viral hepatitis. J Hepatol. 2011;54(1):64-71.
125. Lupsor M, Badea R, Stefanescu H, Grigorescu M, Serban A, Radu C, et al. Performance of unidimensional transient elastography in staging non-alcoholic steatohepatitis. Journal of Gastrointestinal & Liver Diseases. 2010;19(1):53-60.
126. Myers RP, Elkashab M, Ma M, Crotty P, Pomier-Layrargues G**.** Transient elastography for the noninvasive assessment of liver fibrosis: a multicentre Canadian study. Canadian Journal of Gastroenterology. 2010;24(11):661-70.
127. Wong VW-S, Vergniol J, Wong GL-H, Foucher J, Chan HL-Y, Le Bail B, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. Hepatology. 2010;51(2):454-62.
128. Yoneda M, Suzuki K, Kato S, Fujita K, Nozaki Y, Hosono K, et al. Nonalcoholic fatty liver disease: US-based acoustic radiation force impulse elastography. Radiology. 2010;256(2):640-7.
129. Wong VW-S, Vergniol J, Wong GL-H, Foucher J, Chan AW-H, Chermak F, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. American Journal of Gastroenterology. 2012;107(12):1862-71.

130. Petta S, Amato MC, Di Marco V, Camma C, Pizzolanti G, Barcellona MR, et al. Visceral adiposity index is associated with significant fibrosis in patients with non-alcoholic fatty liver disease. Alimentary Pharmacology & Therapeutics. 2012;35(2):238-47.
131. Echosens. Fibroscan (R) 502 Touch. 2013.
132. Kwok R, Choi KC, Wong GL, Zhang Y, Chan HL, Luk AO, et al. Screening diabetic patients for non-alcoholic fatty liver disease with controlled attenuation parameter and liver stiffness measurements: a prospective cohort study. Gut. 2016;65(8):1359-68.
133. Wong VW, Chu WC, Wong GL, Chan RS, Chim AM, Ong A, et al. Prevalence of non-alcoholic fatty liver disease and advanced fibrosis in Hong Kong Chinese: a population study using proton-magnetic resonance spectroscopy and transient elastography. Gut. 2012;61(3):409-15.
134. Wieckowska A, Zein NN, Yerian LM, Lopez AR, McCullough AJ, Feldstein AE. In vivo assessment of liver cell apoptosis as a novel biomarker of disease severity in nonalcoholic fatty liver disease. Hepatology. 2006;44(1):27-33.
135. Yilmaz Y, Dolar E, Ulukaya E, Akgoz S, Keskin M, Kiyici M, et al. Soluble forms of extracellular cytokeratin 18 may differentiate simple steatosis from nonalcoholic steatohepatitis. World Journal of Gastroenterology. 2007;13(6):837-44.
136. Diab DL, Yerian L, Schauer P, Kashyap SR, Lopez R, Hazen SL, et al. Cytokeratin 18 fragment levels as a noninvasive biomarker for nonalcoholic steatohepatitis in bariatric surgery patients. Clinical Gastroenterology & Hepatology. 2008;6(11):1249-54.
137. Younossi ZM, Jarrar M, Nugent C, Randhawa M, Afendy M, Stepanova M, et al. A novel diagnostic biomarker panel for obesity-related nonalcoholic steatohepatitis (NASH). Obesity Surgery. 2008;18(11):1430-7.
138. Feldstein AE, Wieckowska A, Lopez AR, Liu YC, Zein NN, McCullough AJ. Cytokeratin-18 fragment levels as noninvasive biomarkers for nonalcoholic steatohepatitis: a multicenter validation study. Hepatology. 2009;50(4):1072-8.
139. Papatheodoridis GV, Hadziyannis E, Tsochatzis E, Georgiou A, Kafiri G, Tiniakos DG, et al. Serum apoptotic caspase activity in chronic hepatitis C and nonalcoholic Fatty liver disease. Journal of Clinical Gastroenterology. 2010;44(4):e87-95.
140. Musso G, Cassader M, De Michieli F, Saba F, Bo S, Gambino R. Effect of lectin-like oxidized LDL receptor-1 polymorphism on liver disease, glucose homeostasis, and postprandial lipoprotein metabolism in nonalcoholic steatohepatitis. American Journal of Clinical Nutrition. 2011;94(4):1033-42.
141. Joka D, Wahl K, Moeller S, Schlue J, Vaske B, Bahr MJ, et al. Prospective biopsy-controlled evaluation of cell death biomarkers for prediction of liver fibrosis and nonalcoholic steatohepatitis. Hepatology. 2012;55(2):455-64.
142. Papatheodoridis M, Crisan D, Radu C, Grigorescu MD, Sparchez Z, Serban A. A novel pathophysiological-based panel of biomarkers for the diagnosis of nonalcoholic steatohepatitis. Journal of Physiology & Pharmacology. 2012;63(4):347-53.
143. Pirvulescu I, Gheorghe L, Csiki I, Becheanu G, Dumbrava M, Fica S, et al. Noninvasive clinical model for the diagnosis of nonalcoholic steatohepatitis in overweight and morbidly obese patients undergoing bariatric surgery. Chirurgia (Bucuresti). 2012;107(6):772-9.
144. Shen J, Chan HL, Wong GL, Choi PC, Chan AW, Chan HY, et al. Non-invasive diagnosis of non-alcoholic steatohepatitis by combined serum biomarkers. Journal of Hepatology. 2012;56(6):1363-70.
145. Palmeri ML, Wang MH, Rouze NC, Abdelmalek MF, Guy CD, Moser B, et al. Noninvasive evaluation of hepatic fibrosis using acoustic radiation force-based shear stiffness in patients with nonalcoholic fatty liver disease. J Hepatol. 2011;55(3):666-72.
146. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003;3:25.

147. Reitsma JB GA, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH**.** Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58:982-90.
148. Rutter CM GC**.** A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med. 2001;20:2865-884.
149. Metandi HR**.** Stata module for meta-analysis of diagnostic accuracy. Statistical Software Components. Boston College, Department of Economics; 2008.
150. Review Manager (RevMan). The Cochrane Collaboration. 5.0 ed. Copenhagen, The Nordic Cochrane Centre; 2008.
151. Younossi ZM, Stepanova M, Rafiq N, Makhlouf H, Younoszai Z, Agrawal R, et al. Pathologic criteria for nonalcoholic steatohepatitis: interprotocol agreement and ability to predict liver-related mortality. Hepatology. 2011;53(6):1874-82.
152. Pais R, Charlotte F, Fedchuk L, Bedossa P, Lebray P, Poynard T, et al. A systematic review of follow-up biopsies reveals disease progression in patients with non-alcoholic fatty liver. J Hepatol. 2013;59(3):550-6.
153. Bantel H, Ruck P, Gregor M, Schulze-Osthoff K**.** Detection of elevated caspase activation and early apoptosis in liver diseases. European Journal of Cell Biology. 2001;80(3):230-9.
154. Tamimi TI, Elgouhari HM, Alkhouri N, Yerian LM, Berk MP, Lopez R, et al. An apoptosis panel for nonalcoholic steatohepatitis diagnosis. Journal of Hepatology. 2011;54(6):1224-9.
