STRATIFIED MEDICINE: AN EXPLORATION OF THE UTILITY OF NON-INVASIVE SERUM MARKERS FOR THE MANAGEMENT OF CHRONIC LIVER DISEASES

Sudeep Tanwar

Thesis submitted to the University College of London

for the degree of Doctor of Philosophy

2018

DECLARATION

I, Sudeep Tanwar, confirm that the work presented in this thesis is my own.

Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ABSTRACT

Chronic liver disease (CLD), the 3rd commonest cause of premature death in the UK, is detected late when interventions are often ineffective. Non-alcoholic fatty liver disease (NAFLD) and chronic hepatitis C (CHC) account for a significant proportion of CLD in the UK. Numerous direct (molecules involved in matrix biology) and indirect biomarkers (standard laboratory tests) have been successfully developed to detect advanced liver fibrosis. Less success, however, has been achieved in the detection of alternative diagnostic targets such as early stage fibrosis, non-alcoholic steatohepatitis (NASH) and fibrosis evolution.

In a study of 17 candidate biomarkers amongst patients with NAFLD, terminal peptide of procollagen 3 was identified as the only biomarker demonstrating good performance for the detection of NASH in both a derivation and validation cohort. Thereafter, these results were further validated in another NAFLD cohort.

In a study of 9 biomarkers (indirect and direct) in the detection of fibrosis in NAFLD, direct biomarkers demonstrated better diagnostic performance overall and for early stage fibrosis although some indirect biomarkers identified advanced fibrosis and cirrhosis with good effect.

Thereafter, parallel and serial combinations of 3 biomarkers of advanced fibrosis were proposed and successfully employed in a cohort of patients with NAFLD to improve diagnostic performance.

In a study of 10 biomarkers in CHC, fibrosis detection was enhanced using complex biomarker panels that incorporated direct tests. Of note, the use an alternative assay for a constituent component significantly affected biomarker panel performance both overall and at diagnostic thresholds.

The ability of the biomarkers to monitor fibrosis evolution arising due to putative antifibrotic was then studied in CHC. In the first study, changes in direct biomarker, ELF, could predict fibrosis evolution. In the second study, an improvement of indirect biomarker scores in patients with CHC cirrhosis during treatment was found to denote an improved prognosis.

ACKOWLEDGEMENTS

I am deeply grateful to Prof William Rosenberg for his supervision and friendship throughout this period of research. I have gained a range of skills which prior to starting this thesis I had thought imaginable. In addition to understanding how one should act as a mentor, I also understand the importance of having a collaborative approach to research. This body of work is the product of regional, national and international collaboration with centres in the UK (London, Southampton, Nottingham, Plymouth, Newcastle, Glasgow), Austria and Germany.

I am also indebted to Prof Geoffrey Dusheiko particularly for providing funding at the start of my period of 'out of programme research'. The several manuscripts that we published together formed an excellent companion to the translational biomarker studies presented in this thesis.

This project would have also been far more difficult without the help of Dr Julie Parkes who provided an invaluable introduction into biostatistics and translational research. Thereafter, many ideas were shared with Dr Paul Trembling and Carmen Velasquez in both Bloomsbury and Hampstead.

At the Nottingham biomedical research unit, special thanks go to Dr Neil Guha and Professor Guru Aithal for facilitating our collaboration on NAFLD which proved so productive.

Finally, I would like to thank my wife (Manjari), my children (Ishaan and Arjun) and my parents (Shiv Raj and Saroj) for their unwavering support without which this thesis would not have been completed.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

CHAPTER 8

CHAPTER 9

CHAPTER 10

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

CHAPTER 8

CHAPTER 9

CHAPTER 10

CHAPTER 11

COMMUNICATIONS ARISING DIRECTLY FROM THESIS

PEER REVIEWED PUBLICATIONS

- **The Enhanced Liver Fibrosis (ELF) Panel: analyte stability under common sample storage conditions used in clinical practice.**
  Kennedy O, Parkes J, Tanwar S, Trembling P, Rosenberg WM
  Applied Laboratory Medicine May 2017, 1 (6) 720-728
  doi:101.1373/jalm.2016.022806
- **Non-invasive markers of liver fibrosis: on-treatment changes of serum markers predict the outcome of antifibrotic therapy.**
  Tanwar S, Trembling PM, Hogan BJ, Srivastava A, Parkes J, Harris S, Grant P, Nastouli E, Ocker M, Wehr K, Herold C, Neureiter D, Schuppan D, Rosenberg WM.
  Eur J Gastroenterol Hepatol 2017 Mar;29(3):289-296
  doi: 10.1097/MEG.0000000000000789. PMID: 27906753
- **Biomarkers of Hepatic Fibrosis in Chronic Hepatitis C: A Comparison of 10 Biomarkers Using 2 Different Assays for Hyaluronic Acid.**
  Tanwar S, Trembling PM, Hogan BJ, Parkes J, Harris S, Grant P, Nastouli E, Ocker M, Wehr K, Herold C, Neureiter D, Schuppan D, Rosenberg WM.
  J Clin Gastroenterol 2017 Mar; 51(3):268-277
  doi: 10.1097/MCG.0000000000000581. PMID:27380461
- **Complexities of HCV management in the new era of direct-acting antiviral agents**
  Rosenberg WM, Tanwar S, Trembling P
  QJM. 2014 Jan;107(1):17-9.
  doi: 10.1093/qjmed/hct181. PMID: 24065837
- **Treatment decisions and contemporary versus pending treatments for hepatitis C.**
  Trembling PM, Tanwar S, Rosenberg WM, Dusheiko GM
  Nat Rev Gastroenterol Hepatol. 2013 Dec;10(12):713-28.
  doi: 10.1038/nrgastro.2013.163. PMID: 24019151
- **Validation of PIIINP for the detection and assessment of non-alcoholic steatohepatitis in patients with non-alcoholic fatty liver disease.**

Tanwar S, Trembling PM, Guha IN, Parkes J, Kaye P, Burt AD, Ryder SD, Aithal GP, Day CP, Rosenberg WM.
Hepatology. 2013 Jan;57(1):103-11.
doi: 10.1002/hep.26030. PMID: 22930399

- **TMC435 for the treatment of chronic hepatitis C.**
  Tanwar S, Trembling PM, Dusheiko GM
  Expert Opin Investig Drugs. 2012 Aug;21(8):1193-209.
  doi: 10.1517/13543784.2012.69039 PMID: 22616626
- **Randomized clinical trial: a pilot study investigating the safety and effectiveness of an escalating dose of peginterferon α-2a monotherapy for 48 weeks compared with standard clinical care in patients with hepatitis C cirrhosis.**
  Tanwar S, Wright M, Foster GR, Ryder SD, Mills PR, Cramp ME, Parkes J, Rosenberg WM.
  Eur J Gastroenterol Hepatol. 2012 May;24(5):543-50.
  doi: 10.1097/MEG.0b013e3283513e69. PMID: 22337287
- **Boceprevir: an oral protease inhibitor for the treatment of chronic HCV infection**.
  Trembling PM, Tanwar S, Dusheiko GM
  Expert Rev Anti Infect Ther. 2012 Mar;10(3):269-79.
  doi: 10.1586/eri.12.8. PMID: 22397560
- **Hepatitis C Therapy: Lessons of the Last Two Decades.**
  Tanwar S, Trembling PM, Dusheiko GM
  Curr Hepatitis Rep (2012) 11: 119.
  doi:10.1007/s11901-012-0141-7

BOOK CHAPTERS

- **Tietz Textbook of Clinical Chemistry and Molecular Diagnostics 6th Edition. Chapter 61: Liver Disease.**
  William Rosenberg, Tony Badrick, Sudeep Tanwar.
  ISBN: 978-0-323-35921-4

ORAL PRESENTATIONS AT INTERNATIONAL CONFERENCES

- **An algorithm combining direct and indirect panels for fibrosis can result in enhanced diagnostic performance for the detection of advanced fibrosis (F3-4) in NAFLD thus minimising the need for liver biopsy**
  Tanwar S, Trembling PM, Guha IN, Parkes J, Kaye P, Burt AD, Ryder SD, Aithal GP, Day CP, Rosenberg WM.
  Oral Presentation at AASLD Annual Meeting 2012, Boston, USA.
- **Changes in ELF score during treatment with Pegylated Interferon and Silymarin are associated with histological liver disease progression in non-responder patients with chronic hepatitis C.**
  Tanwar S, Trembling PM, Ellis E, Parkes J, Herold C, Rosenberg WM
  Oral Presentation at EASL Annual Meeting 2011, Berlin, Germany.

POSTER PRIZES AT INTERNATIONAL CONFERENCES

- **Direct serum markers are more accurate than simple marker panels for the detection of fibrosis in nonalcoholic fatty liver Disease (NAFLD).**
  Tanwar S, Trembling PM, Guha IN, Parkes J, Kaye P, Burt AD, Ryder SD, Aithal GP, Day CP, Rosenberg WM.
  Presidential Poster of Distinction at DDF Annual Meeting 2012, Birmingham, UK.
- **Non-invasive assessment of hepatic fibrosis in prior non-responders to HCV treatment- a comparison of 8 marker panels of liver fibrosis.**
  Tanwar S, Ellis E, Parkes J, Herold C, Rosenberg WM
  Presidential Poster of Distinction at AASLD Annual Meeting 2010, Boston USA.

POSTER PRESENTATIONS AT INTERNATIONAL CONFERENCES

- **Primary Care Sequential Use of FIB-4 and the Enhanced Liver Fibrosis Test to Stratify Patients with Non-alcoholic Fatty Liver Doubles Cirrhosis Detection and Reduces Referrals of Patients with Mild Disease**

Srivastava A, Gailer R, Demma S, Warner A, Suri D, Morgan S, Tanwar S, Sennett K, Thorburn D, Parkes J, Tsochatzis E, Rosenberg W

Poster Presentation at EASL Annual Meeting 2016, Barcelona, Spain

- **A one year retrospective review of new patient attendances at a tertiary hepatology centre highlighting the increasing challenge of NAFLD and the need to develop clinical pathways**

  Srivastava A, Gailer R, Tanwar S, Trembling P, Warner A, Morgan S, Sennett K, Thorburn D, Tsochatzis M, Rosenberg W

  Poster Presentation at EASL Annual Meeting 2015, Vienna, Austria

- **Further Validation of Terminal Peptide of Procollagen III (PIIINP) for the Detection and Assessment of Nonalcoholic Steatohepatitis in Patients with Nonalcoholic Fatty Liver Disease.**

  Tanwar S, Parkes J, Trembling PM, Hogan B, Schuppan D, Pinzani M, Arthur M, Burt AD, Rosenberg WM.

  Poster Presentation at EASL Annual Meeting 2013, Amsterdam, Netherlands.

- **Comparison of 4 serum marker panels of fibrosis in chronic hepatitis C (CHC): variants of the hyaluronic acid assay significantly affect their diagnostic performance.**

  Tanwar S, Trembling PM, Hogan B, Ellis E, Parkes J, Herold C, Schuppan D, Rosenberg WM

  Poster Presentation at BSG Annual Meeting 2013, Glasgow, UK.

- **Direct non-invasive serum markers of liver fibrosis predict fibrosis evolution in chronic hepatitis C but are increased by Interferon-based therapy.**

  Tanwar S, Trembling PM, Ellis E, Parkes J, Herold C, Schuppan D, Rosenberg WM

  Poster Presentation at EASL Annual Meeting 2012, Barcelona, Spain.

- **Validation of a novel biomarker model for the prediction of non-alcoholic steatohepatitis in patients with non-alcoholic fatty liver disease.**

  Tanwar S, Trembling PM, Guha IN, Parkes J, Kaye P, Burt AD, Ryder SD, Aithal GP, Day CP, Rosenberg WM.

  Poster Presentation at AASLD Annual Meeting 2011, San Francisco, USA.

- **PACIFIC: A phase III, randomized, multicenter, dose escalation, efficacy and safety study examining the effects of treatment with peginterferon alfa-2a in patients with Child's A or B cirrhosis in chronic hepatitis C virus infection.**
  Tanwar S, Wright M, Foster GR, Ryder SD, Mills PR, Cramp ME, Parkes J, Rosenberg WM.
  Poster Presentation at AASLD Annual Meeting 2010, Boston USA.

INTERNET PRESENTATIONS (PODCASTS)

- **"A New Non-invasive test for NASH"**
  Drs. Stephen Harrison and Sudeep Tanwar
  https://soundcloud.com/aasld/a-new-non-invasive-test-for
  Described by AASLD as a 'high-level discussion', following my Hepatology publication, I was invited by AASLD to be interviewed to produce a 15 minute podcast.

LIST OF ABBRIEVIATIONS

2D-SWE	2D-shear wave elastography
ADAM	A Disintegrin and Metallopoteinase-domain protein
ALD	Alcoholic liver disease
ALT	Alanine Aminotransferase
APRI	AST to platelet ratio index
ARFI	Acoustic radiation impulse imaging
AST	Aspartate Aminotransferase
AUROC	Area under receiver operating characteristic curve
BCR	B cell receptors
BMI	Body mass index
CD	Cluster of Differentiation
CDS	Cirrhosis discriminant score
CHC	Chronic hepatitis C
CI	Confidence interval
CK-18	Cytokeratin-18
CLD	Chronic liver disease
CLR	C-like lectin receptors
Coll IV	Collagen IV
CTFG	Connective tissue growth factor
DAA	direct-acting antiviral agents
DAMPS	Damage associated molecular patterns
DANA	Difference between the mean fibrosis stage of advanced fibrosis minus the mean fibrosis stage of non-advanced fibrosis
DOR	Diagnostic odds ratio
ECM	Extracellular matrix
EFCAB4B	EF-hand calcium binding domain 4B
EGF	Epidermal growth factor
ELF	Enhanced Liver Fibrosis Test
EOT	End of treatment
F1	Metavir mild liver fibrosis
F2	Metavir moderate liver fibrosis

F3	Metavir advanced liver fibrosis
F4	Metavir cirrhosis
FDFT1	farnesyl diphosphate farnesyl transferase 1
FFA	Free fatty acids
FLI	Fatty liver index
FN	False negative
FP	False positive
FXR	Farnesoid X receptor
GCKR	glucokinase receptor
GGT	Gamma Glutamyltransferase
GLP-1	Glucagon-like peptide 1
GWAS	Genome-wide association studies
HA	Hyularonic acid
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HNP	Human neutrophil peptides
HRQL	Health related quality of life
HSC	Hepatic Stellate Cell
HSI	Hepatic steatosis index
HVPG	Hepatic Venous Pressure Gradient
IFN	Interferon
IL	Interleukin
ION	Index of NASH
IQR	Interquartile range
IRF	Interferon regulatory transcription factor
ITT	Intention to treat
IU	International units
JNK	Jun-(N)-terminal Kinase
kPa	Kilopascal
KC	Kupffer cells
LAP	Lipid accumulation product
LFT	Liver function tests
LOXL2	Lysyl oxidase 2
LR	Likelihood ratio

LR+	Positive likelihood ratio
LR-	Negative likelihood ratio
LRE	Liver related event
LS	Liver stiffness
LSEC	Liver Sinusoidal Endothelial Cells
MAP	Mitogen activated protein
MDA	Malondialdehyde
MLKL	Mixed-lineage kinase domain-like protein
MMP	Matrix Metalloproteinase
MRI	Magnetic resonance imaging
MRE	Magnetic resonance elastography
NAFLD	Non-alcoholic fatty liver disease;
NAS	NAFLD activity score
NASH	Non-alcoholic steatohepatitis
NICE	National institute for clinical excellence
NIT	Non-invasive test
NKT	Natural Killer T cells
NPV	Negative predictive value
NS	Non-significant
NLR	Nucleotide oligomerisation receptors
OR	Odds radio
ordAUROC	AUROC corrected by Obuchowski measure
ox-LDL	oxidised low-density lipoprotein
PAMPS	Pathogen-associated molecular patterns
PEG-INFα2a	Pegylated interferon alfa-2a
PEG-INFα2b	Pegylated interferon alfa-2b
PIIINP	Terminal peptide of pro-collagen III
PLT	Platelets
PMN	Polymorphonuclear leukocytes
PNPLA3	Patatin-like phospholipase domain containing 3
PPARγ	Proliferator-activated receptor γ
PPV	Positive predictive value
PRR	Pattern Recognition Receptors
RLR	RIG-1 like receptors

RNA	Ribonucleic acid
ROC	Receiver operator characteristic
ROS	Reactive oxygen species
S1	Scheuer mild liver fibrosis
S2	Scheuer moderate liver fibrosis
S3	Scheuer advanced liver fibrosis
S4	Scheuer stage cirrhosis
SCD	Stearoyl-COa desaturase
SH	Steatohepatitis
SS	Simple steatosis
SBP	Spontaneous Bacterial Peritonitis
SD	Standard Deviation
SEC	Sinusoidal epithelial cells
SVR	Sustained Virologic Response
TE	Transient elastography
TGF	Transforming Growth Factor
TIMP	Tissue inhibitor of matrix metalloproteinase
TLR	Toll Like Receptors
TM6SF2	Transmembrane 6 superfamily member 2
TN	True negative
TNF∝	Tumour necrosis factor alpha
TP	True positive
TRAIL	TNF-related apoptosis inducing ligand
WCC	White Cell Count

CHAPTER 1

RATIONALE FOR THESIS

ABSTRACT

Chronic liver disease (CLD), the 3rd commonest cause of premature death in the UK, has experienced a 400% increase in mortality since 1970. At present, most CLD is detected late, when interventions are ineffective, resulting in considerable morbidity and mortality. Obesity related non-alcoholic fatty liver disease (NAFLD) and chronic hepatitis C (CHC) account for a significant proportion of CLD in the UK. In both conditions a toxic insult to the liver causes inflammation, resulting in fibrosis leading to cirrhosis, portal hypertension, liver failure, liver cancer and death. However, the identification of CLD is challenging in the population at risk: standard liver function tests lack sensitivity and specificity for liver fibrosis and the reference test, liver biopsy, is not practicable as the primary method for detecting liver disease. Whilst serum-based biomarkers have been developed for the cross-sectional detection of liver disease, they have not been widely introduced into clinical practice.

Chapter 2 reviews the pathophysiology of inflammation and fibrosis in CLD with a focus on NAFLD and CHC.