155. Abiru S, Migita K, Maeda Y, Daikoku M, Ito M, Ohata K, et al. Serum cytokine and soluble cytokine receptor levels in patients with non-alcoholic steatohepatitis. Liver Int. 2006;26(1):39-45.
156. Copaci I, Micu L, Voiculescu M**.** The role of cytokines in non-alcoholic steatohepatitis. A review. J Gastrointestin Liver Dis. 2006;15(4):363-73.
157. Wigg AJ, Roberts-Thomson IC, Dymock RB, McCarthy PJ, Grose RH, Cummins AG**.** The role of small intestinal bacterial overgrowth, intestinal permeability, endotoxaemia, and tumour necrosis factor alpha in the pathogenesis of non-alcoholic steatohepatitis. Gut. 2001;48(2):206-11.
158. Wong VW, Hui AY, Tsang SW, Chan JL, Tse AM, Chan KF, et al. Metabolic and adipokine profile of Chinese patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol. 2006;4(9):1154-61.
159. Wieckowska A, Papouchado BG, Li Z, Lopez R, Zein NN, Feldstein AE**.** Increased hepatic and circulating interleukin-6 levels in human nonalcoholic steatohepatitis. Am J Gastroenterol. 2008;103(6):1372-9.
160. Grigorescu M, Crisan D, Radu C, Grigorescu MD, Sparchez Z, Serban A**.** A novel pathophysiological-based panel of biomarkers for the diagnosis of nonalcoholic steatohepatitis. Journal of Physiology & Pharmacology. 2012;63(4):347-53.
161. Jarrar MH, Baranova A, Collantes R, Ranard B, Stepanova M, Bennett C, et al. Adipokines and cytokines in non-alcoholic fatty liver disease. Aliment Pharmacol Ther. 2008;27(5):412-21.
162. Haukeland JW, Damas JK, Konopski Z, Loberg EM, Haaland T, Goverud I, et al. Systemic inflammation in nonalcoholic fatty liver disease is characterized by elevated levels of CCL2. J Hepatol. 2006;44(6):1167-74.
163. Chitturi S, Abeygunasekera S, Farrell GC, Holmes-Walker J, Hui JM, Fung C, et al. NASH and insulin resistance: Insulin hypersecretion and specific association with the insulin resistance syndrome. Hepatology. 2002;35(2):373-9.
164. Boza C, Riquelme A, Ibanez L, Duarte I, Norero E, Viviani P, et al. Predictors of nonalcoholic steatohepatitis (NASH) in obese patients undergoing gastric bypass. Obes Surg. 2005;15(8):1148-53.
165. Yoneda M, Mawatari H, Fujita K, Iida H, Yonemitsu K, Kato S, et al. High-sensitivity C-reactive protein is an independent clinical feature of nonalcoholic steatohepatitis (NASH) and also of the severity of fibrosis in NASH. J Gastroenterol. 2007;42(7):573-82.

166. Hui JM, Farrell GC, Kench JG, George J**.** High sensitivity C-reactive protein values do not reliably predict the severity of histological changes in NAFLD. Hepatology. 2004;39(5):1458-9.
167. Miura K, Yang L, van Rooijen N, Ohnishi H, Seki E**.** Hepatic recruitment of macrophages promotes nonalcoholic steatohepatitis through CCR2. Am J Physiol Gastrointest Liver Physiol. 2012;302(11):G1310-21.
168. Kirovski G, Dorn C, Huber H, Moleda L, Niessen C, Wobser H, et al. Elevated systemic monocyte chemoattractrant protein-1 in hepatic steatosis without significant hepatic inflammation. Exp Mol Pathol. 2011;91(3):780-3.
169. Yoneda M, Uchiyama T, Kato S, Endo H, Fujita K, Yoneda K, et al. Plasma Pentraxin3 is a novel marker for nonalcoholic steatohepatitis (NASH). BMC Gastroenterol. 2008;8:53.
170. Tarcin O, Gedik N, Karakoyun B, Tahan V, Sood G, Celikel C, et al. Serum prolidase and IGF-1 as non-invasive markers of hepatic fibrosis during four different periods after bile-duct ligation in rats. Dig Dis Sci. 2008;53(7):1938-45.
171. Kayadibi H, Gultepe M, Yasar B, Ince AT, Ozcan O, Ipcioglu OM, et al. Diagnostic value of serum prolidase enzyme activity to predict the liver histological lesions in non-alcoholic fatty liver disease: a surrogate marker to distinguish steatohepatitis from simple steatosis. Dig Dis Sci. 2009;54(8):1764-71.
172. Basta G, Sironi AM, Lazzerini G, Del Turco S, Buzzigoli E, Casolaro A, et al. Circulating soluble receptor for advanced glycation end products is inversely associated with glycemic control and S100A12 protein. J Clin Endocrinol Metab. 2006;91(11):4628-34.
173. Geroldi D, Falcone C, Emanuele E, D'Angelo A, Calcagnino M, Buzzi MP, et al. Decreased plasma levels of soluble receptor for advanced glycation end-products in patients with essential hypertension. J Hypertens. 2005;23(9):1725-9.
174. Yilmaz Y, Ulukaya E, Gul OO, Arabul M, Gul CB, Atug O, et al. Decreased plasma levels of soluble receptor for advanced glycation endproducts (sRAGE) in patients with nonalcoholic fatty liver disease. Clin Biochem. 2009;42(9):802-7.
175. Geroldi D, Falcone C, Emanuele E**.** Soluble receptor for advanced glycation end products: from disease marker to potential therapeutic target. Curr Med Chem. 2006;13(17):1971-8.
176. Uysal S, Armutcu F, Aydogan T, Akin K, Ikizek M, Yigitoglu MR**.** Some inflammatory cytokine levels, iron metabolism and oxidan stress markers in subjects with nonalcoholic steatohepatitis. Clin Biochem. 2011;44(17-18):1375-9.
177. Rensen SS, Slaats Y, Nijhuis J, Jans A, Bieghs V, Driessen A, et al. Increased hepatic myeloperoxidase activity in obese subjects with nonalcoholic steatohepatitis. Am J Pathol. 2009;175(4):1473-82.
178. Horoz M, Bolukbas C, Bolukbas FF, Sabuncu T, Aslan M, Sarifakiogullari S, et al. Measurement of the total antioxidant response using a novel automated method in subjects with nonalcoholic steatohepatitis. BMC Gastroenterol. 2005;5:35.
179. Bonnefont-Rousselot D, Ratziu V, Giral P, Charlotte F, Beucler I, Poynard T**.** Blood oxidative stress markers are unreliable markers of hepatic steatosis. Aliment Pharmacol Ther. 2006;23(1):91-8.
180. Lomonaco R, Sunny NE, Bril F, Cusi K**.** Nonalcoholic fatty liver disease: current issues and novel treatment approaches. Drugs. 2013;73(1):1-14.