Chapter 3 reviews the literature regarding the non-invasive detection of inflammation and fibrosis in CLD.

Chapter 4 explores the ability of candidate markers to detect non-alcoholic steatohepatitis (NASH) prior to the development of advanced fibrosis in both a derivation and validation cohort.

Chapter 5 further validates the findings derived in chapter 4 in another cohort of patients with NAFLD.

Chapter 6 compares the performance of 9 serum markers for the detection of mild, moderate, advanced fibrosis and cirrhosis due to NAFLD.

Chapter 7 explores whether the diagnostic performance of biomarkers of advanced fibrosis in NAFLD can be enhanced by their application as diagnostic algorithms.

Chapter 8 validates and compares the performance of 10 validated serum markers of fibrosis for the cross-sectional detection of fibrosis due to CHC whilst also exploring

the effect on diagnostic performance of employing an alternative assay for a biomarker constituent component.

Chapter 9 explores the ability of direct serum markers to longitudinally monitor changes in fibrosis arising due to putative anti-fibrotic therapy.

Chapter 10 explores whether the evolution of indirect serum marker scores arising from putative antifibrotic therapy is associated with development of liver related outcomes.

Finally, chapter 11 discusses the main conclusions drawn from this body of work as well as potential future studies.

Liver Disease in the UK

CLD, the 3rd commonest cause of premature death in the UK, and mortality has increased 400% since 1970.[1] This is in direct contrast to other diseases such as cardiovascular disease that have witnessed a decrease in mortality (figure 1.1). Moreover, the mortality from CLD in the UK continues to be one of the highest in western Europe (figure 1.2).

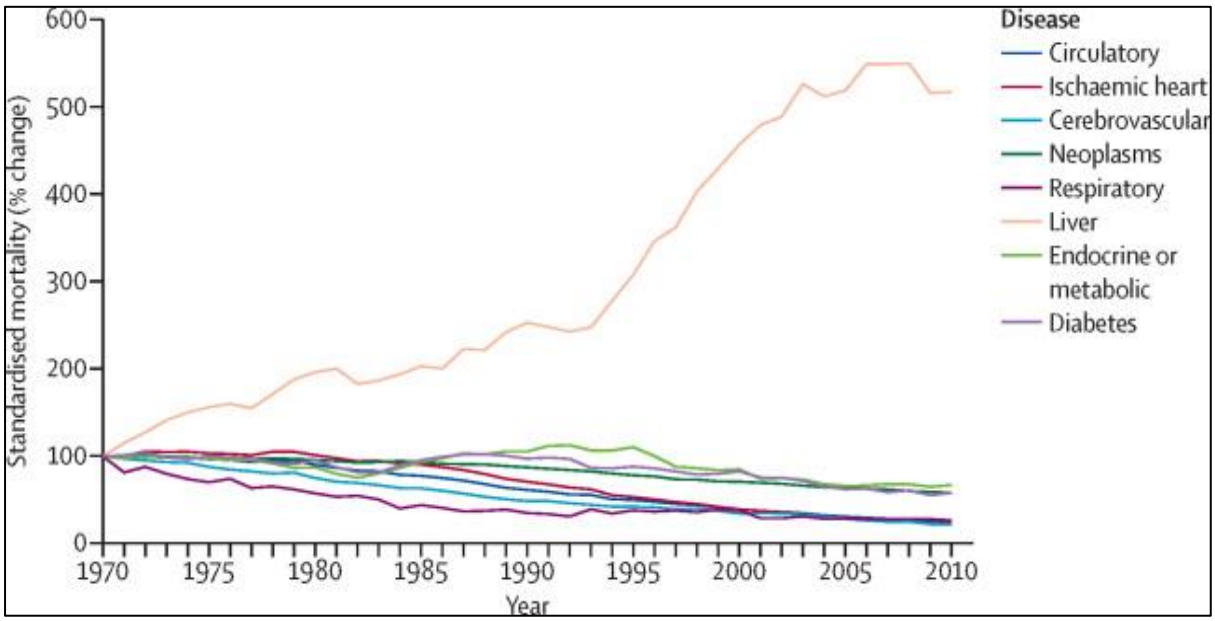


Figure 1.1 Standardised UK mortality rate data

Data were normalised to 100% in 1970, and subsequent trends plotted using the software Statistical Package for the Social Sciences. Data are from the WHO-HFA database[1]

Reproduced with permission from Williams et al. 2014.[2]

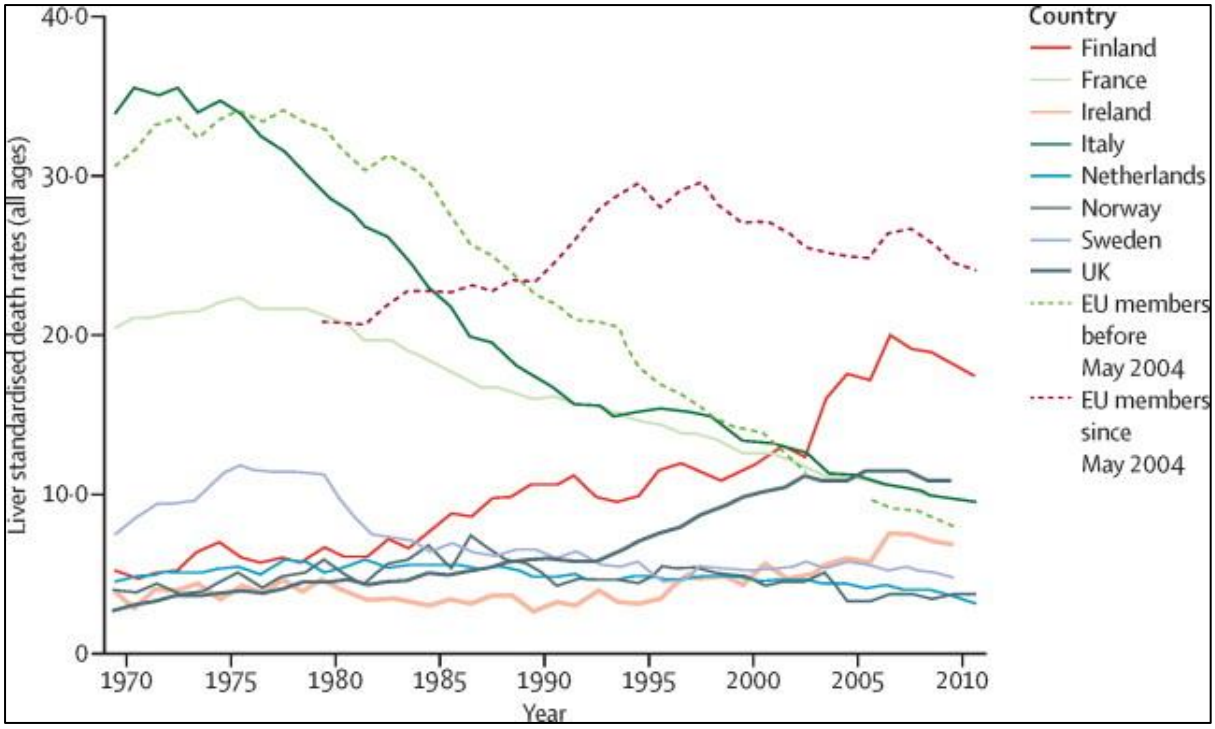


Figure 1.2 Standardised liver death rates in countries in the European Union before 2004

Decrease in liver mortality in selected European Union countries compared with UK mortality. Data are from the WHO-HFA database.[1]

Reproduced with permission from William et al. 2014.[2]

In the UK, the three main causes of chronic liver disease are alcohol, obesity related NAFLD and CHC.[2] In these conditions a toxic insult to the liver causes inflammation, resulting in fibrosis leading to cirrhosis, portal hypertension, liver failure, liver cancer and death.[3]

Non-Alcoholic Fatty Liver Disease

NAFLD is already the commonest cause of CLD worldwide [4, 5] and its prevalence is increasing (figure 1.3).[6] The prevalence of NAFLD is thought to be 25% within the general population.[5, 7, 8] Up to 10% of individuals with NAFLD can develop NASH that can lead to progressive hepatic fibrosis and cirrhosis.[9, 10] Although NAFLD normally follows the development of the metabolic syndrome and obesity, NAFLD can also affect between 5-8% of individuals who are not overweight or obese.[11]

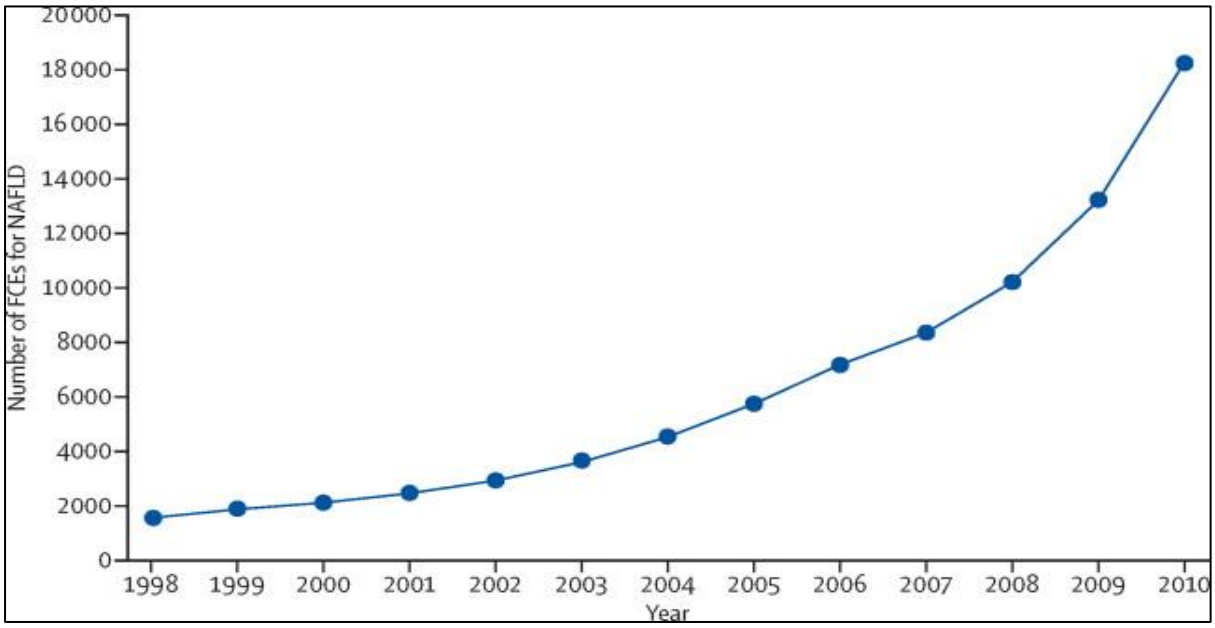


Figure 1.3 Number of hospital admissions for NAFLD 1998–2010

Admissions to hospital defined as first finished consultant episodes. Data are from Hospital Episode Statistics.[1] FCE=finished consultant episodes. NAFLD=non-alcoholic fatty liver disease.

Reproduced with permission from Williams et al. 2014.[2]

Chronic Hepatitis C

CHC is a global healthcare problem with a worldwide prevalence of over 200 million.[12] Within England, conservative estimates suggest the more than 160,000 individuals have CHC.[13, 14] However, the true prevalence may be greater as a large proportion of the CHC disease burden is concentrated in marginalised populations. Fifty-five to 85% of individuals who develop acute HCV infection will develop CHC. In contrast to NAFLD, the majority of individuals with CHC will develop progressive hepatic fibrosis during their lifetime[15] although progression to cirrhosis and its complications is identified in less than 30% of those infected.[16] Six major genotypes and their worldwide distribution have been recognised.

Mortality from Chronic Liver Diseases is attributable to Fibrosis and Cirrhosis

Fibrosis develops over many years but is asymptomatic; cirrhosis develops insidiously so that opportunities for interventions that might modify disease progression or result in a cure are missed. At present, most CLD is diagnosed at the stage of cirrhosis when interventions are ineffective, resulting in considerable morbidity and mortality.[17, 18] As a result, the index presentation of patients with CLD may be with the complications of cirrhosis itself (decompensation or acute on chronic liver failure) which include portal hypertension, ascites, encephalopathy, jaundice, and liver failure. Findings from a large UK study identified that the respective 1 and 5 year survival rates for patients diagnosed with cirrhosis in an outpatient setting were 0.84 and 0.66 and only 0.55 and 0.31 for those admitted to hospital respectively.[18] In addition, cirrhosis is a premalignant condition conferring an increased risk of hepatocellular cancer (incidence approximately 5% per annum).[19]

Progression of CLD to cirrhosis continues to place a considerable burden on health resources. Although liver transplantation can extend life expectancy for end-stage CLD, this requires suitable recipients to be identified and transplanted in a timely fashion. At present however, demand for liver transplantation continues to far exceed the supply of suitable organs. As a result, approximately 50% of patients accepted for liver transplantation die whilst awaiting organ allocation.[20] Earlier detection of CLD would permit the initiation of more timely treatment that could translate into improved survival with less cost to the health service. Preferably CLD should be detected early in its disease course to arrest progression to cirrhosis.

Available treatments can be categorised into those that improve the outcome of established cirrhosis and those that are specific to the aetiology of CLD.

Treatment of Cirrhosis

The major treatable complications of cirrhosis are portal hypertension (incidence approximately 8% per annum) and hepatocellular cancer (incidence approximately 5% per annum). Evidence-based guidelines of high methodological quality advocate endoscopic surveillance for varices and prophylactic intervention with beta-blockers (to reduce portal hypertension) or band ligation, prior to bleeding based on strong evidence of patient benefit in terms of morbidity and mortality and health-economic justification.[21] The case for surveillance and early detection of hepatocellular cancer in patients with cirrhosis has been evaluated in several observational studies and randomised controlled trials. Surveillance of cirrhotic patients for hepatocellular cancer has been shown to identify smaller tumours at a point where potentially curative therapies can be offered. However, unequivocal evidence of benefit for surveillance in terms of improved survival or health economic advantage is lacking but ethical considerations make randomised controlled trials of surveillance challenging.

Treatment for NAFLD

Whilst lifestyle interventions such as weight loss have been shown to arrest the progression of NAFLD, there are an increasing number of treatments for NASH in development.[22] However, current treatments are unable to arrest or reverse NAFLD and NASH in the majority of patients.

Treatment for CHC

The treatment of CHC has evolved significantly since the original discovery of the virus in 1989.[23] Treatment for CHC is now highly efficacious, abbreviated and tolerable.[24] **‘**Virologic cure’ is now possible for more than 95% of patients including those with advanced liver disease. However, modelling suggests that cases of cirrhosis related

to CHC will continue to rise at least until 2030 unless there is a dramatic increase in the rate of CHC treatment.[25]

Current use of non-invasive testing

Standard 'liver function tests' have poor sensitivity and specificity for the detection of liver fibrosis.[26, 27] The reference standard for staging liver fibrosis remains the histological staging of a liver biopsy specimen. Liver biopsy even in experienced centres is associated with complications such as pain (20%), serious morbidity (0.6%) and even death (0.01%).[28] Moreover, the accuracy and reliability liver biopsy is limited due to sampling error and both inter- and intra-observer variability. In addition, as liver biopsy is an invasive procedure with associated morbidity, there are ethical considerations that limit use of this modality for longitudinal disease assessment.

Given these limitations, a non-invasive test that can accurately stage liver disease is attractive is desirable. Over the past 2 decades numerous serum markers have been developed to stage liver disease severity. Serum based tests are cheaper, more reproducible and less invasive than liver biopsy, with minimal associated procedural morbidity. The term "indirect biomarkers" has been used to describe tests that are calculated using standard biochemical and haematology tests and demographic data. They can be calculated at minimal extra cost from tests and data that are routinely acquired. By contrast "direct biomarkers" are tests that measure molecules and enzymes involved in matrix biology. Many of these direct biomarkers incorporate special assays, are patented and incur additional cost in addition to routine biochemistry and haematology. The performance of biomarkers for the detection of advanced fibrosis and cirrhosis in CLD has generally been good or excellent; performance in the detection of mild and moderate fibrosis has not been as successful.

It has already been demonstrated that the widespread introduction of non-invasive testing for CLD, accompanied by effective treatment, such as has occurred in France can lead to an improvement in morbidity and mortality.[2]

Recently NICE Guidance number 49 has recommended the use of the ELF test for the detection of advanced fibrosis due to NAFLD.[29] However, there may be potential to further improve diagnostic performance combining ELF with other tests for fibrosis in NAFLD. Regardless, there has been less success in the development of a reliable non-invasive test diagnosis for NASH.

What changes are needed to reduce the morbidity and mortality related to CLD?

In CLD both morbidity and mortality are directly related to the severity of hepatic fibrosis. However, most individuals who are at risk of CLD will not develop progressive hepatic fibrosis and cirrhosis. As a result, the staging of liver fibrosis severity is essential to determine disease severity and to ascertain prognosis. The reference standard for assessing liver disease is not practicable for the detection of liver injury amongst individuals at risk of CLD. Preferably, the staging of liver disease should be performed using a safe, accurate and reproducible test. Ideally this test could be repeated at reasonable frequency to monitor disease progression or regression. Due to the need for multiple assessments, it is even more important that the test is safe, repeatable and reproducible when liver disease is monitored longitudinally perhaps in response to a specific anti-fibrotic or anti-inflammatory agent. Patients with evidence of advancing hepatic fibrosis or inflammatory states such as NASH should be identified early and offered treatment for their underlying liver disease aetiology. This may be in the form of lifestyle modification (for example weight loss for NAFLD) or targeted drug therapy (antiviral therapy in CHC). Patients who have also developed advanced hepatic fibrosis should be offered screening to reduce the risk of complications related to cirrhosis and portal hypertension. Furthermore, it is essential that non-invasive tests are employed for maximal clinical benefit. Given the ubiquity of CLD, non-invasive testing should ideally be employed in primary care to stratify patient referrals thus best utilising secondary and tertiary care resources.

Aims of thesis

This thesis aims to further investigate and explore the role of serum markers to improve the management of CLD by enhancing disease detection and helping refine the development of anti-fibrotic therapy:

a) Non-alcoholic Fatty Liver Disease

- *Non-invasive diagnosis of non-alcoholic steatohepatitis*

Whereas biomarkers have demonstrated good performance for the detection of advanced fibrosis in NAFLD, their performance in detecting lesser stages of fibrosis have been modest. Prior to the development of advanced fibrosis, the accurate detection of NASH is therefore an alternative way in which patients with progressive disease can be identified. Hitherto, a reliable non-invasive test for NASH has not been identified.

- *Direct head to head comparison of biomarkers of fibrosis*

Numerous biomarkers of fibrosis in NAFLD have been identified but few have been subjected to direct head to head comparisons. Indirect biomarkers are calculated from standard liver function tests and demographic data thus at minimal extra cost. By contrast, direct and hybrid biomarkers are often patented and incur extra cost. It is therefore essential to understand the differential performance of each type of biomarker.

- *Improvement of diagnostic performance by employing biomarkers in diagnostic algorithms*

In the absence of a perfect diagnostic test, the use of a biomarker at a diagnostic threshold will generate false negative and false positive results. Diagnostic strategies

that employ multiple biomarkers could potentially increase diagnostic accuracy and reduce the need for liver biopsy.

b) Chronic Hepatitis C

- *Direct head to head comparison of biomarkers of fibrosis*
- *Effect on direct biomarker diagnostic performance of using alternative assays for constituent components*

Again, numerous biomarkers of fibrosis in CHC have been identified but few have been subjected to direct head to head comparisons. It is therefore essential to understand the differential performance of indirect and direct biomarkers. Moreover, as the formulae for biomarkers are published, it is essential to understand whether using alternative component assays will result in altered diagnostic performance.

c) Monitoring the effects of putative antifibrotic therapy as assessed by:

- *Histology*
- *Serum markers*
- *Clinical outcomes*

Serum markers have been developed for the cross-sectional detection of fibrosis. Confirmation that serum markers can longitudinally monitor changes in fibrosis occurring during anti-fibrotic therapy is essential for their use as a modality to assess fibrosis evolution.

CHAPTER 2

LIVER INJURY, INFLAMMATION AND FIBROSIS IN CHRONIC LIVER DISEASES INCLUDING NAFLD AND CHRONIC HEPATITIS C

ABSTRACT

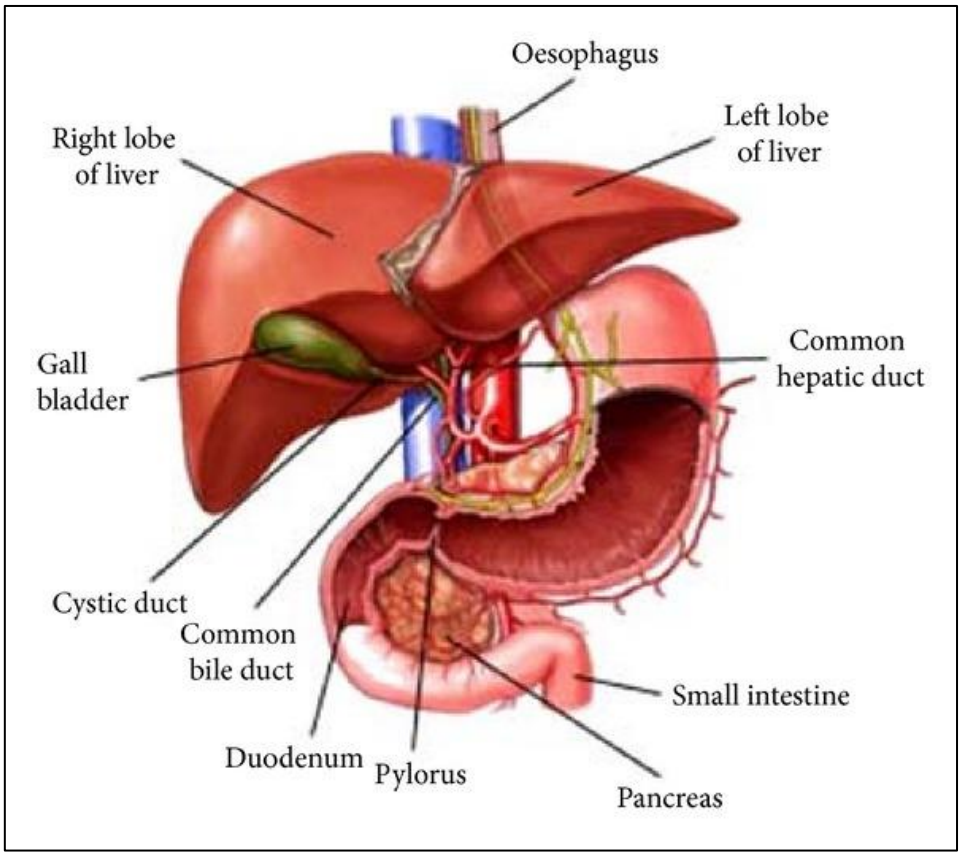
Injury to the liver, the largest solid organ in the body, leads to a cascade of inflammatory events. Chronic inflammation leads to the activation of Hepatic Stellate Cells (HSC) that undergo trans-differentiation to become myofibroblasts, the main ECM producing cells in the liver; over time increased ECM production results in the formation of liver fibrosis. Although fibrogenesis may be viewed as having evolved as a “wound healing” process that preserves tissue integrity, sustained chronic fibrosis can become pathogenic culminating in CLD, cirrhosis and its associated complications. In contrast to the reference standard liver biopsy, the diagnostic assessment of CLD by non-invasive testing is attractive. However, a comprehensive understanding of the pathophysiology of chronic liver disease is required to identify appropriate and meaningful diagnostic targets for clinical practice. Accordingly, in this chapter the mechanisms by which liver inflammation and fibrosis develop in chronic liver diseases are explored. Due to differing disease prevalence and treatment efficacy, disease specific diagnostic targets are required to optimally manage individual CLDs such as NAFLD and chronic hepatitis C infection. To facilitate this, a review of the pathogenesis of both conditions is also conducted. Finally, the evidence for hepatic fibrosis regression and the mechanisms by which this occurs are discussed, including the current use of antifibrotic therapy.

Structure and Function of the Liver

The liver is the largest solid organ in the body and has a median weight of 1.6kg and 1.4kg in adult males and females respectively.[30] The liver receives 75% of its blood supply via the portal vein and 25% via the hepatic arterial system. The portal vein carries blood from the entire capillary system of the digestive tract, spleen, pancreas and gallbladder. The hepatic artery is the second major branch of the celiac axis. The venous drainage of the liver is via the hepatic veins which open into the superior vena cava (figure 2.1). The liver can be divided into eight functional segments or “lobules” based upon blood supply and biliary drainage. Hepatic lobules are comprised of a central hepatic vein and peripheral portal tracts that contain the final tributaries of the bile ducts (bile ductule), portal vein (portal venule) and hepatic vein (hepatic venule). Blood is drained from the portal tracts to the central vein by specialised capillaries known as the hepatic sinusoids.[31]

Figure 2.1 Anatomy of the liver and its macroscopic relationship to the intestinal tract and vasculature

Reproduced with permission (Creative Commons Attribution License) from Ebaid et al. 2013.[32]



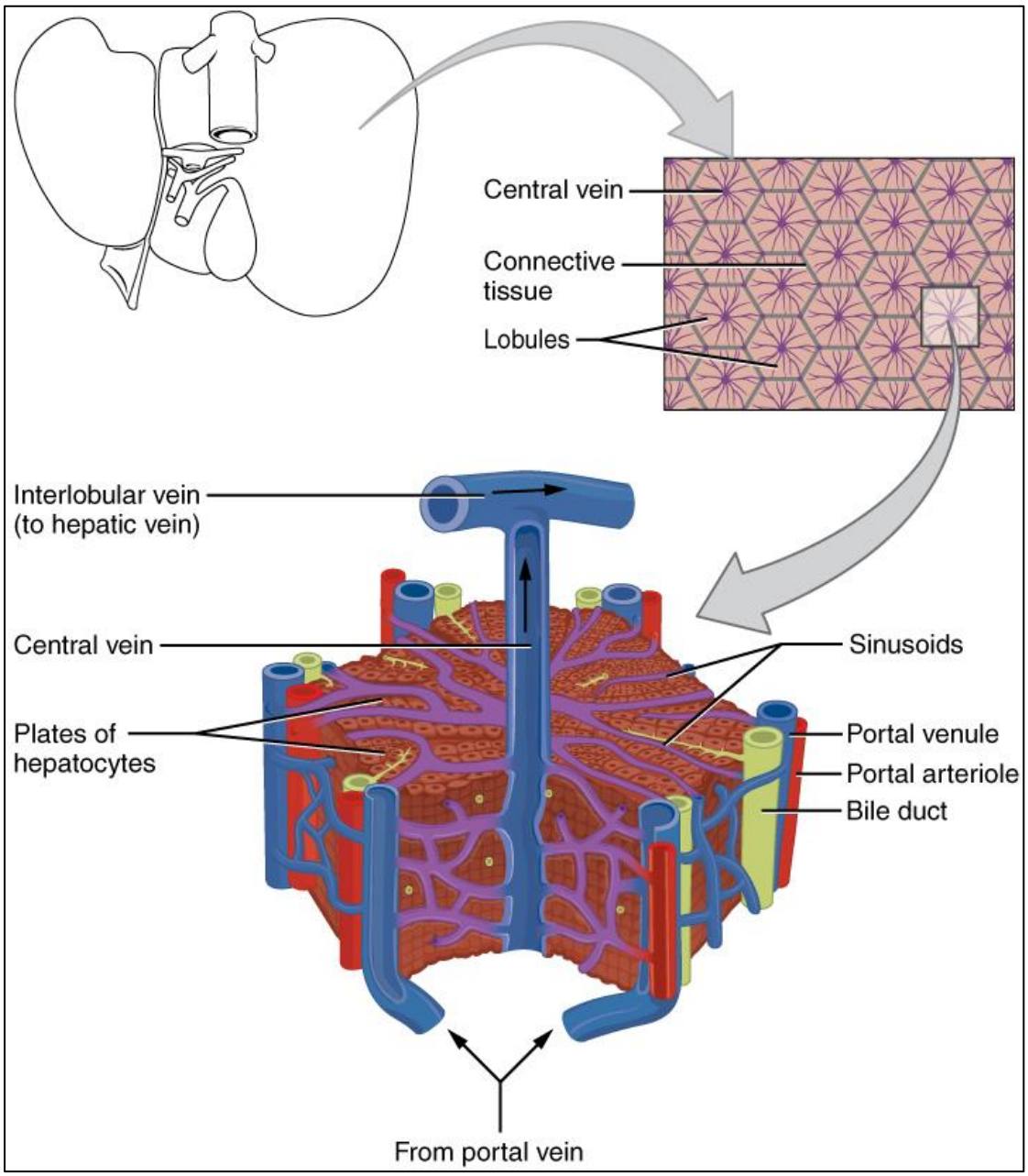
The sinusoids are lined by a fenestrated endothelial layer containing numerous microvilli.[31] This structural organisation facilitates exchange of solutes between the portal tracts and the hepatocytes through the space of Disse. Endothelial cells, Kupffer cells and hepatic stellate cells lie in juxtaposition with the hepatic sinusoid (figures 2.2 and 2.3). Kupffer cells are the resident macrophages of the liver and their major functions include the clearance of particles, immune complexes, senescent red blood cells and endotoxins. In addition, Kupffer cells have a role in the innate immune response and produce pro-inflammatory cytokines including interleukin 1 and 6, tumour necrosis factor-α (TNF-α) and interferons. Hepatic stellate cells are distributed throughout the liver and form the main perisinusoidal cell type with a diverse range of functions.

Hepatic Extracellular Matrix

The hepatic extra-cellular matrix (ECM) is the array of macromolecules that forms the liver 'scaffolding'.[33] In the normal liver, ECM contributes to approximately 0.5% of the total weight of the liver, comprising less than 3% of the area on cross sectional imaging.[33] Normal ECM is composed of collagens (types I, III, IV, V, VI, XIV and XVIII), elastin, structural glycoproteins (laminin, fibronectin, nidogen/enactin, tenascin, osteopontin, various acidic proteins), proteoglycans (heparan sulfate, syndecan, biglycan and decorin), and hyaluronic acid (a glycosaminoglycan).[34] All 3 of the cell types (hepatocytes, endothelial cells, HSC) that surround the space of Disse produce matrix components.

Figure 2.2 Schematic diagram representing the relationship of the macroscopic structure of the liver with the functional hepatic lobule with hepatic venules (blue), hepatic arteriole (red), bile ductules (yellow)

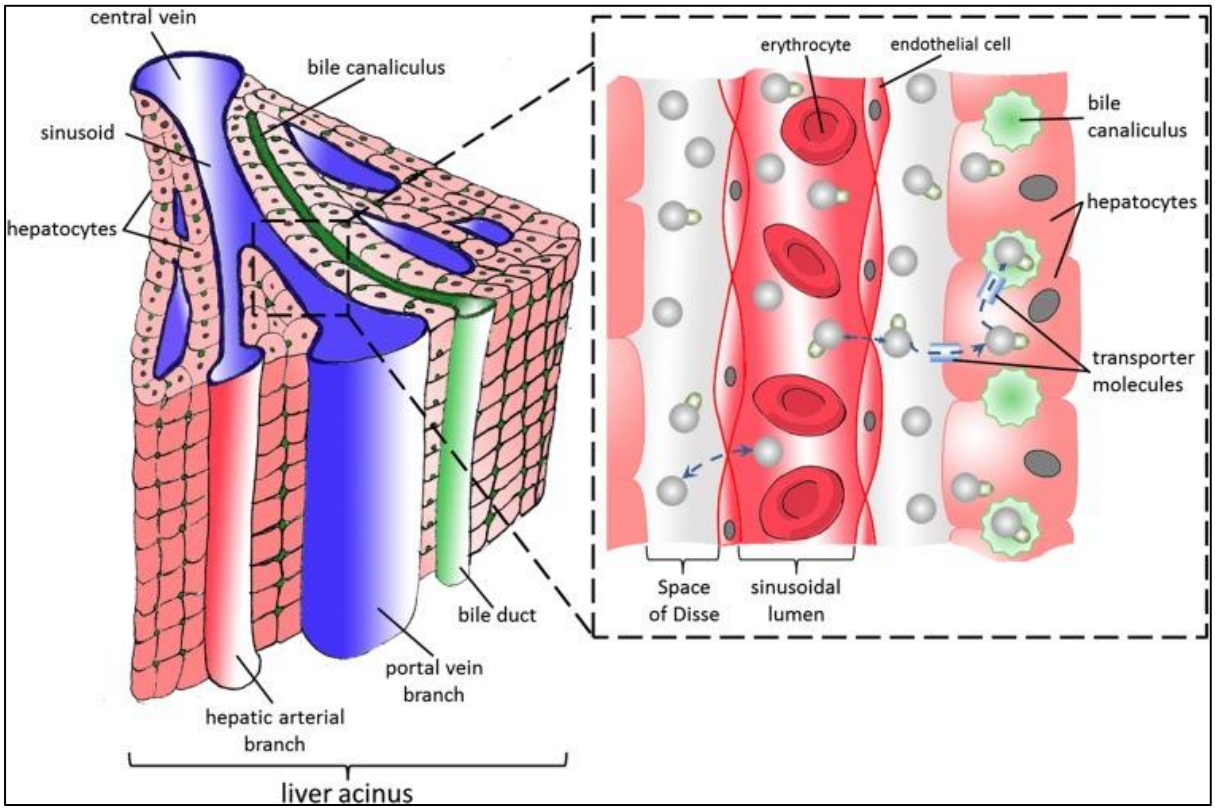
Reproduced with permission (Creative Commons Attribution License) from Anatomy & Physiology text book.[35]



The liver has a number functions which in broad terms can be defined as ‘the regulation of the concentrations of solutes in the blood that affect the function of other organs’.[30] Through the uptake, metabolism and secretion of solutes, the liver performs an integral role in the metabolism of amino acids (for example transamination), carbohydrates (for example gluconeogenesis), lipids (for example lipid production), haemoglobin, bile salts, iron, copper, vitamins, ammonia and drugs. The liver is the major synthetic organ producing albumin, serum binding proteins (for example haptoglobin) and clotting factors. Furthermore, the liver is an important immunological site with functions such as cytokine signalling, antigen surveillance and immune tolerance.

Figure 2.3 Schematic diagram representing functional hepatic acinus with hepatic venules (blue), hepatic arteriole (red), bile ductules (green) together with the relationship to the Space of Disse and the sinusoidal lumen

Reproduced with permission (Creative Commons Attribution License) from Chauhan MD et al. 2016.[36]



Hepatic Inflammation in Response to Chronic Injury

Whereas immune responses in CLD may contribute to the restoration of tissue function they may also lead to tissue injury. An overactive or exaggerated immune response, for example in NASH or CHC, can result in organ dysfunction by the replacement of hepatic parenchyma by scar tissue and by vascular architectural distortion. These immune responses are the subject of active investigation to not only extend understanding of the pathology of liver injury but to also identify diagnostic and therapeutic targets.

Innate immune response

Pattern recognition receptors (PRRs) are an essential part of the innate immune system that facilitate the detection of pathogens. The liver is enriched with both parenchymal liver cells and non-parenchymal liver cells that express PRRs. Hepatocytes express PRRs and are the main liver parenchymal cells. Non-parenchymal cells that express PRRs include liver sinusoidal endothelial cells (LSEC), HSC and bone marrow derived immune cells (Kupffer cells, and dendritic cells).[37] PRR can be divided into 4 classes: Toll-like receptors (TLR), Nucleotide oligomerisation receptors (NLR), C-like lectin receptors (CLR) and RIG-1 like receptors (RLR). PRR recognise conserved molecular structures called pathogen-associated molecular patterns (PAMPs) that are pathogen specific. Of the 4 types of PRR, TLRs are the key sensor of the innate immune system for the recognition of pathogens[38] including viruses.[39] For example, TLR-3 is able to recognise viral dsRNA leading to the production of type 1 Interferons (IFN α or β).[40] Three main pathways are activated by TLRs: mitogen activated protein (MAP) kinase pathway (ERK, p38 and JNK), nuclear

factor κB (NFκB)pathway and interferon regulatory transcription factor (IRF) pathway. Pro-inflammatory and anti-inflammatory cytokines can be induced following TLR stimulation. Inflammation is characterized by both the activation of innate immune cells and the production of pro-inflammatory cytokines such as IL-1α, IL-1β, and TNFα.