181. Cheung O, Kapoor A, Puri P, Sistrun S, Luketic VA, Sargeant CC, et al. The impact of fat distribution on the severity of nonalcoholic fatty liver disease and metabolic syndrome. Hepatology. 2007;46(4):1091-100.
182. Nomura K, Yano E, Shinozaki T, Tagawa K**.** Efficacy and effectiveness of liver screening program to detect fatty liver in the periodic health check-ups. J Occup Health. 2004;46(6):423-8.
183. Poynard T, Ratziu V, Charlotte F, Messous D, Munteanu M, Imbert-Bismut F, et al. Diagnostic value of biochemical markers (NashTest) for the prediction of non alcoholo steato hepatitis in patients with non-alcoholic fatty liver disease. BMC Gastroenterol. 2006;6:34.

184. Lassailly G, Caiazzo R, Hollebecque A, Buob D, Leteurtre E, Arnalsteen L, et al. Validation of noninvasive biomarkers (FibroTest, SteatoTest, and NashTest) for prediction of liver injury in patients with morbid obesity. Eur J Gastroenterol Hepatol. 2011;23(6):499-506.
185. Younossi ZM, Page S, Rafiq N, Birerdinc A, Stepanova M, Hossain N, et al. A biomarker panel for non-alcoholic steatohepatitis (NASH) and NASH-related fibrosis. Obes Surg. 2011;21(4):431-9.
186. Anty R, Iannelli A, Patouraux S, Bonnafous S, Lavallard VJ, Senni-Buratti M, et al. A new composite model including metabolic syndrome, alanine aminotransferase and cytokeratin-18 for the diagnosis of non-alcoholic steatohepatitis in morbidly obese patients. Alimentary Pharmacology & Therapeutics. 2010;32(11-12):1315-22.
187. Feldstein AE, Lopez R, Tamimi TA, Yerian L, Chung YM, Berk M, et al. Mass spectrometric profiling of oxidized lipid products in human nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. J Lipid Res. 2010;51(10):3046-54.
188. Roskams T, Yang SQ, Koteish A, Durnez A, DeVos R, Huang X, et al. Oxidative stress and oval cell accumulation in mice and humans with alcoholic and nonalcoholic fatty liver disease. Am J Pathol. 2003;163(4):1301-11.
189. Dixon JB, Bhathal PS, O'Brien PE**.** Nonalcoholic fatty liver disease: predictors of nonalcoholic steatohepatitis and liver fibrosis in the severely obese. Gastroenterology. 2001;121(1):91-100.
190. Paradis V, Zalinski S, Chelbi E, Guedj N, Degos F, Vilgrain V, et al. Hepatocellular carcinomas in patients with metabolic syndrome often develop without significant liver fibrosis: a pathological analysis. Hepatology. 2009;49(3):851-9.
191. Yasui K, Hashimoto E, Komorizono Y, Koike K, Arii S, Imai Y, et al. Characteristics of patients with nonalcoholic steatohepatitis who develop hepatocellular carcinoma. Clin Gastroenterol Hepatol. 2011;9(5):428-33; quiz e50.
192. Wong VW, Wong GL, Tsang SW, Hui AY, Chan AW, Choi PC, et al. Metabolic and histological features of non-alcoholic fatty liver disease patients with different serum alanine aminotransferase levels. Aliment Pharmacol Ther. 2009;29(4):387-96.
193. Hui AY, Wong VW, Chan HL, Liew CT, Chan JL, Chan FK, et al. Histological progression of non-alcoholic fatty liver disease in Chinese patients. Aliment Pharmacol Ther. 2005;21(4):407-13.
194. Poonawala A, Nair SP, Thuluvath PJ**.** Prevalence of obesity and diabetes in patients with cryptogenic cirrhosis: a case-control study. Hepatology. 2000;32(4 Pt 1):689-92.
195. Struben VM, Hespenheide EE, Caldwell SH**.** Nonalcoholic steatohepatitis and cryptogenic cirrhosis within kindreds. Am J Med. 2000;108(1):9-13.
196. Angulo P, Hui JM, Marchesini G, Bugianesi E, George J, Farrell GC, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. Hepatology. 2007;45(4):846-54.
197. Wong VW, Wong GL, Chim AM, Tse AM, Tsang SW, Hui AY, et al. Validation of the NAFLD fibrosis score in a Chinese population with low prevalence of advanced fibrosis. Am J Gastroenterol. 2008;103(7):1682-8.
198. McPherson S, Stewart SF, Henderson E, Burt AD, Day CP**.** Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. Gut. 2010;59(9):1265-9.
199. Shah AG, Lydecker A, Murray K, Tetri BN, Contos MJ, Sanyal AJ, et al. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol. 2009;7(10):1104-12.
200. Mehta SH, Lau B, Afdhal NH, Thomas DL**.** Exceeding the limits of liver histology markers. J Hepatol. 2009;50(1):36-41.
201. Harrison SA, Oliver D, Arnold HL, Gogia S, Neuschwander-Tetri BA**.** Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. Gut. 2008;57(10):1441-7.
202. Cales P, Laine F, Boursier J, Deugnier Y, Moal V, Oberti F, et al. Comparison of blood tests for liver fibrosis specific or not to NAFLD. J Hepatol. 2009;50(1):165-73.

203. Sheth SG, Flamm SL, Gordon FD, Chopra S. AST/ALT ratio predicts cirrhosis in patients with chronic hepatitis C virus infection. Am J Gastroenterol. 1998;93(1):44-8.
204. Wai C-T, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology. 2003;38(2):518-26.
205. Guha IN, Parkes J, Roderick P, Chattopadhyay D, Cross R, Harris S, et al. Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: Validating the European Liver Fibrosis Panel and exploring simple markers. Hepatology. 2008;47(2):455-60.
206. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology. 2006;43(6):1317-25.
207. Ratziu V, Massard J, Charlotte F, Messous D, Imbert-Bismut F, Bonyhay L, et al. Diagnostic value of biochemical markers (FibroTest-FibroSURE) for the prediction of liver fibrosis in patients with non-alcoholic fatty liver disease. BMC Gastroenterol. 2006;6:6.
208. Adams LA, George J, Bugianesi E, Rossi E, De Boer WB, van der Poorten D, et al. Complex non-invasive fibrosis models are more accurate than simple models in non-alcoholic fatty liver disease. J Gastroenterol Hepatol. 2011;26(10):1536-43.
209. Hernaez R, Lazo M, Bonekamp S, Kamel I, Brancati FL, Guallar E, et al. Diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver: a meta-analysis. Hepatology. 2011;54(3):1082-90.
210. Saadeh S, Younossi ZM, Remer EM, Gramlich T, Ong JP, Hurley M, et al. The utility of radiological imaging in nonalcoholic fatty liver disease. Gastroenterology. 2002;123(3):745-50.
211. Osawa H, Mori Y. Sonographic diagnosis of fatty liver using a histogram technique that compares liver and renal cortical echo amplitudes. Journal of Clinical Ultrasound. 1996;24(1):25-9.
212. Ballestri S, Lonardo A, Romagnoli D, Carulli L, Losi L, Day CP, et al. Ultrasonographic fatty liver indicator, a novel score which rules out NASH and is correlated with metabolic parameters in NAFLD. Liver International. 2012;32(8):1242-52.
213. Park SH, Kim PN, Kim KW, Lee SW, Yoon SE, Park SW, et al. Macrovesicular hepatic steatosis in living liver donors: use of CT for quantitative and qualitative assessment. Radiology. 2006;239(1):105-12.
214. Rofsky NM, Fleishaker H. CT and MRI of diffuse liver disease. Seminars in Ultrasound, CT & MR. 1995;16(1):16-33.
215. Schwenzer NF, Springer F, Schraml C, Stefan N, Machann J, Schick F. Non-invasive assessment and quantification of liver steatosis by ultrasound, computed tomography and magnetic resonance. Journal of Hepatology. 2009;51(3):433-45.
216. Fishbein M, Castro F, Cheruku S, Jain S, Webb B, Gleason T, et al. Hepatic MRI for fat quantitation: its relationship to fat morphology, diagnosis, and ultrasound. J Clin Gastroenterol. 2005;39(7):619-25.
217. McPherson S, Jonsson JR, Cowin GJ, O'Rourke P, Clouston AD, Volp A, et al. Magnetic resonance imaging and spectroscopy accurately estimate the severity of steatosis provided the stage of fibrosis is considered. J Hepatol. 2009;51(2):389-97.
218. Machado MV, Cortez-Pinto H. Non-invasive diagnosis of non-alcoholic fatty liver disease. A critical appraisal. J Hepatol. 2013;58(5):1007-19.
219. Yin M, Talwalkar JA, Glaser KJ, Manduca A, Grimm RC, Rossman PJ, et al. Assessment of hepatic fibrosis with magnetic resonance elastography. Clin Gastroenterol Hepatol. 2007;5(10):1207-13 e2.
220. Huwart L, Sempoux C, Vicaut E, Salameh N, Annet L, Danse E, et al. Magnetic resonance elastography for the noninvasive staging of liver fibrosis. Gastroenterology. 2008;135(1):32-40.
221. Chen J, Talwalkar JA, Yin M, Glaser KJ, Sanderson SO, Ehman RL. Early detection of nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease by using MR elastography. Radiology. 2011;259(3):749-56.

222. Castera L, Foucher J, Bernard PH, Carvalho F, Allaix D, Merrouche W, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. Hepatology. 2010;51(3):828-35.
223. Wong GL, Wong VW, Chim AM, Yiu KK, Chu SH, Li MK, et al. Factors associated with unreliable liver stiffness measurement and its failure with transient elastography in the Chinese population. J Gastroenterol Hepatol. 2011;26(2):300-5.
224. Friedrich-Rust M, Hadji-Hosseini H, Kriener S, Herrmann E, Sircar I, Kau A, et al. Transient elastography with a new probe for obese patients for non-invasive staging of non-alcoholic steatohepatitis. European Radiology. 2010;20(10):2390-6.
225. Myers RP, Pomier-Layrargues G, Kirsch R, Pollett A, Duarte-Rojo A, Wong D, et al. Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients. Hepatology. 2012;55(1):199-208.
226. Osaki A, Kubota T, Suda T, Igarashi M, Nagasaki K, Tsuchiya A, et al. Shear wave velocity is a useful marker for managing nonalcoholic steatohepatitis. World Journal of Gastroenterology. 2010;16(23):2918-25.
227. Friedrich-Rust M, Romen D, Vermehren J, Kriener S, Sadet D, Herrmann E, et al. Acoustic radiation force impulse-imaging and transient elastography for non-invasive assessment of liver fibrosis and steatosis in NAFLD. Eur J Radiol. 2012;81(3):e325-31.
228. Friedrich-Rust M, Nierhoff J, Lupsor M, Sporea I, Fierbinteanu-Braticevici C, Strobel D, et al. Performance of Acoustic Radiation Force Impulse imaging for the staging of liver fibrosis: A pooled meta-analysis. Journal of Viral Hepatitis. 2012;19(2):e212-e9.
229. Masuda K, Ono M, Fukumoto M, Hirose A, Munekage K, Ochi T, et al. Usefulness of Technetium-99 m-2-methoxy-isobutyl-isonitrile liver scintigraphy for evaluating disease activity of non-alcoholic fatty liver disease. Hepatol Res. 2012;42(3):273-9.
230. Kikuchi M, Tomita K, Nakahara T, Kitamura N, Teratani T, Irie R, et al. Utility of quantitative 99mTc-phytate scintigraphy to diagnose early-stage non-alcoholic steatohepatitis. Scand J Gastroenterol. 2009;44(2):229-36.
231. Wong VW, Wong GL, Yeung DK, Lau TK, Chan CK, Chim AM, et al. Incidence of non-alcoholic fatty liver disease in Hong Kong: A population study with paired proton-magnetic resonance spectroscopy. J Hepatol. 2015;62(1):182-9.
232. Amarapurkar DN, Hashimoto E, Lesmana LA, Sollano JD, Chen PJ, Goh KL, et al. How common is non-alcoholic fatty liver disease in the Asia-Pacific region and are there local differences? J Gastroenterol Hepatol. 2007;22(6):788-93.
233. Ratziu V, Bellentani S, Cortez-Pinto H, Day C, Marchesini G**.** A position statement on NAFLD/NASH based on the EASL 2009 special conference. J Hepatol. 2010;53(2):372-84.
234. Wong VW, Chan HL**.** Transient elastography. J Gastroenterol Hepatol. 2010;25(11):1726-31.
235. Sasso M, Beaugrand M, de Ledinghen V, Douvin C, Marcellin P, Poupon R, et al. Controlled attenuation parameter (CAP): a novel VCTE guided ultrasonic attenuation measurement for the evaluation of hepatic steatosis: preliminary study and validation in a cohort of patients with chronic liver disease from various causes. Ultrasound Med Biol. 2010;36(11):1825-35.
236. Sasso M, Tengher-Barna I, Ziol M, Miette V, Fournier C, Sandrin L, et al. Novel controlled attenuation parameter for noninvasive assessment of steatosis using Fibroscan((R)): validation in chronic hepatitis C. J Viral Hepat. 2012;19(4):244-53.
237. de Ledinghen V, Vergniol J, Capdepont M, Chermak F, Hiriart JB, Cassinotto C, et al. Controlled attenuation parameter (CAP) for the diagnosis of steatosis: a prospective study of 5323 examinations. J Hepatol. 2014;60(5):1026-31.
238. Piwernetz K, Home PD, Snorgaard O, Antsiferov M, Staehr-Johansen K, Krans M**.** Monitoring the targets of the St Vincent Declaration and the implementation of quality management in diabetes care: the DIABCARE initiative. The DIABCARE Monitoring Group of the St Vincent Declaration Steering Committee. Diabetic Medicine. 1993;10(4):371-7.

239. Prati D, Taioli E, Zanella A, Della Torre E, Butelli S, Del Vecchio E, et al. Updated definitions of healthy ranges for serum alanine aminotransferase levels. Ann Intern Med. 2002;137(1):1-10.