Innate immune cells recruited to sites hepatic inflammation

- Neutrophils (also known as neutrophilic granulocytes or polymorphonuclear leukocytes (PMNs)) infiltrate the site of injury within minutes in response to the injury itself following the release of damage-associated molecular patterns (DAMPS)[41] from damaged cells. Neutrophil recruitment peaks within hours.[42] Neutrophils are key effectors of the innate immune system with high phagocytic potential and possess a vast number of antimicrobial molecules. In addition to clearing pathogens, neutrophils may also exacerbate macrophage cytotoxicity and help promote a chronic inflammatory state.[43]
- Monocyte infiltration following tissue injury peaks between 24-48 hours.[44] Recruited monocytes demonstrate a variety of functions including production of inflammatory mediators, clearance of neutrophils, ECM production and angiogenesis. In humans, 3 types of monocyte subsets have been identified: classic ($CD14^{++}CD16^{-}$), intermediate ($CD14^{++}CD16^{+}$), and non-classic ($CD14^{++}CD16^{++}$).[45] Furthermore, as derived from murine studies, monocytes can also be categorised into inflammatory ($CCR2^{hi}$, CX_3CR1^{lo}) and anti-inflammatory (CX_3CR1^{hi}, $CCR2^{lo}$) subtypes. Pro-inflammatory monocytes produce inflammatory cytokines and chemokines, produce proteases and facilitate clearance. Anti-inflammatory monocytes produce anti-inflammatory cytokines (IL-10, transforming growth factor β and VEGF) that promote resolution and restitution.[46]

- Kupffer cells are the resident macrophages of the liver.[47] In addition to detecting tissue injury, Kupffer cells can recruit inflammatory cells, promote tissue repair and remodelling. In addition, new populations of macrophages can be recruited following acute inflammation (emergency repopulation). Macrophages may be proinflammatory (classic M1 type) or anti-inflammatory (alternative M2 type) and their relative balance contributes to either injury or repair.[48]
- Platelets are metabolically and synthetically active cells that can produce many inflammatory cytokines and chemokines thus attracting neutrophils to sites of inflammation.[49]
- Natural killer T cells (NKT) and natural killer (NK) cells (granular lymphocytes) can be thought of as being part of the innate immune system as they can kill target cells without priming.[50] NK cells can contribute to HSC clearance in a TNF-related apoptosis inducing ligand (TRAIL) dependent manner.[51]

Adaptive immune response to hepatic injury

Both T (thymus) and B (bone marrow derived) lymphocytes mediate the adaptive immune response. Whereas T cells are involved in cell-mediated immunity, B cells are primarily responsible for humoral immunity. Regardless, the function of both T and B cells is to recognise ‘non self’ antigens by generating specific responses to eliminate specific pathogens or cells expressing ‘non self’ antigens.[52]

T helper (CD4+) cells can activate other immune cells (including B cells), switch antibody classes, activate cytotoxic (CD8+) T cells and enhance macrophage phagocytosis. In response to cytokine stimulation, T helper cells can either assume a proinflammatory phenotype (Th1) or an anti-inflammatory phenotype (Th2). Th1 T cells secrete the inflammatory cytokines interferon-γ and TGF-β. Th2 T cells are

characterised by the secretion of IL-4, IL-5 and IL-10. Proinflammatory M1 macrophages are induced by the release of interferon-γ by Th1 T cells.

B cells secrete antibodies which provides humoral immunity within the adaptive immune system. In addition, B cells secrete cytokines and function as antigen presenting cells. All B cells express B cell receptors (BCRs) on their cell membrane. BCRs allow B cells to bind specific antigens which is followed by the production of antibodies. When naïve or memory B cells are activated by antigen, they proliferate and differentiate into effector B cells with the formation of large plasma cell representing the end stage of their maturation pathway.

Cell death and ongoing inflammatory response

Cell death can be categorised in several ways including non-inflammatory cell death (apoptosis) and inflammatory cell death (necrosis, pyroptosis and necroptosis). Regardless, the nature of the accompanying immune response is dependent on the signals liberated from dead cells. Apoptosis fails to generate an immune response as apoptotic cells retain membrane integrity. In addition, apoptotic cells also release factors that inhibit the recruitment and activation of neutrophils. By contrast, membrane integrity is disrupted during necrosis which results in the release of DAMPS.[41] Broadly, DAMPS are molecules that can activate inflammation. Necrosis also produces inflammatory signals by modifying extracellular matrix components including hyaluronic acid.[53] Moreover, necrosis can also result from programmed cell death pathways: pyroptosis and necroptosis. Pyroptosis is initiated by inflammasome mediated activation of capsase-1 which results in lytic cell death and the production of the interleukins (IL) IL-1β and IL-18.[54] Necroptosis is activated by the presentation of

‘external death signals’ to cells such as tumour necrosis factor, interferon receptor and selected TLR pathways.[55] Following activation by these signals, necroptosis involves inhibition of caspase-8 (promotes apoptosis), mixed-lineage kinase domain-like protein (MLKL) and RIP kinase family members (RIPK1 and RIPK3).

Inflammasome

Inflammasomes are an intracellular multiprotein scaffolding that are expressed in both parenchymal and non-parenchymal cells of the liver. Functionally, inflammasomes are both sensors and receptors of the innate immune system that can induce inflammation in response to pathogens and molecules derived from host proteins (DAMPs). In response to these ‘cellular danger signals’ inflammasomes activate caspase-1 and release both IL-1β and IL-18.[56]

Pathophysiology of Liver Fibrosis

Whereas the development of liver fibrosis itself is often a prerequisite for the morbidity associated with CLD it is important to highlight that fibrogenesis is also a part of the normal wound healing process in response to noxious stimuli.

Liver fibrosis is a structural change defined by an accumulation of ECM proteins such as collagen often triggered by chronic, sustained inflammation. The pattern of fibrosis deposition is dependent on the aetiology of liver injury. Pericellular and perisinusoidal fibrosis deposition in the centrilobular areas are characteristic of fibrosis related to NASH and alcohol related liver disease. By contrast, periportal fibrosis deposition is characteristic of autoimmune and viral liver diseases.

Regardless, it is important to recognise that maintenance of the liver matrix is a dynamic process in which deposition and resorption of matrix are balanced. In a steady physiologic state, both ongoing ECM deposition and removal are equal resulting in an unchanged amount of ECM.[57] At a cellular level, the HSC is recognised to be the most important cell lineage in the development of liver fibrosis. In its quiescent or resting state, the HSC is a lipid storing cell representing the body's major location of Vitamin A. However, on activation the HSC undergoes transformation to become a myofibroblast capable of regulating matrix deposition and resorption whilst also possessing contractile properties. In addition to those derived from HSC, profibrogenic myofibroblasts can also be derived from portal fibroblasts, recruited bone marrow cells and epithelial cells (for example hepatocytes and cholangiocytes) that have undergone epithelial to mesenchymal transition.[58] HSC reside in the space of Disse and are the resident non-paranchymal cell type. Their embryonic origin is likely to be mesenchymal given that they produce α-smooth muscle actin when activated in addition to vimentin and desmin.[59] HSC themselves represent approximately 10% of

the total liver cell number and 1.5% of the total liver cell volume.[60] The location of HSC within the space of Disse allows for direct (within 140μm) contact of the HSC with their other cell types including hepatocytes, endothelial cells and Kupfer cells[60] thus facilitating the intercellular transport of soluble mediators and cytokines. In addition to this intimate location within the space of Disse, intercellular communication between HSC and the neighbouring cells are also facilitated by their prominent dendritic cytoplasmic processes. Their contractility may contribute to the regulation of portal pressure.[61] When quiescent, the HSC serve as a reservoir for retinol (a precursor for vitamin A) and other lipid soluble compounds. Indeed, the presence of Vitamin A esters in cytoplasmic perinuclear lipid droplets is the characteristic microscopic feature of quiescent HSC. Hepatic stellate cells represent resting profibrogenic myofibroblasts which are a major constituent of the extracellular matrix in both the healthy and diseased liver. The phenotypic transformation of hepatic stellate cells into profibrogenic myofibroblasts is associated with the acquisition of α-smooth muscle reactivity. The activation of HSC is comprised of 2 stages: initiation and perpetuation.

HSC Activation

In response to liver injury, lipid peroxides and apoptotic bodies that accumulate in damaged hepatocytes initiate HSC activation in a process mediated by Fas and TRAIL.[62] This activation process is initiated by profibrogenic cytokines (for example TGF-β), fibronectin, PDGF, ROS and apoptotic bodies derived from neighbouring cells, immune cells and platelets (figure 2.4).[57] Thereafter, HSC undergo characteristic phenotypic changes and resemble myofibroblasts.[58]

Figure 2.4 Matrix and cellular alteration in hepatic fibrosis.

Normal liver parenchyma contains epithelial cells (hepatocytes) and nonparenchymal cells: fenestrated sinusoidal endothelium, hepatic stellate cells (HSCs), and Kupffer cells (KCs).
(A) After injury, the stellate cells become activated and secrete large amounts of extracellular matrix (ECM).
(B) Deposition of ECM in the space of Disse leads to the loss of both endothelial fenestrations and hepatocyte microvilli.

Reproduced with permission from Hernandez-Gea V and Friedman SL 2011.[63]

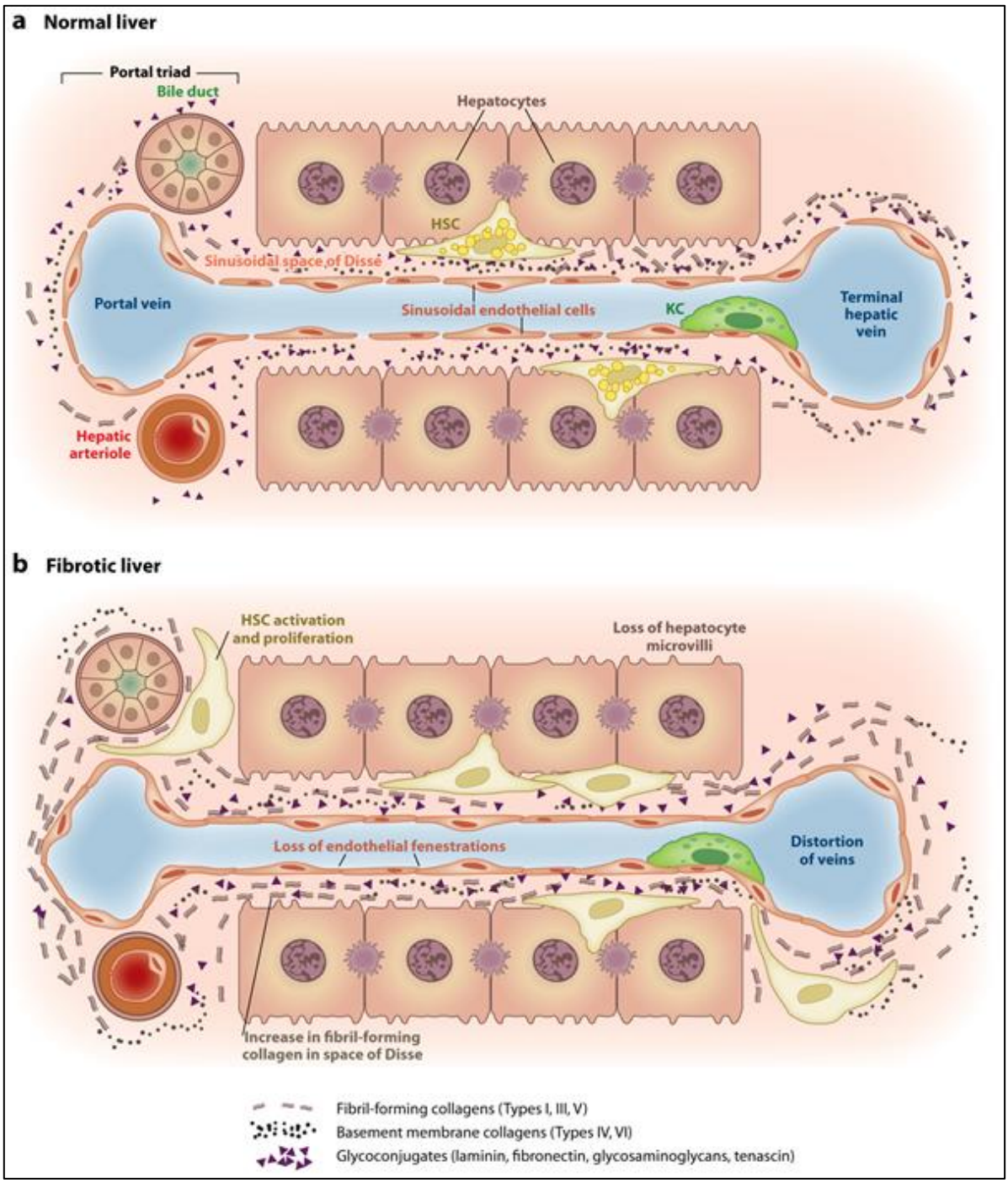
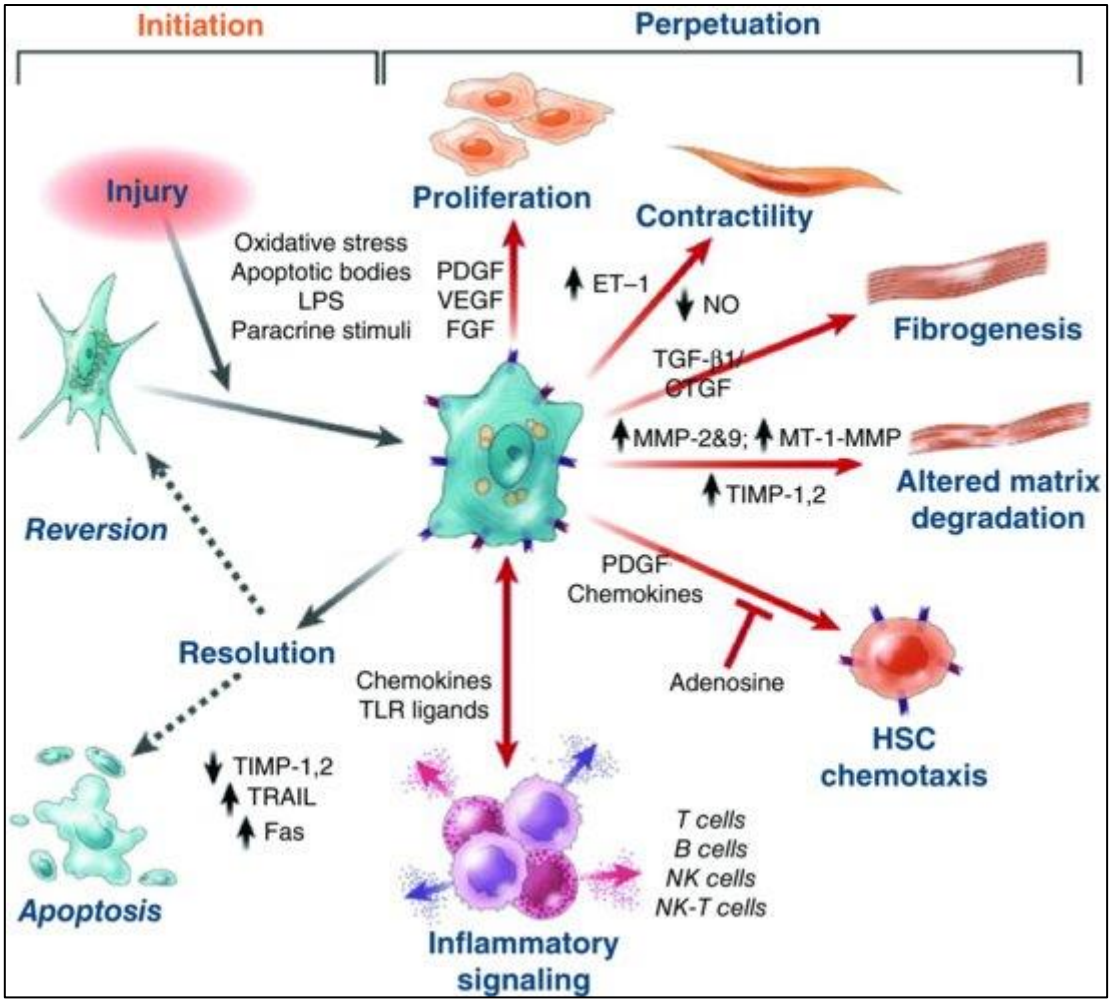


Figure 2.5 Pathways of hepatic stellate cell activation including those contributing to initiation and perpetuation.

Initiation is provoked by soluble stimuli that include oxidant stress signals, apoptotic bodies, lipopolysaccharide, and cytokine stimuli from neighbouring cells.
Perpetuation is characterised by specific phenotypic changes including proliferation, contractility, fibrogenesis, altered matrix degradation, chemotaxis, and cytokine signalling.

Reproduced with permission from Friedman SL 2008.[57]



HSC Perpetuation

Following on from activation, the activated HSC enters the phase of perpetuation in which ECM is accumulated resulting in scar tissue formation (figure 2.5). Tissue hypoxia, apoptosis and cell matrix interactions maintain ongoing HSC activation. HSC perpetuation is comprised of number of functional responses including proliferation, fibrogenesis, chemotaxis, contractility, matrix degradation, retinoid loss and cytokine expression.

- **Proliferation**

HSC proliferate rapidly and gain a profibrogenic phenotype; this occurs primarily in response to both an increase in PDGF and PDGF responsivity.[64] Other mediators involved in HSC proliferation include TNFα, VEGF, thrombin and EGF.[57]

- **Fibrogenesis**

An accumulation of ECM (particularly collagen type I) occurs follows on from increased synthesis by activated myofibroblasts and decreased degradation. TGF-β1 is the main driver for ECM production by activated HSCs. Connective tissue growth factor (CTGF) also acts as a profibrogenic cytokine towards HSC.