240. Ma YC, Zuo L, Chen JH, Luo Q, Yu XQ, Li Y, et al. Modified glomerular filtration rate estimating equation for Chinese patients with chronic kidney disease. J Am Soc Nephrol. 2006;17(10):2937-44.
241. Wong VW, Vergniol J, Wong GL, Foucher J, Chan HL, Le Bail B, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. Hepatology. 2010;51(2):454-62.
242. Wong VW, Vergniol J, Wong GL, Foucher J, Chan AW, Chermak F, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. Am J Gastroenterol. 2012;107(12):1862-71.
243. Hosmer DW, Lemeshow S**.** Applied logistic regression. 2nd ed. New York: John Wiley and Sons; 2000.
244. Gupte P, Amarapurkar D, Agal S, Baijal R, Kulshrestha P, Pramanik S, et al. Non-alcoholic steatohepatitis in type 2 diabetes mellitus. Journal of Gastroenterology & Hepatology. 2004;19(8):854-8.
245. Fan JG, Zhu J, Li XJ, Chen L, Li L, Dai F, et al. Prevalence of and risk factors for fatty liver in a general population of Shanghai, China. J Hepatol. 2005;43(3):508-14.
246. Jimba S, Nakagami T, Takahashi M, Wakamatsu T, Hirota Y, Iwamoto Y, et al. Prevalence of non-alcoholic fatty liver disease and its association with impaired glucose metabolism in Japanese adults. Diabetic Medicine. 2005;22(9):1141-5.
247. Park SH, Jeon WK, Kim SH, Kim HJ, Park DI, Cho YK, et al. Prevalence and risk factors of non-alcoholic fatty liver disease among Korean adults. J Gastroenterol Hepatol. 2006;21(1 Pt 1):138-43.
248. Wong VW, Hui AY, Tsang SW, Chan JL, Wong GL, Chan AW, et al. Prevalence of undiagnosed diabetes and postchallenge hyperglycaemia in Chinese patients with non-alcoholic fatty liver disease. Aliment Pharmacol Ther. 2006;24(8):1215-22.
249. El-serag HB, Tran T, Everhart JE**.** Diabetes increases the risk of chronic liver disease and hepatocellular carcinoma. Gastroenterology. 2004;126(2):460-8.
250. Jacqueminet S, Lebray P, Morra R, Munteanu M, Devers L, Messous D, et al. Screening for liver fibrosis by using a noninvasive biomarker in patients with diabetes. Clin Gastroenterol Hepatol. 2008;6(7):828-31.
251. Casey SP, Kemp WW, McLean CA, Topliss DJ, Adams LA, Roberts SK**.** A prospective evaluation of the role of transient elastography for the detection of hepatic fibrosis in type 2 diabetes without overt liver disease. Scand J Gastroenterol. 2012;47(7):836-41.
252. de Ledinghen V, Vergniol J, Gonzalez C, Foucher J, Maury E, Chemineau L, et al. Screening for liver fibrosis by using FibroScan((R)) and FibroTest in patients with diabetes. Dig Liver Dis. 2012;44(5):413-8.
253. Morling JR, Fallowfield JA, Guha IN, Nee LD, Glancy S, Williamson RM, et al. Using non-invasive biomarkers to identify hepatic fibrosis in people with type 2 diabetes mellitus: the Edinburgh type 2 diabetes study. J Hepatol. 2014;60(2):384-91.
254. Das K, Das K, Mukherjee PS, Ghosh A, Ghosh S, Mridha AR, et al. Nonobese population in a developing country has a high prevalence of nonalcoholic fatty liver and significant liver disease. Hepatology. 2010;51(5):1593-602.
255. Mofrad P, Contos MJ, Haque M, Sargeant C, Fisher RA, Luketic VA, et al. Clinical and histologic spectrum of nonalcoholic fatty liver disease associated with normal ALT values. Hepatology. 2003;37(6):1286-92.
256. Sanyal AJ, Chalasani N, Kowdley KV, McCullough A, Diehl AM, Bass NM, et al. Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. New England Journal of Medicine. 2010;362(18):1675-85.

257. Zhang X, Harmsen WS, Mettler TA, Kim WR, Roberts RO, Therneau TM, et al. Continuation of metformin use after a diagnosis of cirrhosis significantly improves survival of patients with diabetes. Hepatology. 2014;60(6):2008-16.
258. Lavanchy D. Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. Journal of Viral Hepatitis. 2004;11(2):97-107.
259. Liaw Y-F, Kao J-H, Piratvisuth T, Chan HLY, Chien R-N, Liu C-J, et al. Asian-Pacific consensus statement on the management of chronic hepatitis B: a 2012 update. Hepatology International. 2012;6(3):531-61.
260. Chen CJ, Yang HI. Natural history of chronic hepatitis B REVEALed. J Gastroenterol Hepatol. 2011;26(4):628-38.
261. Ganem D, Prince AM. Hepatitis B virus infection--natural history and clinical consequences. New England Journal of Medicine. 2004;350(11):1118-29.
262. Liaw YF. Prevention and surveillance of hepatitis B virus-related hepatocellular carcinoma. Seminars in Liver Disease. 2005;25 Suppl 1:40-7.
263. Hoofnagle JH, Doo E, Liang TJ, Fleischer R, Lok AS. Management of hepatitis B: summary of a clinical research workshop. Hepatology. 2007;45(4):1056-75.
264. Nguyen VT, Razali K, Amin J, Law MG, Dore GJ. Estimates and projections of hepatitis B-related hepatocellular carcinoma in Australia among people born in Asia-Pacific countries. Journal of Gastroenterology & Hepatology. 2008;23(6):922-9.
265. Raimondo G, Allain JP, Brunetto MR, Buendia MA, Chen DS, Colombo M, et al. Statements from the Taormina expert meeting on occult hepatitis B virus infection. J Hepatol. 2008;49(4):652-7.
266. Liaw YF, Jia JD, Chan HL, Han KH, Tanwandee T, Chuang WL, et al. 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. 2011;54(5):1591-9.
267. Ono A, Suzuki F, Kawamura Y, Sezaki H, Hosaka T, Akuta N, et al. Long-term continuous entecavir therapy in nucleos(t)ide-naive chronic hepatitis B patients. J Hepatol. 2012;57(3):508-14.
268. Chaung KT, Ha NB, Trinh HN, Garcia RT, Nguyen HA, Nguyen KK, et al. High frequency of recurrent viremia after hepatitis B e antigen seroconversion and consolidation therapy. J Clin Gastroenterol. 2012;46(10):865-70.
269. Fung J, Lai CL, Tanaka Y, Mizokami M, Yuen J, Wong DK, et al. The duration of lamivudine therapy for chronic hepatitis B: cessation vs. continuation of treatment after HBeAg seroconversion. Am J Gastroenterol. 2009;104(8):1940-6; quiz 7.
270. Seto WK, Hui AJ, Wong VW, Wong GL, Liu KS, Lai CL, et al. Treatment cessation of entecavir in Asian patients with hepatitis B e antigen negative chronic hepatitis B: a multicentre prospective study. Gut. 2014.
271. Australia DHFotGSo. Australian and New Zealand Chronic Hepatitis B (CHB) Receommendations. 2nd ed; 2009.
272. Terrault NA, Bzowej NH, Chang KM, Hwang JP, Jonas MM, Murad MH. AASLD guidelines for treatment of chronic hepatitis B. Hepatology. 2016;63(1):261-83.
273. Liu J, Yang HI, Lee MH, Lu SN, Jen CL, Wang LY, et al. Incidence and determinants of spontaneous hepatitis B surface antigen seroclearance: a community-based follow-up study. Gastroenterology. 2010;139(2):474-82.
274. Liaw YF, Sheen IS, Chen TJ, Chu CM, Pao CC. Incidence, determinants and significance of delayed clearance of serum HBsAg in chronic hepatitis B virus infection: a prospective study. Hepatology. 1991;13(4):627-31.
275. Alward WL, McMahon BJ, Hall DB, Heyward WL, Francis DP, Bender TR. The long-term serological course of asymptomatic hepatitis B virus carriers and the development of primary hepatocellular carcinoma. J Infect Dis. 1985;151(4):604-9.
276. Liaw YF. Antiviral therapy of chronic hepatitis B: opportunities and challenges in Asia. J Hepatol. 2009;51(2):403-10.
277. Lok AS, McMahon BJ. Chronic hepatitis B: update 2009. Hepatology. 2009;50(3):661-2.

278. Shimamatsu K, Kage M, Nakashima O, Kojiro M**.** Pathomorphological study of HCV antibody-positive liver cirrhosis. J Gastroenterol Hepatol. 1994;9(6):624-30.
279. Kojiro M, Shimamatsu K, Kage M**.** Pathomorphologic comparison of hepatitis C virus-related and hepatitis B virus-related cirrhosis bearing hepatocellular carcinoma. Princess Takamatsu Symp. 1995;25:179-84.
280. Wong GL, Wong VW, Choi PC, Chan AW, Chum RH, Chan HK, et al. Assessment of fibrosis by transient elastography compared with liver biopsy and morphometry in chronic liver diseases. Clin Gastroenterol Hepatol. 2008;6(9):1027-35.
281. Park GJ, Katelaris PH, Jones DB, Seow F, Le Couteur DG, Ngu MC**.** Validity of the 13C-caffeine breath test as a noninvasive, quantitative test of liver function. Hepatology. 2003;38(5):1227-36.
282. Park GJ, Katelaris PH, Jones DB, Seow F, Lin BP, Le Couteur DG, et al. The C-caffeine breath test distinguishes significant fibrosis in chronic hepatitis B and reflects response to lamivudine therapy. Alimentary Pharmacology & Therapeutics. 2005;22(5):395-403.
283. Hui AY, Chan HL, Wong VW, Liew CT, Chim AM, Chan FK, et al. Identification of chronic hepatitis B patients without significant liver fibrosis by a simple noninvasive predictive model. Am J Gastroenterol. 2005;100(3):616-23.
284. Zeng MD, Lu LG, Mao YM, Qiu DK, Li JQ, Wan MB, et al. Prediction of significant fibrosis in HBeAg-positive patients with chronic hepatitis B by a noninvasive model. Hepatology. 2005;42(6):1437-45.
285. Kim BK, Kim SA, Park YN, Cheong JY, Kim HS, Park JY, et al. Noninvasive models to predict liver cirrhosis in patients with chronic hepatitis B. Liver International. 2007;27(7):969-76.
286. Chen YP, Dai L, Wang JL, Zhu YF, Feng XR, Hou JL**.** Model consisting of ultrasonographic and simple blood indexes accurately identify compensated hepatitis B cirrhosis. Journal of Gastroenterology & Hepatology. 2008;23(8 Pt 1):1228-34.
287. Vallet-Pichard A, Mallet V, Nalpas B, Verkarre V, Nalpas A, Dhalluin-Venier V, et al. FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. Hepatology. 2007;46(1):32-6.
288. Giannini E, Risso D, Botta F, Chiarbonello B, Fasoli A, Malfatti F, et al. Validity and clinical utility of the aspartate aminotransferase-alanine aminotransferase ratio in assessing disease severity and prognosis in patients with hepatitis C virus-related chronic liver disease. Archives of Internal Medicine. 2003;163(2):218-24.
289. Poynard T, Bedossa P**.** 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. Journal of Viral Hepatitis. 1997;4(3):199-208.
290. Ohta T, Sakaguchi K, Fujiwara A, Fujioka S, Iwasaki Y, Makino Y, et al. Simple surrogate index of the fibrosis stage in chronic hepatitis C patients using platelet count and serum albumin level. Acta Medica Okayama. 2006;60(2):77-84.
291. Williams AL, Hoofnagle JH**.** Ratio of serum aspartate to alanine aminotransferase in chronic hepatitis. Relationship to cirrhosis. Gastroenterology. 1988;95(3):734-9.
292. Bonacini M, Hadi G, Govindarajan S, Lindsay KL**.** Utility of a discriminant score for diagnosing advanced fibrosis or cirrhosis in patients with chronic hepatitis C virus infection. American Journal of Gastroenterology. 1997;92(8):1302-4.
293. Imbert-Bismut F, Ratziu V, Pieroni L, Charlotte F, Benhamou Y, Poynard T, et al. Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study. Lancet. 2001;357(9262):1069-75.
294. Forns X, Ampurdanes S, Llovet JM, Aponte J, Quinto L, Martinez-Bauer E, et al. Identification of chronic hepatitis C patients without hepatic fibrosis by a simple predictive model. Hepatology. 2002;36(4 Pt 1):986-92.
295. Rosenberg WM, Voelker M, Thiel R, Becka M, Burt A, Schuppan D, et al. Serum markers detect the presence of liver fibrosis: a cohort study. Gastroenterology. 2004;127(6):1704-13.

296. Adams LA, Bulsara M, Rossi E, DeBoer B, Speers D, George J, et al. Hepascore: an accurate validated predictor of liver fibrosis in chronic hepatitis C infection. Clinical Chemistry. 2005;51(10):1867-73.
297. Adams LA, Bulsara M, Rossi E, DeBoer B, Speers D, George J, et al. Hepascore: an accurate validated predictor of liver fibrosis in chronic hepatitis C infection. Clin Chem. 2005;51(10):1867-73.
298. Islam S, Antonsson L, Westin J, Lagging M**.** Cirrhosis in hepatitis C virus-infected patients can be excluded using an index of standard biochemical serum markers. Scand J Gastroenterol. 2005;40(7):867-72.
299. Koda M, Matunaga Y, Kawakami M, Kishimoto Y, Suou T, Murawaki Y**.** FibroIndex, a practical index for predicting significant fibrosis in patients with chronic hepatitis C. Hepatology. 2007;45(2):297-306.
300. Zhou K, Gao CF, Zhao YP, Liu HL, Zheng RD, Xian JC, et al. Simpler score of routine laboratory tests predicts liver fibrosis in patients with chronic hepatitis B. J Gastroenterol Hepatol. 2010;25(9):1569-77.