- **Chemotaxis**

HSCs migrate towards chemokines allowing cells to organise within regions of injury. Examples of chemokines toward which HSCs migrate to include PDGF, VEGF, Ang-1, TGF-β1, EGF, b-FGF, CCL2, CXR3 and CXR4.[65] In addition, tissue hypoxia enhances HSC migration with ROS activating extracellular signal-regulated kinase 1/2 (ERK) and JNK1/2 pathways.

- **Contraction**

Hepatic sinusoidal remodelling occurs following HSC activation which is mediated by collagen matrix deposition, loss of fenestration and an increase in the number of

contractile HSCs.[66] These events contribute in an increase in sinusoidal resistance; in the context of advanced fibrosis this can contribute to portal hypertension. Other factors that stimulate HSC contraction include nitrous oxide deficiency, ET-1, angiotensinogen II, eicosanoids, atrial naturetic peptide and somatostatin.

- **Retinoid Loss**

HSC activation is characterised by the loss of perinuclear retinoid droplets.[67] As described earlier HSC are the largest reserve of retinoids in the body and conversion of retinol into retinyl ester is a characteristic feature of HSC activation

- **Matrix Degradation**

Matrix metalloproteinases (MMPs) are the are the main enzymes responsible for ECM degradation; the MMP activity is in turn regulated by tissue inhibitors of metalloproteinases (TIMPs). Both MMPs and TIMPs are produced by several liver cell populations including Kupffer cells, myofibroblasts and hepatocytes.[68] However, TIMPs are predominantly expressed by activated HSCs. The A Disintegrin and Metallopoteinase-domain proteins (ADAMs) form another stimulus for stellate cell collagen production via TGF-β activation.[69]

- **Cytokine Expression**

HSCs express a vast array of chemokines that have been identified to recruit neutrophils, macrophages, NK/NKT cells, dendritic cells and T cells. As a result, HSCs play an important role in immune cell infiltration.

Progression of fibrosis in chronic liver diseases to cirrhosis

Cirrhosis can be defined as 'the histological development of regenerative nodules surrounded by fibrous bands in response to chronic liver injury that leads to portal hypertension and end stage liver disease'.[3] Cirrhosis is a consequence of long standing excessive fibrogenesis resulting in encapsulation and/or replacement of injured liver parenchyma by a collagenous scar. Histologically, cirrhosis is characterized by fibrotic septa that link portal tracts with each other and central vein. This produces liver parenchyma that is composed of hepatocyte islands that are surrounded by fibrotic septa and devoid of a central vein. These changes result in an increase in intravascular resistance within the portal venous system and decreased hepatic perfusion. In addition, the development of cirrhosis confers a significant increase in the risk of developing hepatocellular carcinoma (incidence up to 30% over a 5 year period).[3] Progression of fibrosis to cirrhosis is variable and is dependent on the cause of liver disease, environmental and host factors.[3] Cirrhosis is frequently asymptomatic and unsuspected until complications of liver disease present with include variceal bleeding, ascites, spontaneous bacterial peritonitis, encephalopathy and death.

Exploring the Pathogenesis of Inflammation and Fibrosis In NAFLD

The pathogenesis of NASH is complex involving hepatic parenchymal and non-parenchymal cells together with immune cells. With regard to pathogenesis, a “two hit” hypothesis was first proposed in 1998[70]; this has subsequently been modified into “three hit”[71] and “multiple hit” hypotheses.[72] In the original “two hit” hypothesis, the development of insulin resistance results in excessive lipid accumulation within hepatocytes (the first hit). The first hit is followed by lipotoxic metabolite-induced mitochondrial dysfunction, oxidative stress and endoplasmic reticular (ER) stress which leads to hepatocyte death (the second hit). Under normal physiological circumstances, existing hepatocyte replication results in the replacement of dead hepatocytes. In NASH however, it is believed that progenitor cell replication is enhanced as hepatocyte replication is impaired. Although progenitor cell proliferation results in the replacement of dead hepatocytes it also results in hepatic stellate cell activation and fibrogenesis (the third hit). The “multiple hit” hypothesis describes further additional insults derived from other sites such as the GI tract (such as gut-derived endotoxins due to impaired gut permeability and stasis) and adipose tissue (adipokines). Moreover, the “multiple hit” hypothesis allows one to understand that the development of NASH is the result of a complex interplay between factors such as the genetic variation in immune balance and the influence of additional aetiologies such as alcohol consumption and obesity. Whereas several pathways (including direct lipotoxicity, inflammasome activation, toll-like receptor signalling, hedgehog signalling) have all been implicated in the pathogenesis of cellular inflammation in NASH, the final pathway appears to be the development of a profibrotic state.

Epigenetic and Genetic Regulation in NASH

Recent genome-wide association studies (GWAS) have identified several genes that confer an increased risk of NASH amongst individuals with NAFLD.[73] These include patatin-like phospholipase domain containing 3 (PNPLA3), transmembrane 6 superfamily member 2 (TM6SF2), farnesyl diphosphate farnesyl transferase 1 (FDFT1), EF-hand calcium binding domain 4B (EFCAB4B) and glucokinase receptor (GCKR). Of these identified genes, the most widely studied is the PNPLA3 gene located on chromosome 22. PNPLA3 gene encodes a 481 amino acid protein that mediates triacylglycerol hydrolysis. In studies of Huh7 hepatoma cell line cells, PNPLA3 variant p.I148M (rs738409 substitution of cytosine to guanine) was associated with reduced enzymatic activity of emulsified triglycerides following hydrolysis. Within clinical studies, the PNPLA3 variant p.I148M has also been found to be associated with a high risk of NASH in adult[74], paediatric[75] and also lean patients.[11]

In addition, GWAS have identified that epigenetics (a reversible phenomenon affecting gene expression) also appears to contribute to the pathogenesis of NASH. A study of rat liver tissues and high fat emulsion induced fatty liver identified that the expression of genes involved in apoptosis, biosynthesis and inflammation increase in NASH.[76] By contrast, these studies have revealed down-regulation of expression of genes involved in DNA damage response signal transduction, cholesterol biosynthesis and carbohydrate metabolism.

Lipotoxic Hepatocyte Injury

The accumulation of excess free fatty acids (FFA) in hepatocytes results in the generation of toxic lipid metabolites. This process, lipotoxicity, is thought to be fundamental in the development of NASH. Following exposure to these toxic metabolites, injured hepatocytes appear enlarged and swollen in a process termed ballooning. In addition to steatosis and lobular inflammation, the presence of ballooning is an essential diagnostic criterion for the presence of histologic NASH.[77]

The source of excess FFA is multifactorial and excess dietary intake is certainly a contributing factor. In addition, an increase in intrahepatic FFA can occur because of *de novo* lipogenesis, adipose lipolysis and impaired FFA oxidation. Under normal circumstances, hepatocytes store FFA as triglycerides and it is postulated that the conversion of FFA into triglycerides may be protective against lipotoxicity.[78] However, in the context of FFA excess, the conversion to tryglycerides becomes saturated and alternative highly toxic lipid metabolites are formed: ceramides, diacylglycerols and lysophosphatidylcholine and oxidised cholesterol metabolites.[79] These toxic metabolites cause liver injury through the overproduction of reactive oxygen species (ROS).

Two major mechanisms of oxidative stress in NASH have been identified. The first, direct cell injury, occurs when there is an imbalance between pro-oxidants and antioxidants. The second, indirect cell injury, occurs when damaging cellular pathways are activated such as those involving NF-κB. Increased production of ROS induces activation of NF-κB which regulates the production of pro-inflammatory cytokines such as interleukin-1β (IL-1β), tumour necrosis α (TNFα) and interleukin-6 (IL-6).

ER stress is activated by FFA induced oxidative stress and appears to be an important mechanism in the development of NASH.[80] Prolonged and severe ER stress can lead

to cell death. Studies have identified that ER stress markers are elevated in NASH[81], and that ER stress can activate inflammatory pathways such as Jun-(N)-terminal Kinase (JNK) and NF-κB.[82]

Inflammatory and immune mediators in NAFLD

Analogous to other liver diseases, a variety of inflammatory and immunologic mechanisms contribute to NASH and NAFLD progression. These include innate immunity (neutrophils, macrophages, NK cells and NK T cells), adaptive immunity (T and B cells), inflammasome activation and the gut-liver axis.[83]

Gut microbiota and macrophages in the pathogenesis of NASH

As described above, the liver is exposed to low level endotoxaemia and antigenaemia via the portal vein. Thereafter, Kupffer cells are the principle cell type responsible for antigen and endotoxin clearance thus maintaining immune tolerance and homeostasis. This balance can, however, be impaired in the context of changes to gut flora, gut permeability and Kupffer cell responsivity. Studies have identified evidence for gut dysbiosis as a potential factor in the aetiology of NAFLD. Lower levels of *Bateriodetes* together with higher levels of *Prevotella* and *Porphoyromonas* have been identified in patients with NAFLD as compared to healthy controls.[84] Moreover, in mouse models, hepatic steatosis and inflammation driven by inflammasome mediated dysbiosis were associated with enhanced hepatic TNF-α expression.[85] In another study examining differences in gut microbiota between patients health subjects, obese patients and those with NASH, marked differences were identified in the composition of gut flora in either obese (without NASH) or obese (with NASH) patients as compared with healthy subjects.[86] Whereas the gut composition of patients with obese and NASH patients were more similar, there were significant differences in the concentrations of

Proteobacteria, *Enterobacteriaceae*, and *Escherichia* species. Interestingly, whilst similar blood-ethanol concentrations were observed between healthy subjects and obese non-NASH patients, obese NASH patients exhibited significantly elevated blood ethanol levels. It was therefore hypothesised that elevated ethanol blood levels in patients with NASH may be the result of alcohol producing gut bacteria.

Macrophages may be proinflammatory (classic M1 type) or anti-inflammatory (alternative M2 type) and their relative balance contributes to either injury or repair.[48] In both mouse models and human subjects, an increase in M2 Kupffer cells was found to promote M1 Kupffer cell apoptosis which in turn was found to inhibit progression of NAFLD.[87] Toll-like receptor activation (particularly TLR4) is associated with macrophage mediated inflammation in NASH[88] and is associated with the release of IL-1β, TNFα and IL-6.[89] Studies have identified that both TLR4 inhibition and macrophage depletion reduces NAFLD progression in both animal and human liver biopsy studies.[90, 91] In addition, liver macrophages transform into 'foam cells' in NASH by internalising oxidised low-density lipoprotein (ox-LDL) resulting in storage of cholesterol and cholesterol crystals in enlarged lysosomes.[92] Studies in mice have revealed that inhibition of ox-LDL recognition and ox-LDL uptake into foam cells is associated with reduced hepatic histologic progression.[93]

Neutrophils in the pathogenesis of NASH

A number have studies have identified a possible role of neutrophils in the progression of NAFLD and NASH. Neutrophil infiltration is a common histologic finding in patients

with NASH with neutrophils frequently surrounded steatotic hepatocytes, resembling the crown-like structures found in obese adipose tissue.[94] Moreover, patients with NASH and also advanced fibrosis related to NASH have been found to have a higher neutrophil-to-lymphocyte ratio than patients without either NASH or advanced fibrosis.[95] Human neutrophil peptides (HNPs) are proteins produced by neutrophils that induce cytokine and chemokine production under inflammatory conditions. Mouse studies have identified that HNPs can enhance hepatic fibrosis in fatty liver by inducing hepatic stellate cell proliferation.[96] In addition, the deletion of elastase (a protease secreted by neutrophils) in high fat diet induced obese mice was found to improve hepatic histologic inflammation with reduced of neutrophil and macrophage infiltration.[97]

T and B lymphocytes in the pathogenesis of NASH

Recent studies have confirmed that the innate immune system appears to play an important role in the pathogenesis of NASH. Analysis of liver histology obtained from patients with NASH has identified that the inflammatory infiltrate in NASH is heavily enriched with both lymphocytes and macrophages.[98] Moreover, the inflammatory infiltrate at the portal tracts is composed predominantly of CD8+ lymphocytes.[99] Regulatory T cells play a critical role in regulating inflammatory processes in NASH. Th17 T cells functionally oppose regulatory T cells and produce IL-17. Elevated levels of IL-17 have been identified in patients with obesity.[100] In mouse studies, the neutralisation of IL-17 was found to reduce the severity of NASH, implicating IL-17 as a mediator of NASH and identifying this as a potential therapeutic target.[101, 102]

More recently, B cells have emerged as additional protagonists in the development of NASH.[103] Studies have identified that NASH is associated with the presence of

circulating antibodies targeting neoantigens formed from the interaction between lipid peroxidation products and cellular proteins.[104] These IgG antibodies against malondialdehyde (MDA)-derived products may be present in up to 60% of patients with NASH.[105] Furthermore, raised anti-MDA IgG titres have been identified in patients with severe histologic NASH and advanced fibrosis.[65] In both serum and adipose tissue, B cell levels have been found to increase markedly in mice fed a high fat diet.[105] Interestingly, whilst B cell deficient mice display exhibit reduced insulin resistance, the transfer of B cells or specific IgG (isolated from wild type mice fed a high fat diet) is able to induce insulin resistance.

Inflammasome activation in the pathogenesis of NASH

Several studies have identified that inflammasome activation contributes to the development of NASH and fibrosis in NAFLD. In both mouse and human studies, both pro-IL-1β and pro-IL-18 have been found to be markedly increased in NASH.[106, 107] In addition, mice deficient of either NLRP3, IL1-α or IL-1β have been found to be from both from diet induced hepatic inflammation and fibrosis.[107, 108] In a further mouse study, inhibition of caspase-1 was also able to prevent the development of diet induced NASH and fibrosis.[109]

HSC activation in the pathogenesis of NASH

Prior to the development of NASH, HSC are hypothesised to be in a quiescent state. Thereafter, the development of both lipotoxicity and steatohepatitis is followed by HSC activation which results in ongoing fibrogenesis. Hitherto, ongoing studies have attempted to elucidate the mechanism for HSC activation in NASH:

- Studies of human hepatocytes have identified that lipid metabolite accumulation increases TGF-β and impairs adiponectin-mediated induction of activin A.[110]
- The Notch single-pass transmembrane receptor signalling pathway has been implicated in HSC activation. In a recent study, Notch pathway components were found to be upregulated in TGF-β activated HSC and in fibrotic livers with inhibition of Notch signalling decreasing HSC activation.[111]
- The stress-activated protein kinases, c-Jun N-terminal kinase (JNK) and p38 have been identified to have distinct and opposed roles in rat HSC.[112] Whereas JNK was identified to promote HSC proliferation, P38 was identified to inhibit HSC proliferation.
- Hedgehog pathway activation (sonic hedgehog) has been identified to corollate with the severity of histologic injury and fibrosis in human NAFLD biopsy samples.[113]

EXPLORING THE PATHOGENESIS OF INFLAMMATION AND FIBROSIS IN HCV

HCV Structure and replication

HCV is an enveloped, single stranded RNA molecule that is approximately 9600 nucleotides in length and encodes a polyprotein of approximately 3000 amino acids.[114] During HCV replication, the HCV polyprotein is cleaved by proteolytic enzymes into four structural and six non-structural proteins (figure 2.6).[115] New viral particles are assembled by the four structural proteins whilst the six n proteins support viral replication. The post-translational processing of the non-structural proteins from the polypeptide is catalysed by the NS3 serine protease and cofactor NS4A. NS3/NS4A together complete the post translational processing of the NS proteins at the NS3/NS4A, NS4A/NS4B, NS4B/NS5A and NS5A/NS5B junctions.[116] These non-structural replicative products are then released and form a complex responsible for forming viral RNA.

The lack of a proof-reading capacity in the HCV RNA polymerase contributes to the extensive genetic variation within HCV isolates. Six major genotypes of HCV are recognised with numerous sub-types. Within an infected individual an HCV may exist as multiple quasi-species or minor genetic variants within a sub-type. The geographic distribution of HCV genotypes varies at an epidemiological level so that specific genotypes and sub-types are more common in certain countries.

Fifty five to 85% of individuals who develop acute HCV infection will develop CHC which is associated with progressive liver fibrosis and cirrhosis, portal hypertension, liver failure and hepatocellular carcinoma.[16]

The primary goal of HCV treatment is to cure infection: successfully treated patients may attain a sustain virologic response (SVR) which is considered tantamount to a 'virological cure' (defined as the absence of HCV RNA 12-24 weeks after the discontinuation of antiviral therapy). The attainment of SVR is typically associated with normalisation of liver function tests and either significant improvement or resolution of hepatic necroinflammation. In addition, this may also be associated with stabilisation or regression of liver fibrosis.

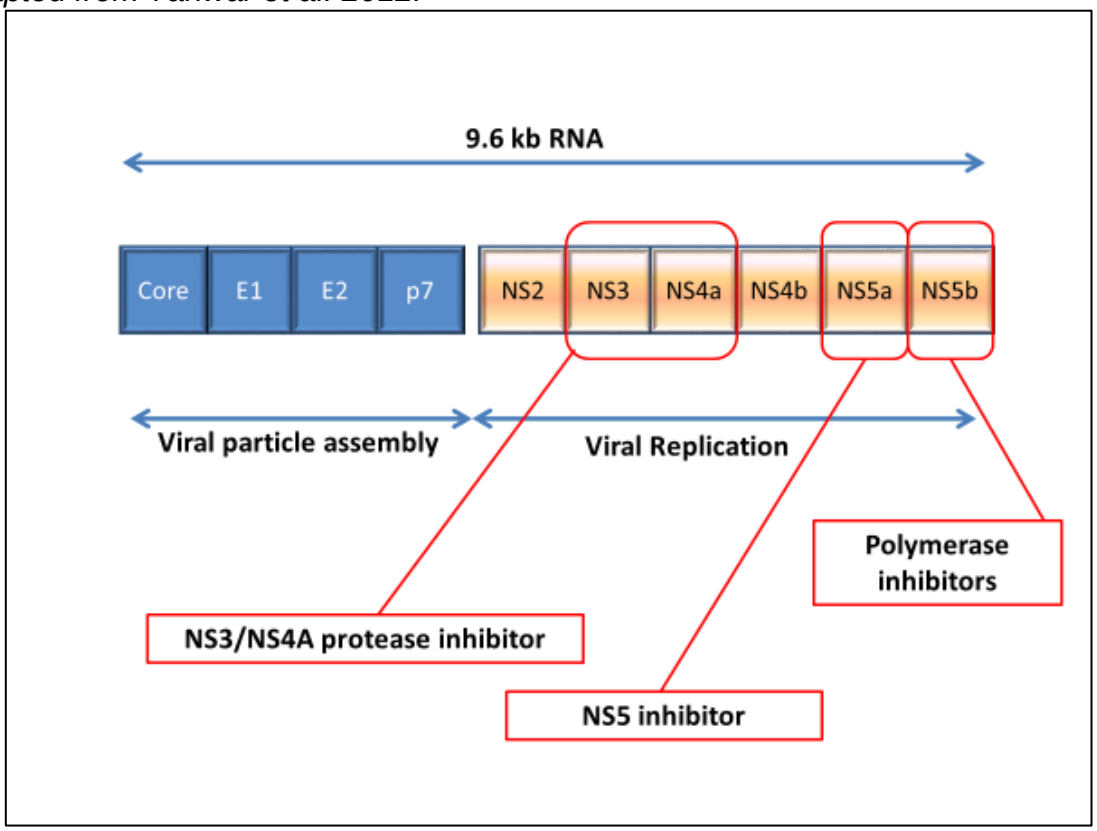
More recently, direct-acting antiviral (DAA) agents for the treatment of HCV have been developed that specifically target HCV viral replication. The development of DAA therapies is a consequence of the development of in-vitro HCV replication models [117] including the HCV replicon system. DAA agents were discovered by screening for their ability to inhibit viral replication.[118] In addition, the replicon systems were used to select and characterise resistant mutations to specific DAAs and also assess replication fitness.[119]

The introduction of DAA agents has led to HCV therapy in treatment regimens that are highly efficacious (SVR 95-100%), well tolerated and abbreviated even in difficult to treat populations such as those with advanced liver disease.[24]

Whilst therapy for established infection has rapidly evolved, the development of a vaccine for HCV remains in its infancy. There is currently no vaccine available against HCV, however greater understanding of host immunological factors and development of cell culture systems and animal models now offer the possibility of development of prophylactic and therapeutic vaccines

Figure 2.6 Depiction of HCV Genome structure and Drug Targets

Adapted from Tanwar et al. 2012.[120]



TLRs and HCV

As described earlier, TLRs are an important component of virus-mediated innate immune responses; this has also been confirmed in HCV. TLR-3 has been found to mediate antiviral responses against HCV in hepatoma cells.[121] In addition to inducing the production of type I interferons, TLR-3 activation also induces type III interferon (IFN-λ / IL-28) production.[122] A recent study also identified that a pro-fibrogenic HSC phenotype is mediated by TLR-2 signalling pathway following HCV core and NS3 protein exposure.[123]

Direct activation of fibrogenesis by HCV

The interaction between hepatic stellate cells and HCV is now well established. Hepatocytes infected with HCV have been found to release several pro-fibrotic factors including TGF-β that modify the expression of several profibrogenic genes in HSCs.[124] HSC have also been identified to engulf apoptotic bodies derived from HCV infection-induced hepatocyte apoptosis; this process has been found to elicit a profibrogenic response.[125] There is also evidence that HSC and HCV directly communicate with each other[126, 127]; there is evidence that this may be mediated through both the core and non-structural proteins directly. Studies have identified that the core protein can active pro-mitogenic intracellular pathways with HSC with the NS3 and NS5 proteins stimulating pro-inflammatory pathways via NF-κB and c-JNK.[127] In addition, the E2 structural protein has also been identified as profibrotic with MMP-2 activation occurring after binding to CD81 of HSC.[126] In HCV replicon models studies using rat and human hepatic stellate cells, HCV replication has been followed by the activation of TGFβ1 and other profibrotic cytokines.[124, 128]

Indirect activation of fibrogenesis by HCV

Analogous to other liver diseases, the ongoing immune response to HCV infection results in ongoing fibrogenesis via the activation of HSC. The mediators involved in the activation and proliferation of HSC include PDGF, TGFα, VEGF and activation of epidermal growth factor (EGF) receptor. Thereafter, several factors promote ECM production via TGF-β1 (stimulating collagen production via Smad signalling). Importantly, connective tissue growth factor (CTGF) and leptin also activate fibrogenesis via TGFβ1 signalling. Leptin also acts by suppressing proliferator-activated receptor γ (PPARγ). As described earlier, NF-κB signalling promotes a variety of chemotactic and mitogenic HSC effects.

Reversibility of Liver Fibrosis

Fibrosis regression in CHC

Hitherto, treatment for chronic liver disease has been primarily aetiology specific in which the causative agent or factor has been eliminated or modified. The treatment of CHC is an excellent example of this in which successful antiviral therapy results in SVR which is tantamount to a virologic cure. The attainment of SVR abolishes ongoing liver injury and studies have confirmed that not only is fibrogenesis arrested but fibrosis reversal can also manifest.[129-131]

Clinical regression of fibrosis in cirrhosis

Studies have identified that the potential for liver fibrosis regression may also be dependent on the duration of fibrosis and on several scar factors including scar composition, cellularity and spatial distribution.[132] Areas that may not be completely degraded include those that are acellular, rich in elastin and heavily cross-linked. Both animal and human studies have also identified that the regression of cirrhosis results in the transformation of cirrhotic micronodules into larger macronodules.[132, 133] Furthermore, due to the significant vascular distortion, shunting and angiogenesis that is associated with cirrhosis, significant fibrosis regression in a previously cirrhotic liver may not result in portal pressure reduction.[134]

Mechanisms for fibrosis regression

An understanding of the mechanisms that underlie fibrosis regression have led to the identification of antifibrotic targets and the developments of antifibrotic agents.[135]

Studies have confirmed that rather than being a fixed entity, the hepatic scar is in a state of continuous dynamic flux with respect to its cellular and matrix composition. In keeping with fibrosis progression, both myofibroblasts and hepatic macrophages are key protagonists in the regression of fibrosis. MMPs are a group of proteolytic enzymes that are responsible for the degradation of extracellular matrix components. The MMP proteolytic activity is regulated by their potent specific inhibitors tissue inhibitors TIMPs as expressed by activated HSCs.[136] Moreover, the apoptosis of activated HSC may be inhibited by TIMPs.

In both animal models and human studies documenting fibrosis regression with scar degradation, TIMP levels decrease dramatically and MMP levels rise.[132, 136] Thereafter, profibrogenic myofibroblasts are removed from the receding hepatic scar by apoptosis. In addition, recent studies have identified that hepatic macrophages are important regulators of ECM remodelling[137] highlighting the importance of specific macrophage subsets in either fibrogenesis or fibrosis regression (Ly-6C^{lo}).[138]

HSCs and fibrosis regression

As described earlier, HSCs play an important contribution to ECM remodelling in fibrosis evolution. Three mechanisms have been described through which HSCs can contirubute to fibrosis regression: apoptosis, senescence or quiescence.

- **HSC apoptosis**

HSC apoptosis leads to a decrease in the number of activated HSCs and has been observed during fibrosis regression.[139] This results in a decrease in TIMP-1 expression and thus inhibiting MMP activity. Apoptosis can be initiated in HSCs via CD95L (Fas ligand), Bcl-2 and p53.[140] Moreover, apoptosis can be initiated by hepatocytes (via NGF)[141], NK cells (via TRAIL and NKG2D)[51] and Kupffer cells (via capsase-9).[142]

- **HSC senescence**

There is evidence that HSCs can enter a senescent phase adopting a more inlammatory but less fibrogenic phenotype.[143] Both in culture and in vivo there is evidenc that HSC senescence is mediated by p53.

- **HSC quiescence**

Studies have demonstrated that a large proportion (~50%) of activated HSCs can revert to a quiescent phenotype.[144, 145] What has also been demonstrated, however, is that quiescent HSCs have a lower threshold to enter an activated state than naïve HSCs. HSC quiescence is thought to be mediated by PPAR-γ.

Targets of HSC mediated antifibrotic therapy

- Reduction of the inflammatory/immune response or inhibition hepatocyte apoptosis/injury to avoid HSC activation
- Promotion of fibrosis regression by inhibiting scar formation; promotion of matrix degradation; inhibition HSC activation or stimulate HSC apoptosis/ senescence/quiescence.
- Inhibition of signalling pathways (extracellular and intracellular) that initiate and perpetuate HSCs

Classes of antifibrotic therapy

Treatments for liver fibrosis can be categorised into several ways. Antifibrotic treatment may target the underlying aetiology of liver disease specifically (for example in CHC). Conversely, antifibrotic treatments may act with a direct anti-fibrotic action in the absence of any effect on the underlying aetiology of liver disease. These may be established drugs licenced for other indications and rediscovered to investigate their potential antifibrotic effect (drug repositioning).

Type 1 interferons

In CHC, both reduction in portal pressure[146] and significant histological improvement with antiviral therapy has been observed in the absence of achieving SVR.[147] These observations are supported by studies in animal models where type 1 interferons have been shown to have a direct antifibrotic effect.[148, 149] These basic science observations were translated into the three large clinical trials that addressed the use of long term interferon in patients with advanced fibrosis.[150-152] Although these trials showed that low maintenance doses of PEG-IFN did not improve outcomes in patients with

compensated cirrhosis and CHC who had not responded to a lead-in phase of interferon, this will be further discussed in chapter 11.

Specific antifibrotic drug targets:

- **Activated myofibroblasts**

Activated myofibroblasts can contribute to fibrosis regression by the release of proteolytic enzymes (mainly MMPs) that degrade ECM when subjected to favourable stimuli lintegrin receptor-mediated) in a process called ‘stress relaxation’. Myofibroblast stress relaxation resulting in a reduction in fibrogenesis and portal hyper tension has been demonstrated in cirrhotic rats by inhibiting Rho kinase.[153] Myofibroblasts have also been targeted by liposomes loaded with siRNA-Loaded Cationic Nanohydrogel Particles.[154]

Damaged Hepatocytes

Myofibroblasts can become activated by the phagocytosis of apoptotic hepatocytes. Biliary fibrosis has been found to be ameliorated in mice by the inhibition of hepatocyte apoptosis using a pan-cases inhibitor or an antagonist of cathespin B (a lysosomal trigger of apoptosis).[155, 156]

Biliary progenitors

The ‘ductular reaction’ is classically seen in biliary fibrosis where biliary progenitor cells proliferate and secrete cytokines that promote HSC activation. In advanced fibrosis of ‘non-biliary’ liver diseases, liver progenitor cell proliferation is also marked after the development of portal fibrosis. Several drugs targeting biliary progenitor cells have been used as antifibrotics including those inhibiting integrin αvβ6 (a receptor for fibronectin and tenascin-C)[157] and the hedgehog pathway.[158, 159]

Liver sinusoidal endothelial cells

During perisinusoidal fibrosis, activated LSECs produce ECM and secrete fibrogenic cytokines. LSECs themselves are activated by angiogenic factors produced by myofibroblasts including VEGF and angiopoetin-1.[160] Polykinase inhibitors such as sunitinib and sorafenib are antiangiogenic and have been identified to ameliorate experimental liver fibrosis.[66, 161]

Monocytes

Monocytes are the precursors of myofibroblasts, macrophages and dendritic cells playing an important role in inflammation and fibrosis. The chemokine CXCL9 has been identified to inhibit fibrogenesis by activating the monocyte receptor CX3CR.[162, 163]

Lysyl oxidase (LOXL2)

The enzyme LOXL2 mediates ECM crosslinking of matrix proteins including collagen. Antifibrotic activity has been demonstrated using LOXL an experiment models of liver fibrosis.[164] Simtuzumab is a monoclonal antibody targeting LOX2 and is currently being evaluated in phase II trials in patients with NAFLD with and without cirrhosis.

TLRs

TLR activation promotes a proinflammatory environment and TLR3 and TLR4 are being investigated for their antifibrotic properties.

Treatments and Therapies for NASH

Numerous medications are currently being studied for the treatment of NASH and fibrosis related to NAFLD. Whereas many of these medications have been shown to be effective in improving liver histology, the magnitude of the treatment effect is at best modest and there may also be associated adverse events.

Diet and physical activity

Sustained weight loss of more than 10% has been identified to be effective at improving liver histology (inflammation and fibrosis).[165-167]

PPAR agonists

Perioxisome proliferator-activator receptors (PPARs) are nuclear receptors that are expressed in a variety of organs including the liver. There are 3 types of PPAR receptor: α, β/δ and γ. PPARs regular metabolic processes including oxidation and lipid transport. PPARα agonists (for example fibrates) have not shown significant histologic benefit in NAFLD. In a phase IIb study, the dual PPARα/δ agonist was found to increase resolution of histologic NASH[168]; phase III studies are currently being conducted. Thiazolidinediones (for example pioglitazone) are PPARγ agonists that commonly used to treat diabetes and act as insulin sensitisers. In patients with NAFLD, treatment with pioglitazone has been shown to improve liver histology.[169-171] There is however, a risk of weight gain and congestive heart failure with pioglitazone treatment.

FXR (Farnesoid X receptor) bile acid axis

FXR is a bile acid intracellular receptor that both regulates bile acid synthesis and decreases hepatic gluconeogenesis.[172] Obeticholic acid is a synthetic bile acid

derivative and a FXR agonist. In phase II studies treatment with Obeticholic acid significantly improved liver histology (NASH and fibrosis) in patients with NAFLD.[173] Side effects of Obeticholic acid include reversible pruritus and worsening lipid profile. Phase III trials of Obeticholic acid are currently being conducted. Other agents work via the FXR-bile acid are also under investigation (FGF-19 and NGM-282)

Lipid-altering agents

Aramchol is Stearoyl-CoA desaturase (SCD) inhibitor that is currently being investigated in phase III studies for NASH. In a small phase II study Aramchol was associated with a decrease in hepatic fat content.[174] HMG-CoA reductase inhibitors (statins) are used extensively in both primary and secondary care for the treatment of hyperlipidaemia. There is evidence that statin use may improve liver histology (steatosis, steatohepatitis and fibrosis) in patients with NAFLD.[175]

Incretin based therapies

Glucagon-like peptide 1 (GLP-1) agonists such as liraglutide and exenatide are primarily used for the treatment of diabetes. GLP-1 belongs to the incretin family of proteins and acts on the pancreas to cause beta-cell proliferation and enhance insulin biosynthesis. A meta-analysis of several studies identified that treatment with GLP-1 agonist was associated with improved hepatic steatosis and fibrosis.[176]

Agents targeting inflammation, cell injury and inflammation

Vitamin E is an antioxidant that has been shown in both paediatric and adult NAFLD studies to improve hepatic steatosis, inflammation and NASH but not hepatic fibrosis.[177, 178] Pentoxyphilline is a methylxanthine derivative that modulates several cytokines including inhibiting TNFα. In a small study of 55 patients, as compared to

placebo, patients with biopsy proven NASH treated with pentoxyphilline were found to have improved NAS scores and fibrosis scores.[179] Larger studies will be needed to confirm these findings.

CHAPTER 3

NON-INVASIVE ASSESSMENT OF LIVER DISEASE

ABSTRACT

The histologic assessment of a needle core liver biopsy specimen is an imperfect reference standard. Despite the introduction of semi-continuous scoring systems to more accurately stage and grade liver histology, sampling variability and poor sample size can lead to both under and over staging of liver disease. Moreover, liver biopsy is an invasive procedure with associated morbidity and even mortality.

These limitations have driven the development of non-invasive tests (NIT) for liver disease using either serum or imaging-based methods. Serum based NITs can be classified into indirect markers (measuring standard biochemical parameters), direct markers (measures of fibrogenesis and fibrinolysis) or a combination of both. Imaging based techniques have assessed liver fibrosis by measuring liver stiffness as a surrogate for fibrosis itself. NITs have performed well in identifying the cross-sectional diagnostic targets of advanced fibrosis and cirrhosis with more variable results in the detection of mild and moderate fibrosis. Regardless, several NITs of liver fibrosis have also outperformed liver histologic staging in the prediction of long term liver related outcomes. NITs have also been used to detected other diagnostic targets such as steatosis and portal hypertension.

Current limitations of NITs include the inability to accurately detect of mild liver fibrosis and NASH. Furthermore, the ability of NITs to monitor fibrosis evolution during antifibrotic therapy is not established. Finally, given numerous NITs that have been developed, equipoise remains over how to optimally apply these NITs in clinical care for maximal patient benefit particularly in primary care.

Liver Biopsy – an imperfect reference standard

Chronic liver disease is characterised by progressive hepatic fibrosis which over time can culminate in cirrhosis and its complications of portal hypertension, hepatocellular carcinoma and liver failure.[3] Specifically, liver related morbidity correlates with increasing hepatic fibrosis in CLDs.[5, 180] The reference standard for staging liver fibrosis remains the histological staging of a liver biopsy specimen. Liver biopsy is an invasive procedure and is resource-intensive. Moreover, even in experienced centres, liver biopsy is associated with complications such as pain (20%), serious morbidity (0.6%) and even death (0.01%).[28] However, even under optimal conditions, liver biopsy is far from a perfect reference standard. Aside from the hazards associated with the procedure of liver biopsy itself, sampling variability may under or over stage fibrosis in as many of 20% of liver biopsies.[28] The accuracy and reliability liver biopsy is limited not only because of sampling error [181] but due to inter- and intra-observer variability.[182] In order to obtain a representative sample it has been proposed that a liver biopsy specimen must be of adequate length (minimum 20mm in length containing ≥11 portal tracts) and width (preferably using 16 gauge needle) to optimise accurate histopathologic fibrosis staging.[183, 184] At this size a liver biopsy specimen represents approximately 1/50,000th of the liver and it is well documented that liver histologic changes can vary within a particular liver segment or lobe.[183] In clinical practice however, biopsy specimens are often much smaller resulting in a decrease in diagnostic accuracy.[183] Thus, due to biopsy sampling error the calculated diagnostic accuracy of even a perfect non-invasive test would be reduced. Furthermore, as liver biopsy is invasive with associated morbidity, there are ethical considerations related to repeated procedures that hinder its use for longitudinal disease assessment.