301. Park GJ, Wiseman E, George J, Katelaris PH, Seow F, Fung C, et al. Non-invasive estimation of liver fibrosis in non-alcoholic fatty liver disease using the 13 C-caffeine breath test. Journal of Gastroenterology & Hepatology. 2011;26(9):1411-6.
302. Cholongitas E, Papatheodoridis GV, Vangeli M, Terreni N, Patch D, Burroughs AK**.** Systematic review: The model for end-stage liver disease--should it replace Child-Pugh's classification for assessing prognosis in cirrhosis? Aliment Pharmacol Ther. 2005;22(11-12):1079-89.
303. Chen YP, Peng J, Hou JL**.** Non-invasive assessment of liver fibrosis in patients with chronic hepatitis B. Hepatol Int. 2013;7(2):356-68.
304. Chon YE, Choi EH, Song KJ, Park JY, Kim do Y, Han KH, et al. Performance of transient elastography for the staging of liver fibrosis in patients with chronic hepatitis B: a meta-analysis. PLoS One. 2012;7(9):e44930.
305. Wong GL**.** Update of liver fibrosis and steatosis with transient elastography (Fibroscan). Gastroenterol Rep (Oxf). 2013;1(1):19-26.
306. Du D, Zhu X, Kuno A, Matsuda A, Tsuruno C, Yu D, et al. Comparison of LecT-Hepa and FibroScan for assessment of liver fibrosis in hepatitis B virus infected patients with different ALT levels. Clinica Chimica Acta. 2012;413(21-22):1796-9.
307. Degos F, Perez P, Roche B, Mahmoudi A, Asselineau J, Voitot H, et al. Diagnostic accuracy of FibroScan and comparison to liver fibrosis biomarkers in chronic viral hepatitis: a multicenter prospective study (the FIBROSTIC study). Journal of Hepatology. 2010;53(6):1013-21.
308. Kim BK, Kim SU, Kim HS, Park JY, Ahn SH, Chon CY, et al. Prospective validation of FibroTest in comparison with liver stiffness for predicting liver fibrosis in Asian subjects with chronic hepatitis B. PLoS ONE [Electronic Resource]. 2012;7(4):e35825.
309. Cardoso AC, Carvalho-Filho RJ, Stern C, Dipumpo A, Giuily N, Ripault MP, et al. Direct comparison of diagnostic performance of transient elastography in patients with chronic hepatitis B and chronic hepatitis C. Liver International. 2012;32(4):612-21.
310. Kim do Y, Kim SU, Ahn SH, Park JY, Lee JM, Park YN, et al. Usefulness of FibroScan for detection of early compensated liver cirrhosis in chronic hepatitis B. Digestive Diseases & Sciences. 2009;54(8):1758-63.
311. Chon YE, Choi EH, Song KJ, Park JY, Kim do Y, Han KH, et al. Performance of transient elastography for the staging of liver fibrosis in patients with chronic hepatitis B: a meta-analysis. PLoS ONE [Electronic Resource]. 2012;7(9):e44930.
312. Guirgis M, Manoharan S, Kwok R, Scott DR, Lee AU, Connor SJ, et al. Transient Elastography (FIBROSCAN) in chronic hepatitis B infection: Reduce number of biopsies and save money. Hepatology. 2009;50:543A.

313. Poynard T, Bedossa P**.** 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. 1997;4(3):199-208.
314. Kim BK, Kim do Y, Park JY, Ahn SH, Chon CY, Kim JK, et al. Validation of FIB-4 and comparison with other simple noninvasive indices for predicting liver fibrosis and cirrhosis in hepatitis B virus-infected patients. Liver International. 2010;30(4):546-53.
315. Xiao G, Yang J, Yan L**.** 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. 2015;61(1):292-302.
316. Chen YP, Liang XE, Dai L, Zhang Q, Peng J, Zhu YF, et al. Improving transient elastography performance for detecting hepatitis B cirrhosis. Digestive & Liver Disease. 2012;44(1):61-6.
317. Shin WG, Park SH, Jang MK, Hahn TH, Kim JB, Lee MS, et al. Aspartate aminotransferase to platelet ratio index (APRI) can predict liver fibrosis in chronic hepatitis B. Digestive & Liver Disease. 2008;40(4):267-74.
318. Zhang Z, Wang G, Kang K, Wu G, Wang P**.** The Diagnostic Accuracy and Clinical Utility of Three Noninvasive Models for Predicting Liver Fibrosis in Patients with HBV Infection. PLoS One. 2016;11(4):e0152757.
319. Ding D, Li H, Liu P, Chen L, Kang J, Zhang Y, et al. FibroScan, aspartate aminotransferase and alanine aminotransferase ratio (AAR), aspartate aminotransferase to platelet ratio index (APRI), fibrosis index based on the 4 factor (FIB-4), and their combinations in the assessment of liver fibrosis in patients with hepatitis B. Int J Clin Exp Med. 2015;8(11):20876-82.
320. Eminler AT, Ayyildiz T, Irak K, Kiyici M, Gurel S, Dolar E, et al. AST/ALT ratio is not useful in predicting the degree of fibrosis in chronic viral hepatitis patients. Eur J Gastroenterol Hepatol. 2015;27(12):1361-6.
321. Poynard T, Ngo Y, Munteanu M, Thabut D, Ratziu V**.** Noninvasive markers of hepatic fibrosis in chronic hepatitis B. Current Hepatitis Reports. 2011;10(2):87-97.
322. Raftopoulos SC, George J, Bourliere M, Rossi E, de Boer WB, Jeffrey GP, et al. Comparison of noninvasive models of fibrosis in chronic hepatitis B. Hepatol Int. 2012;6(2):457-67.
323. Afdhal NH, Nunes D**.** Evaluation of liver fibrosis: a concise review. Am J Gastroenterol. 2004;99(6):1160-74.
324. Standish RA, Cholongitas E, Dhillon A, Burroughs AK, Dhillon AP**.** An appraisal of the histopathological assessment of liver fibrosis. Gut. 2006;55(4):569-78.
325. Isgro G, Calvaruso V, Andreana L, Luong TV, Garcovich M, Manousou P, et al. The relationship between transient elastography and histological collagen proportionate area for assessing fibrosis in chronic viral hepatitis. J Gastroenterol. 2013;48(8):921-9.
326. Jung KS, Kim SU, Ahn SH, Park YN, Kim do Y, Park JY, et al. Risk assessment of hepatitis B virus-related hepatocellular carcinoma development using liver stiffness measurement (FibroScan). Hepatology. 2011;53(3):885-94.
327. Chan HL, Leung NW, Hui AY, Wong VW, Liew CT, Chim AM, et al. A randomized, controlled trial of combination therapy for chronic hepatitis B: comparing pegylated interferon-alpha2b and lamivudine with lamivudine alone.[Summary for patients in Ann Intern Med. 2005 Feb 15;142(4):I30; PMID: 15710953]. Annals of Internal Medicine. 2005;142(4):240-50.