Techniques to obtain a liver biopsy

Currently, there are three techniques for performing liver biopsy: percutaneous, transjugular and intraoperative.[185] Percutaneous biopsy can be performed without the use of concurrent imaging (using percussion) or with ultrasound guidance. In the former approach, percussion is used to delineate the position for tissue sampling. Percutaneous biopsy in the presence of ascites and coagulopathy confers an increased risk of bleeding. The technique of transjugular liver biopsy can, however, be performed in these circumstances. In a transjugular biopsy, usually the right internal jugular vein is cannulated using ultrasound guidance. Thereafter, fluoroscopy is used to navigate a guidewire through the right heart and into the hepatic vein. A biopsy sampling system is then navigated over the wire and liver tissue is obtained. During the procedure, the hepatic pressure gradient can be calculated by measuring the wedge pressure and the free hepatic venous pressure. As compared with the percutaneous liver biopsy, limitations of transjugular liver biopsy include additional cost, radiation exposure, and increased procedure time. Moreover, many centres do not perform transjugular liver biopsy necessitating additional interventional radiology training. The final technique by which liver tissue can be obtained is intraoperative (open or laparoscopic surgery).

Information gained from the assessment of liver histology

Nevertheless, the histological assessment of a liver biopsy specimen offers a plethora of information on a variety of clinical factors:

Aetiology of liver disease

The acquisition of liver histology may help uncover the underlying aetiology of liver disease if it is not apparent using a combination of clinical history, blood tests and imaging studies. In addition, if more than one aetiological agent has been identified, histologic analysis can offer information on relative contribution of each aetiology.

Stage of Fibrosis

Traditionally liver histologic analysis, including fibrosis severity, was performed in a descriptive fashion. Pathological descriptions of liver histology specimens can range from being extremely detailed to being minimally descriptive or ambiguous. Such descriptions do not readily provide definitive endpoints assessing disease clinical disease progression or for statistical analysis. The development of disease-specific scoring systems for fibrosis severity have provided an additional tool for evaluating disease progression, need for treatment and response to treatment. In 1981, Robert Knodell and colleagues described the Knodell score for staging liver biopsy specimens in patients with chronic hepatitis (non-A non-B hepatitis or hepatitis B). This was the first time a semi-continuous numerical score was used to stage fibrosis and grade inflammation in patients with CLD[186]. In the Knodell score the degree of periportal necrosis, intralobular necrosis, portal inflammation, and fibrosis were numerically quantified. Thereafter, further semi-continuous scores have been constructed and validated for patients with mixed-aetiology and aetiology specific CLD. Example of these scoring systems are presented below in tables 3.1-3.5.

Table 3.1 The Knodell scoring system for staging liver fibrosis

Knodell[186]	**Chronic Hepatitis**
0	No fibrosis
1	Fibrous portal expansion
3	Bridging fibrosis (portal-portal or portal-central linkage)
4	Cirrhosis

Table 3.2 The Ishak scoring system for staging liver fibrosis

Ishak[187]	**Chronic Hepatitis**
0	No fibrosis
1	Fibrous expansion of some portal areas, with or without short fibrous septa
2	Fibrous expansion of most portal areas, with or without short septa
3	Fibrous expansion of most portal areas with occasional portal to portal bridging
4	Fibrous expansion of portal areas with marked bridging (portal to portal as well as portal to central)
6	Marked bridging with occasional nodules (incomplete cirrhosis)
7	Cirrhosis, probable or definite

Table 3.3 The Metavir scoring system for staging liver fibrosis

Metavir[188]	**Chronic Hepatitis C**
0	No scarring
1	Stellate enlargement of portal tracts without septae formation
2	Enlargement of portal tracts with rare septae formation
3	Bridging fibrosis is spreading and connecting to other areas that contain fibrosis
4	Cirrhosis

Table 3.4 The Scheuer scoring system for staging liver fibrosis

Scheuer[189]	**Chronic Viral Hepatitis**
0	None
1	Enlarged, fibrotic portal tracts
2	Periportal or portal-portal septa, but intact architecture
3	Fibrosis with architectural distortion, but no obvious cirrhosis
4	Probable or definite cirrhosis

Table 3.5 The Kleiner scoring system for staging liver fibrosis

Kleiner[190]	**NAFLD**
0	No fibrosis
1	Mild /moderate zone 3 perisinusoidal fibrosis, or portal fibrosis only
2	Zone 3 and portal/periportal fibrosis
3	Bridging Fibrosis
4	Cirrhosis

In the field of non-invasive markers of fibrosis, scoring systems have provided the reference standard for the development and validation of individual markers and panels, providing a dichotomous target based on the quantity and location of fibrosis.[191]

Grade of Inflammatory Activity

In contrast to hepatic fibrosis, which is an overarching indicator cumulative chronic hepatic injury, the degree of hepatic inflammation is an indicator of current ongoing liver injury. The scoring systems described above all describe the necro-inflammatory grade of the liver including the degree of periportal necrosis, intralobular necrosis and portal inflammation

Non-alcoholic steatohepatitis and the NAFLD activity score

NASH is diagnosed histologically by identifying presence of a characteristic pattern of steatosis, inflammation and hepatocellular ballooning on liver biopsies in the clinical absence of significant alcohol consumption.[192] However, the dichotomous assessment of either having 'steatohepatitis or not' does not facilitate the longitudinal assessment of NASH severity. For this reason, a scoring system was devised that included the full spectrum of NAFLD histologic changes and would be sensitive to histologic changes within an established diagnosis of NASH. The NAFLD activity score (NAS) was derived from the blinded and individual readings of biopsies from 32 adults and 18 children with clinically presumed NAFLD by ten pathologists. The NAS score ranges from (0-8) which incorporates the scores of steatosis (0-3), lobular inflammation (0-3), and ballooning (0-2); a score of ≥5 is suggestive of NASH (table 3.6).

In a validation study of the NAS score, the numeric scores correlated closely but not perfectly with separately derived diagnoses of "definite steatohepatitis (SH)", "not SH" and "borderline SH". Of biopsies with a NAS ≥ 5, 86% had SH and 3% "not SH". Importantly, NAS ≤ 4 did not indicate benign histology; 29% had SH and only 42% had "not SH".[77] Regardless, recent data suggest that amongst patients with advanced fibrosis due to NASH, a change of fibrosis stage is a more important indicator of prognosis than a change in NAS score.[193]

Table 3.6 The NAFLD activity Score (NAS) and its constituent components

NAS Component		Score
1. Steatosis		
Low- to medium-power evaluation of parenchymal involvement by steatosis	<5%	0
	5–33%	1
	34-36	2
	>36	3
2. Lobular inflammation		
Overall assessment of all inflammatory foci	No foci	0
		1
	2–4 foci per × 200 field	2
		3
3. Hepatocellular ballooning		
	None	0
	Few balloon cells	1
	Many cells/prominent ballooning	2

Presence of Pre-Malignant or Malignant Disease

In addition to describing the presence of inflammation and fibrosis, histologic analysis can also detect dysplastic or frankly malignant tissue. As described earlier, cirrhosis is a premalignant condition conferring an increased risk of HCC. If malignant tissue is identified by liver biopsy the risk of tumour seeding has been estimated at approximately 2%.[194]

Collagen proportionate area for fibrosis staging

Histopathologic interobserver variability limits the accuracy and reproducibility of the aforementioned semi-continuous histologic scoring systems for fibrosis staging; this is most marked when quantifying lesser stages of fibrosis. However, it is possible to eliminate interobserver variability completely by using automated systems that accurately quantify the percentage of fibrous tissue (collagen) evident on a single biopsy specimen. Collagen proportionate area[195] is an example of this type of assessment. In this technique, dye is used to stain collagen on the liver biopsy specimen. Thereafter, an automated grid based camera system is used to determine the area stained positive for collagen proportionate to total liver tissue with gating used to exclude major vessels, bile ducts and capsule. Another potential advantage of collagen proportionate area over classical histologic staging is the ability to sub-classify cirrhotic patients into those at higher risk of subsequent decompensation.[196] Despite offering a more robust method of quantifying fibrosis on an individual biopsy specimen, these methods still suffer with the inherent limitations of liver biopsy such as sampling variability, hazard to the patient and cost.

NON-INVASIVE TESTS OR THE ASSESSMENT OF LIVER DISEASE

Diagnostic targets of NITs

Histologic analysis offers a plethora of information on several aspects of liver disease; this information can be reported as a semi-continuous variable (for example the Metavir fibrosis stage). In addition, the histologic analysis of a liver biopsy specimen includes information on the aetiology of liver disease. By contrast, non-invasive tests are examined for their association with a specific single parameter that reflects liver disease severity (diagnostic target). A diagnostic target should be of clinical relevance and its presence or absence should determine patient management or offer prognostic information. As a result, current NITs are not a replacement for histologic assessment.

- **Mild Fibrosis (Metavir or Kleiner F1, Scheuer S1)**

The development of mild fibrosis signifies the onset of liver fibrosis. In the context of NAFLD this diagnostic target is highly relevant as positivity would signify progression from simple steatosis to a more aggressive phenotype that develops only a minority of patients. By contrast, in CHC this diagnostic target is less relevant as almost all patients with CHC will develop mild fibrosis at some stage during their disease course.

- **Moderate Fibrosis (Metavir or Brunt F2, Scheuer S2)**

The development of moderate fibrosis signifies the development of progressive liver fibrosis. In NAFLD this signifies the presence of increasing fibrosis due to NASH. In the context of CHC, if a patient has been infected for only a short period the development of moderate fibrosis would signify rapid fibrosis progression.

- **Advanced Fibrosis (F3 Metavir or Brunt, Scheuer S3)**

Advanced fibrosis is an important diagnostic target as it signifies extensive liver damage. Patients with advanced fibrosis should be offered intervention to arrest progression to cirrhosis.

- **Cirrhosis (F4 Metavir or Brunt, Scheuer S4)**

Cirrhosis undoubtedly remains the most important diagnostic target for NITs in liver disease. All patients with cirrhosis should receive treatment for their underlying liver disease. In addition, as discussed in chapter 1, all patients with cirrhosis regardless of aetiology should be screened for HCC development and the presence of oesophageal varices.

Fibrosis Evolution

The diagnostic targets described identify the presence or absence of a disease phenotype for example ‘advanced’ versus ‘non-advanced’ fibrosis. However, disease severity is not in reality dichotomous but instead a continuum. Liver fibrosis is the result of a dynamic balance between fibrogenesis and fibrinolysis with an increase in either process resulting in fibrosis progression and regression respectively. To identify both fibrosis progression and regression, ideally the evolution of NIT scores should correlate with the evolution of the liver histology.

Presence of Inflammation and Steatohepatitis

The identification of inflammation in CLDs is important as inflammation itself is often the driver of fibrogenesis. In NAFLD the development of NASH is recognised as highly

significant as it signifies the transformation from a benign to an aggressive phenotype. Therefore, along with the detection of increasing fibrosis, the identification of NASH in NAFLD is an alternative method by which to detect patients with progressive disease. In the context of CHC, however, the detection of inflammation is now less relevant given that fibrosis progression can be relatively slow and treatment is now highly efficacious.

Presence of Steatosis (NAFLD)

The development of steatosis is the 'first hit' in the current hypothesis for the pathogenesis of NASH. Accordingly, NAFLD can only be identified after steatosis is confirmed. Thereafter, following a diagnosis of steatosis, further tests for NASH and/or fibrosis can be applied to identify those patients at risk of disease progression.

Portal Hypertension and Hepatic Venous Pressure Gradient

Direct measurement of portal pressure can be performed at the time of transjugular liver biopsy by measuring the hepatic venous pressure gradient (HVPG). HVPG is calculated by subtracting the wedged (sinusoidal) pressure from the free hepatic venous pressure. Normal HVPG is ≤5mmHg with clinically significant and severe portal hypertension is defined as HVPG ≥10mmg and ≥12mmHg respectively.[197] Severe portal hypertension (HVPG ≥12mmHg) can result in the development of oesophageal and gastric varices, ascites, hepatorenal syndrome and spontaneous bacterial peritonitis. The direct measurement of portal pressure is an invasive and time-consuming procedure with availability only in specialist centres.

Liver Related Clinical Outcomes

In clinical practice, discordance between NIT and liver biopsy may relate to the aforementioned limitations of liver biopsy that limit its accuracy. Regardless, the presence and degree of hepatic fibrotic change is used as a surrogate for the development of complications related to CLDs.

In longitudinal studies of patients with CLD, the development of liver related clinical outcomes represents another reference standard against which NITs can be assessed. However, given that liver histological analysis is used to determine prognosis, liver histology can also be viewed as merely a surrogate for clinical outcomes. The ability to predict liver related clinical outcomes allows clinicians to further stratify patients that are at higher risk of future hepatic decompensation. These clinical events may relate to portal hypertension, sepsis, HCC and hepatic failure. Patients deemed to be at high risk of decompensation can be treated and monitored more aggressively with earlier referral for transplantation considered.

NIT Derivation and Validation

Typically, in NIT derivation studies a ‘training set’ is employed in which a putative NIT (such as a potential biomarker or imaging parameter) is measured in a representative patient population and examined for its association with a diagnostic target as assessed by liver biopsy. As described above, fibrosis stages are usually dichotomised into diagnostic targets such as those with and without at least mild (e.g. Metavir F1-4), moderate (e.g. Metavir F2-4), advanced (e.g. Metavir F3-4) liver fibrosis or cirrhosis (e.g. Metavir F4). The diagnostic performance of the putative NIT in the training cohort can then be calculated. Thereafter, logistic regression modelling can be employed to identify whether a predictive algorithm composed of multiple parameters in addition to the putative NIT can offer improved diagnostic performance. Finally, the performance of the individual NIT or identified algorithm can be compared to its performance in an independent cohort of patients (‘validation set’).

Assessment of NIT performance and Spectrum Bias

Area under the receiver operator characteristic curve

The diagnostic accuracy of a NIT can be measured by calculating the area under the receiver operator characteristic (AUROC) curve. A receiver operator characteristic (ROC) curve plots sensitivity against 1-specificity for all possible values of a NIT as a continuous variable. With reference to table 3.7, sensitivity can be defined as the probability of a positive result in a patient with the diagnostic target[198] (TP/TP+FN). Specificity can be defined as the probability of a negative result in a patient without the diagnostic target[198] (TN/FP+TN). The positive predictive value (PPV) is the proportion of positive results that are true positive results (TP/TP+FP). The negative predictive value (NPV) is the proportion of negative results that are true negatives results (TN/TN+FN).

Table 3.7 Two by two contingency table comparing test result and disease state

		Disease State	
		Positive	Negative
Test result	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

AUROC ranges between 0 (a test with 0% sensitivity and 0% specificity) to 1 (a test with 100% sensitivity and 100% specificity). The clinical interpretation of AUROC is presented in table 3.8.

Table 3.8 Clinical interpretation of AUROC values[199]

AUROC	Interpretation of AUROC
>0.90	Excellent
0.80-0.90	Good
0.70-0.80	Fair
0.60-0.70	Poor
0.50-0.60	Fail
<0.5	Worse than chance

Although constructed from all the possible values of sensitivity and specificity for a NIT, the AUROC value does not signify the performance of a NIT at an individual diagnostic threshold.

It should be noted however that if sampling error and inter- and intra-observer variability limit the sensitivity and specificity of liver biopsy itself to 90% then a truly perfect biomarker of fibrosis could obtain a maximum AUROC of no more than 0.9.[200]

The performance of non-invasive tests can vary depending on the distribution of fibrosis stages in a cohorts in a phenomenon called spectrum bias.[201] For example, in a population composed entirely of patients with F0 and F4 fibrosis stages, the AUROC of a biomarker performance in its ability to discriminate between non-advanced (F0-2) and advanced fibrosis (F3-4) will be higher than in a population composed entirely of patients with F2 and F3 fibrosis stages. Due to confounding factors such as spectrum bias, numerical comparisons of diagnostic performance of NITs in different patient populations are subject to type I and type II error. As only a limited number of published studies have directly compared the performance of non-invasive tests within a single patient population, statistical methods[202, 203] have been developed to compare the AUROC derived for each biomarker with a correction applied for spectrum bias. Regardless, while these simulations are useful tools that facilitate theoretical

biomarker comparisons, they are inherently less powerful than a comparative study within a unique population. These statistical methods, Obuchowski measure and DANA are described below:

Obuchowski measure

The Obuchowski measure (ordAUROC) gives a weighted average of the N(N-1)/2 AUROC pairwise comparisons between N categories of gold standard outcome. For example, using the Metavir scale with its N (=5) categories of fibrosis staging (F0-4) there are 10 pairwise comparisons between 2 groupings of the N categories. Each pairwise comparison can be weighted to account for the distance between fibrosis stages. A 'penalty function' proportional to the difference in Metavir units between fibrosis stages can also be applied to differentially penalise incorrect fibrosis classification by a NIT. For example, the penalty function could be 0.25, 0.5, 0.75, and 1 when the difference between Metavir stages is 1, 2, 3 and 4 stages respectively. Using this penalty function, fibrosis misclassification by 2 stages would be penalised twice as severely as misclassification by a single stage with the final Obuchowski score adjusted accordingly.

Difference between the mean fibrosis stage of advanced fibrosis minus the mean fibrosis stage of nonadvanced fibrosis (DANA)

DANA assesses spectrum bias in the context of the Metavir 5 stage fibrosis scoring when specifically discriminating between patients with (F2-4) and without (F0-1) fibrosis. The DANA score is derived by calculating the difference in prevalence distribution between the mean fibrosis stage of patients with F2-4 fibrosis minus the mean fibrosis stage of patients with F0-1 fibrosis:

- (([%F2x2] + [%F3x3] + [%F4x4])/3) - (([%F1x1] + [%F0x0])/2)

A DANA of 4 is attained when the population is composed entirely of patients with F0 and F4 fibrosis. A DANA of 1 is attained when there is central clustering of fibrosis stages and the population is composed entirely of patients with F2 and F3 fibrosis only. Thereafter the AUROC of a non-invasive test calculated in any given population can be standardised to the distribution of fibrosis stages found in a reference population such as that observed in the French population:

- [adjusted AUROC = observed AUROC + (0.1056) (2.16 − DANA)].

Using NITs in clinical practice

NITs are used in clinical practice to help ascertain whether a diagnostic target is present or absent. Most NITs are reported as continuous variables. The presence or absence of a diagnostic target is signified by a NIT score that is greater or less than, respectively, a pre-determined diagnostic threshold. The *a priori* diagnostic threshold is chosen based on the pre-test probability of the diagnostic target and the relative cost of having either a false positive or false negative result. Following the application of a NIT, the PPV and NPV of the test can be calculated. The PPV is the likelihood of a diagnostic target being present when the NIT is positive (or the proportion of positive test results that are truly positive). Conversely the NPV is the likelihood of a diagnostic target being absent when the NIT is negative (or the proportion of negative test results that are truly negative).