328. Chang TT, Liaw YF, Wu SS, Schiff E, Han KH, Lai CL, et al. Long-term entecavir therapy results in the reversal of fibrosis/cirrhosis and continued histological improvement in patients with chronic hepatitis B. Hepatology. 2010;52(3):886-93.
329. Dienstag JL, Goldin RD, Heathcote EJ, Hann HW, Woessner M, Stephenson SL, et al. Histological outcome during long-term lamivudine therapy. Gastroenterology. 2003;124(1):105-17.
330. Yuen MF, Seto WK, Chow DH, Tsui K, Wong DK, Ngai VW, et al. Long-term lamivudine therapy reduces the risk of long-term complications of chronic hepatitis B infection even in patients without advanced disease. Antivir Ther. 2007;12(8):1295-303.

331. Kurokawa M, Hiramatsu N, Oze T, Yakushijin T, Miyazaki M, Hosui A, et al. Long-term effect of lamivudine treatment on the incidence of hepatocellular carcinoma in patients with hepatitis B virus infection. J Gastroenterol. 2012;47(5):577-85.
332. Liaw YF, Leung N, Kao JH, Piratvisuth T, Gane E, Han KH, et al. Asian-Pacific consensus statement on the management of chronic hepatitis B: a 2008 update. Hepatol Int. 2008;2(3):263-83.
333. Fung J, Lai CL, Wong DK, Seto WK, Hung I, Yuen MF**.** Significant changes in liver stiffness measurements in patients with chronic hepatitis B: 3-year follow-up study. J Viral Hepat. 2011;18(7):e200-5.
334. Kim SU, Park JY, Kim do Y, Ahn SH, Choi EH, Seok JY, et al. Non-invasive assessment of changes in liver fibrosis via liver stiffness measurement in patients with chronic hepatitis B: impact of antiviral treatment on fibrosis regression. Hepatol Int. 2010;4(4):673-80.
335. Enomoto M, Mori M, Ogawa T, Fujii H, Kobayashi S, Iwai S, et al. Usefulness of transient elastography for assessment of liver fibrosis in chronic hepatitis B: Regression of liver stiffness during entecavir therapy. Hepatol Res. 2010;40(9):853-61.
336. Wong GL, Wong VW, Choi PC, Chan AW, Chim AM, Yiu KK, et al. On-treatment monitoring of liver fibrosis with transient elastography in chronic hepatitis B patients. Antivir Ther. 2011;16(2):165-72.
337. Fung J, Lai CL, Seto WK, Wong DK, Yuen MF**.** Prognostic significance of liver stiffness for hepatocellular carcinoma and mortality in HBeAg-negative chronic hepatitis B. J Viral Hepat. 2011;18(10):738-44.
338. Saag KG, Teng GG, Patkar NM, Anuntiyo J, Finney C, Curtis JR, et al. American College of Rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. Arthritis Rheum. 2008;59(6):762-84.
339. Kalb RE, Strober B, Weinstein G, Lebwohl M**.** Methotrexate and psoriasis: 2009 National Psoriasis Foundation Consensus Conference. J Am Acad Dermatol. 2009;60(5):824-37.
340. Rajagopalan PT, Zhang Z, McCourt L, Dwyer M, Benkovic SJ, Hammes GG**.** Interaction of dihydrofolate reductase with methotrexate: ensemble and single-molecule kinetics. Proc Natl Acad Sci U S A. 2002;99(21):13481-6.
341. Kremer JM, Lee RG, Tolman KG**.** Liver histology in rheumatoid arthritis patients receiving long-term methotrexate therapy. A Prospective Study with Baseline and Sequential Biopsy Samples. Arthritis & Rheumatism. 1989;32(2):121-7.
342. Guidelines for monitoring drug therapy in rheumatoid arthritis. American College of Rheumatology Ad Hoc Committee on Clinical Guidelines. Arthritis Rheum. 1996;39(5):723-31.
343. Rodenhuis S, Kremer JM, Bertino JR**.** Increase of dihydrofolate reductase in peripheral blood lymphocytes of rheumatoid arthritis patients treated with low-dose oral methotrexate. Arthritis & Rheumatism. 1987;30(4):369-74.
344. Berkowitz RS, Goldstein DP, Bernstein MR**.** Ten year's experience with methotrexate and folinic acid as primary therapy for gestational trophoblastic disease. Gynecologic Oncology. 1986;23(1):111-8.
345. Kremer JM, Kaye GI, Kaye NW, Ishak KG, Axiotis CA**.** Light and electron microscopic analysis of sequential liver biopsy samples from rheumatoid arthritis patients receiving long-term methotrexate therapy. Followup over long treatment intervals and correlation with clinical and laboratory variables. Arthritis & Rheumatism. 1995;38(9):1194-203.
346. Kremer JM, Lee RG, Tolman KG**.** Liver histology in rheumatoid arthritis patients receiving long-term methotrexate therapy. A prospective study with baseline and sequential biopsy samples. Arthritis & Rheumatism. 1989;32(2):121-7.
347. Kremer JM, Alarcon GS, Lightfoot RW, Jr., Willkens RF, Furst DE, Williams HJ, et al. Methotrexate for rheumatoid arthritis. Suggested guidelines for monitoring liver toxicity. American College of Rheumatology. Arthritis & Rheumatism. 1994;37(3):316-28.

348. Barbero-Villares A, Mendoza J, Trapero-Marugan M, Gonzalez-Alvaro I, Dauden E, Gisbert JP, et al. Evaluation of liver fibrosis by transient elastography in methotrexate treated patients. Medicina Clinica. 2011;137(14):637-9.
349. Barbero-Villares A, Mendoza Jimenez-Ridruejo J, Taxonera C, Lopez-Sanroman A, Pajares R, Bermejo F, et al. Evaluation of liver fibrosis by transient elastography (Fibroscan(R)) in patients with inflammatory bowel disease treated with methotrexate: a multicentric trial. Scand J Gastroenterol. 2012;47(5):575-9.
350. Laharie D, Seneschal J, Schaeverbeke T, Doutre M-S, Longy-Boursier M, Pellegrin J-L, et al. Assessment of liver fibrosis with transient elastography and FibroTest in patients treated with methotrexate for chronic inflammatory diseases: a case-control study. J Hepatol. 2010;53(6):1035-40.
351. Zachariae H**.** Have methotrexate-induced liver fibrosis and cirrhosis become rare? A matter for reappraisal of routine liver biopsies. Dermatology. 2005;211(4):307-8.
352. Fleming KM, Aithal GP, Solaymani-Dodaran M, Card TR, West J**.** Incidence and prevalence of cirrhosis in the United Kingdom, 1992-2001: a general population-based study. J Hepatol. 2008;49(5):732-8.
353. Farrell GC**.** Drugs and steatohepatitis. Semin Liver Dis. 2002;22(2):185-94.
354. Menter A, Korman NJ, Elmets CA, Feldman SR, Gelfand JM, Gordon KB, et al. Guidelines of care for the management of psoriasis and psoriatic arthritis: section 4. Guidelines of care for the management and treatment of psoriasis with traditional systemic agents. J Am Acad Dermatol. 2009;61(3):451-85.