Both the PPV and NPV are dependent not only on the diagnostic performance (sensitivity and specificity) of a NIT but also the pre-test probability of the diagnostic target in question.[204] For example, in the context of a very low pre-test probability of a diagnostic target, a positive result from a NIT with exemplary diagnostic performance would only generate a modest PPV. In the same context, a negative result from a NIT with poor performance will result in a very high NPV. Moreover, in order to further increase both the PPV and NPV of a NIT, two diagnostic thresholds can be employed. A lower threshold with higher sensitivity is used to ‘rule out’ disease and an upper one with higher specificity used ‘to rule’ in disease. Values lying between these 2 thresholds are classified as indeterminate in which the presence of the diagnostic target is uncertain. Conventionally the reference standard (for example liver biopsy in the context of staging liver fibrosis) is used to classify these indeterminate cases.

Combining Multiple Diagnostic Tests

Most NITs do not exhibit flawless diagnostic performance; this includes NITs that are currently being used in clinical practice. As a result, the PPV and NPV generated following the application of a NIT may be inadequate to either 'rule' in or 'rule' out a diagnosis. In this context, two or more tests may need to be applied to attain sufficient diagnostic certainty. Tests can be applied in parallel (at the same time and interpreted together) or in series (the application of second test is dependent on the result of the first test).[198] Equations that calculate the incremental gain (sensitivity or specificity) in diagnostic performance achieved by combining tests are described below. For maximal benefit, it is essential that these principles are followed when formulating diagnostic algorithms composed of multiple NITs in clinical practice.

Parallel Testing

Parallel testing requires all patients to be assessed by the proposed NITs. Parallel testing can be used in two ways: (1) OR (2) AND.[198]

- **OR**

  When using two tests with the parallel OR approach, a positive result is conferred if at least one of the tests is positive, and a negative result conferred if both tests are negative. The OR approach results in increased sensitivity as compared to single test application but decreased in specificity.

  The equation for calculating the sensitivity of the combined test in this context is:

  $$Sensitivity_{Test\ 1} + Sensitivity_{Test\ 2} – (Sensitivity_{Test\ 1} \times Sensitivity_{Test\ 2})$$

  The equation for calculating the Specificity of the combined test in this context is:

$Specificity_{Test\ 1} \times Specificity_{Test\ 2}$

- **AND**

Using the parallel AND approach, a positive result is conferred if both tests are positive, and a negative result conferred if both tests are negative. The AND approach results in increased specificity as compared to single test application but decreased sensitivity.

The equation for calculating the sensitivity of a combined test in this context is:

$Sensitivity_{Test\ 1} \times Sensitivity_{Test\ 2}$

The equation for calculating the specificity of a combined test in this context is:

$Specificity_{Test\ 1} + Specificity_{Test\ 2} - (Specificity_{Test\ 1} \times Specificity_{Test\ 2})$

Serial Testing

By contrast, serial testing avoids the application of unnecessary NITs. This may be useful when there are limitations surrounding the use of a second test such as expense or associated complications. Serial testing can also be used in two ways: (1) OR (2) AND.[198]

- **OR**

When using two tests with the serial OR approach, a positive result is conferred if at least one of the tests is positive; a negative result results in the application of a second test. If the second test result is also positive, then the overall result is positive (otherwise it is negative). The OR approach results in increased sensitivity as compared to single test application but decreased specificity.

The equation for calculating the sensitivity of a combined test in this context is:

$Sensitivity_{Test\ 1} + (1 - Sensitivity_{Test\ 1}) \times Sensitivity_{Test\ 2}$

The equation for calculating the specificity of a combined test in this context is:

$Specificity_{Test\ 1} \times Specificity_{Test\ 2}$

- **AND**

Using the serial AND approach, if the first test result is positive then a second test is applied. If the second test is positive, then the overall result is positive (otherwise a negative result is conferred). The serial AND approach results in increased specificity as compared to single test application but decreased sensitivity.

The equation for calculating the sensitivity of the combined test in this context is:

$Sensitivity_{Test\ 1} \times Sensitivity_{Test\ 2}$

The equation for calculating the specificity of the combined test in this context is:

$Specificity_{Test\ 1} + (1 - Specificity_{Test\ 1}) \times Specificity_{Test\ 2}$

NITs for the assessment of Liver Fibrosis

Liver function tests (LFTs), the standard blood tests employed in clinical practice to assess 'liver health', are neither sensitive nor specific for liver injury. Due to these deficiencies and the limitations of liver biopsy, multiple NITs have been developed over the last 20 years to detect progressive liver fibrosis and stage liver disease. In contrast to liver biopsy, NITs for liver fibrosis are less invasive with minimal associated procedural morbidity.

Current NITs for liver fibrosis can be grouped into 2 broad categories:

1. Serum markers of liver fibrosis
2. Imaging modalities to assess liver stiffness

Serum markers of liver fibrosis

Hitherto, numerous serum markers or biomarkers of liver fibrosis have been identified and described. Serum markers of liver fibrosis can be taken at the same time as standard blood tests and are therefore well suited to identification of liver disease in the population at risk.

Serum biomarkers of fibrosis can be categorised in several ways including into indirect serum markers (combinations of serum parameters which are related to liver function including aspartate aminotransferase (AST) and alanine aminotransferase (ALT)) and direct serum markers (measuring parameters directly related both the fibrolytic and fibrogenic processes involved in liver matrix turnover).[3] As derived by logistic regression, panels of biomarkers have been also described. Combinations of both

indirect and direct biomarkers can result in superior diagnostic performance for the detection of fibrosis as compared to their constituent components.

Indirect Biomarkers

Indirect (or class 2) biomarkers reflect parameters that are altered due to changes in hepatic function that arise in the context of advancing hepatic fibrosis. Indirect biomarkers include biochemical or haematological variables that are synthesised or regulated by the liver (for example clotting, cholesterol and bilirubin) or indicate inflammation (for example aminotransaminases).

The constituent components of indirect marker panels (for example AST or bilirubin) can be heavily influenced by factors such as inflammation (AST/ALT), haemolysis (bilirubin) or haemorrhage (platelet count).[205] As a result, it is essential that indirect serum marker results are not interpreted in isolation to avoid false negative and false positive results.

Direct Serum Markers

In contrast to indirect markers, direct (class 1) biomarkers exhibit biological plausibility as they measure alterations to ECM composition that occur during hepatic fibrosis. The most widely studied direct biomarker is Hyaluronic acid (HA), a glycosaminoglycan that is synthesised by hepatic stellate cells and degraded by liver sinusoidal cells.[206] Other direct biomarkers include terminal peptide of procollagen III (PIIINP)[206], TIMP-1[207], type IV collagen[208], and tumour growth factor β.[209] Direct serum markers of fibrosis exhibit biologically plausible but are not liver specific with rising levels also identified in extra-hepatic inflammation and fibrosis. In addition, levels of direct serum markers can also be elevated when biliary excretion or renal function is impaired.

In this thesis, the following direct serum markers have been employed to assess liver fibrosis:

- **Hyaluronic acid**

HA is a polysaccharide composed of 2 alternating subunits of N-acetyl-D-glucosamine and β-glucuronic acid.[210] HA forms un-branched, linear chains of varying size that can reach up to 107 Da in length. HA functions as a major component of ECM in the body, whilst shorter oligosaccharide forms are involved in mediation of proinflammatory cytokines.[211] Uptake and degradation of HA occurs via hyaluronate receptors in sinusoidal epithelial cells (SECs) within the liver.

- **Tissue Inhibitors of Metalloprotease**

TIMPs are a family of 4 glycoproteins with a molecular weight of around 20kDa.[212] The principle functions of TIMPs are inhibition of activated matrix degrading MMPs, and prevention of pro-MMP activation.

- **N-Terminal peptide of type III Procollagen**

PIIINP is a cleavage product produced when type III collagen fibrils are laid down in ECM.[213] Under normal physiological conditions, type III collagen is confined to the portal tracts and around large vessels in the liver.[213] However, under fibrotic conditions, deposition outside of these areas occurs. Increased collagen deposition results in raised levels of cleavage products in the serum. PIIINP serum concentrations offer an indication of the level of type III collagen production. Conversely, elevated serum PIIINP has been shown to be associated with elevated pro-inflammatory cytokines in liver disease and therefore may also be a marker of hepatic inflammation.[214]

- **YKL-40**

Human YKL-40 was first discovered in 1989 and was fully sequenced in 1993.[215] YKL-40 is a single 383 amino acid polypeptide belonging to the family of glycoproteins known as the glycosyl hydrolases. The name describes the 3 N-terminal amino acids (Tyrosine [Y], Lycine [K], and Leucine [L]), and the approximate molecular mass of 40 kDa. Studies have identified the presence of YKL-40 mRNA within the fibrotic liver[216] with suppression subtractive hybridisation confirming that YKL-40 to be among the most over-expressed proteins in cirrhotic liver.[217] Studies have suggested that YKL-40 may act as a signalling molecule, facilitating evasion of apoptosis.[218] It has been shown that an increase in YKL-40 results in upregulation of MAP-kinase and PI-3k signalling in fibroblasts. This in turn leads to phosphorylation of ERK-1/2 MAP-kinase, and AKT cascades, which are involved in apoptotic pathways. This indicates a

potential role of YKL-40 in cell survival as an anti-apoptotic protein.[219] Moreover, studies have found raised serum YKL-40 in up to 93% of patients with active fibrogenesis, suggesting a role for this protein as a marker of ongoing fibrotic change and inflammation.[220]

Summary of Serum NIT for the detection of fibrosis

Several Biomarker panels have been described as NITs for liver fibrosis. The formulae for these panels are described in the appendix. Five panels (ELF[221], Fibrometer[222], FibroSpect II[48], Fibrotest[223] and Hepascore[224]) are now commercially available following patent protection.

The published performances of the currently available biomarker panels as NITs for liver fibrosis are outlined in tables 3.9-3.13. Overall, panels composed of direct serum biomarkers have generally exhibited superior diagnostic performance than indirect serum markers particularly for milder stages of fibrosis.

Table 3.9 Performance of Direct Serum Markers for the non-invasive detection of fibrosis in CLD

Biomarker	AUROC	Sensitivity %	Specificity %	PPV %	NPV %	Ref
Hyaluronic acid	Ishak≥4 0.88	75	-	81	85	Mehta 08 [225]
	F≥3 0.86	65	91	-	-	Guéchot 96 [226]
	F4 0.92	79	89	-	-	
YKL-40	Ishak≥4 0.75	75	-	58	82	Mehta 08 [225]
Type IV collagen	Ishak≥3 0.83	73	85	-	-	Walsh 00 [227]
Laminin	Ishak≥3 0.82	80	83	-	-	Walsh 00 [227]
PIIINP	F≥3 0.69	70	63	-	-	Guéchot 96 [226]
	F4 0.73	60	74	-	-	
TIMP-1	Ishak≥3 0.73	94	57	-	-	Walsh 99 [228]

Table 3.10 Diagnostic Performance of Non-Patented Serum Markers for the Non-Invasive Detection of Fibrosis in CHC and Mixed Aetiology CLD Part 1

Biomarker Panel	AUROC	Sensitivity %	Specificity %	PPV %	NPV %	Ref
MP3	F≥2 0.84	96/82/44/19	24/73/96/99	57/76/91/94	84/80/62/54	Leroy 07 [229]
	F≥3 0.88	100/92/61/31	20/92/61/31	33/59/90/98	100/95/85/78	
APRI	F≥2 0.81	92/80/72/58	27/63/88/94	56/69/86/91	76/75/76/69	Leroy 07 [229]
	F≥3 0.82	94/89/87/74	22/54/75/84	32/43/57/64	90/93/94/92	
	F≥2 0.69/0.77	84/79	77/95	87/96	73/72	Sebastiani 06 [230]
	F4 0.61	38	87	39	87	
	F≥2 0.88	-	-	64/91	90/65	Wai 03 [231]
	F4 0.94	-	-	35/65	100/95	
	Ishak≥3 0.78	-	-	-	-	Fontana 06 [232]
	F4 0.80	30/64	81/94	52/62	81/88	Castera 09 [233]
	F4 0.83	78	75	30	96	Islam 05 [234]
	F4 0.73	-	-	-	-	Fontana 08 [235]
	F≥2 0.69	14/36	91/100	90/100	42/57	Sebastiani 08 [236]
	F≥2 0.75	27/79	57/100	81/100	36/54	
Forn's Index	F≥2 0.78	88/42	42/93	61/86	78/61	Leroy 07 [229]
	F≥3 0.78	92/54	34/87	35/62	91/83	
	F≥2 0.79/0.58	80/79	61/82	78/85	64/75	Sebastiani 06 [230]
	F≥2 0.81	94/30	51/95	40/66	96/80	Forns 02 [237]
	F≥2 0.81	-	-	-	-	Koda 07 [238]
	F≥3 0.90	-	-	-	-	
	F≥2 0.76	21/85	98/49	96/79	35/59	Sebastiani 08 [236]
Fibroindex	F≥2 0.86	-	-	-	-	Koda 07 [238]
	F≥3 0.93	-	-	-	-	
	F≥2 0.58	10-41	100/77	100/75	43/39	Sebastiani 08 [236]
	F≥2 0.74	19-68	100/69	100/84	47/47	

Table 3.11 Diagnostic Performance of Non-Patented Serum Markers for the Non-Invasive Detection of Fibrosis In CHC and Mixed Aetiology CLD Part 2

Biomarker Panel	AUROC	Sensitivity %	Specificity %	PPV %	NPV %	Ref
AST/ALT Ratio	Ishak≥3 0.58	-	-	-	-	Fontana 06 [232]
	F4 0.61	31	89	47	81	Castera 09 [233]
	F≥2 0.51	12	88	70	29	Sebastiani 08 [236]
	F≥2 0.54	43	58	67	35	
Lok Index	F4 0.81	40/98	99/53	84/27	90/99	Lok 05 [239]
	F4 0.80	40/86	94/46	62/32	81/91	Castera 09 [233]
	F4 0.79	-	-	-	-	Fontana 08 [235]
GUCI	F≥4 0.85	80	78	31	97	Islam 05 [234]
FIB-4	F≥3 0.85	38-74	81-98	-	-	Vallet-Pichard 07 [240]
	F4 0.91	-	-	-	-	
HALT-C	F4 0.81	47/88	92/45	78/43	74/86	Fontana 08 [235]

Table 3.12 Diagnostic Performance of Patented Serum Markers for the Non-Invasive Detection of Fibrosis in CHC and Mixed Aetiology CLD

Biomarker Panel	AUROC	Sensitivity %	Specificity %	PPV %	NPV %	Ref
FibroTest	F≥2 0.87	75	85	-	-	Imbert-Bismut 01 [223]
	F≥2 0.84	89/76/45	53/74/90	66/75/82	83/75/62	Leroy 07 [229]
	F≥3 0.87	94/90/67	42/64/88	39/50/68	95/94/87	
	F≥2 0.82	65/58	81/91	80/78	67/81	Sebastiani 06 [230]
	F4 0.71	50	93	58	91	
	F4 0.82	55	86	55	86	Castera 09 [233]
	F≥2 0.70	67	85	88	61	Sebastiani 08 [236]
	F≥2 0.79	82	72	87	64	
	F≥3 0.88	100/92/61/31	20/92/61/31	33/59/90/98	100/95/85/78	
Hepascore	F≥2 0.79	54/33	84/92	78/81	64/57	Leroy 07 [229]
	F≥3 0.85	77/47	81/90	62/65	90/81	
	F≥2 0.82	89-	63	-	-	Adams 05 [241]
	F≥3 0.90	-	-	-	-	
	F4 0.89	71	89	-	-	
	F≥2 0.81	82	65	70	78	Becker 09 [242]
	F≥2 0.76	77	63	59	80	Halfon 07 [243]
Fibrometer	F≥2 0.86	-	-	-		Leroy 07 [229]
	F≥3 0.91	-	-	-	-	
	F≥2 0.88	-	-	-	-	Cales 05 [222]
FibroSpect	F≥2 0.83	-	-	74	76	Patel 04 [244]
	F≥2 0.83	72	74	61	82	Zaman 07 [245]
	Ishak≥4 0.92	75	-	88	86	Mehta 08 [225]
ELF	F≥2 0.77	95/80/30	29/58/99	28/35/90	95/91/83	Rosenberg 04 [246]
	F≥2 0.83	-	-	-	-	Cales 05 [222]
	F≥3 0.85	95/70/54	44/85/95	44/68/82	95/86/81	Parkes 11 [247]
	Ishak>2 0.82	90	65	68	89	Cobbold 09 [248]
	Ishak>4 0.90	79	87	61	94	

Table 3.13 Diagnostic performance of serum markers for the non-invasive Detection of Fibrosis in NAFLD

Panel	AUROC	Sensitivity %	Specificity %	PPV %	NPV %	Ref
ELF	F≥1 0.76	61	80	81	79	Guha 07 [249]
	F≥2 0.82	70	80	70	80	
	F≥3 0.90	80	90	71	94	
Fibrotest-Fibrosure	F≥2 0.71	59	73	60	73	Adams 11 [250]
	F≥3 0.80	61	90	63	89	
	F4 0.86	73	92	49	97	
	F≥2 0.81	77	77	54	90	Ratziu 06 [251]
	F≥3 0.88	15	98	73	76	
Hepascore	F≥2 0.73	51	88	74	73	Adams 11 [250]
	F≥3 0.81	76	84	57	92	
	F4 0.91	87	89	45	99	
NAFLD fibrosis score	F≥3 0.82	77	71	52	88	Angulo 07 [252]
	F≥2 0.69	-	-	-	-	Shah 09 [253]
	F≥3 0.77	-	-	-	-	
Fibrometer	F≥2 0.94	79	96	88	92	Calès 09 [254]
FIB-4	F≥2 0.74	54	88	76	74	Adams 11 [250]
	F≥3 0.86	74	87	61	92	
	F4 0.86	73	89	40	97	
	F≥2 0.75	-	-	-	-	Shah 09 [253]
	F≥3 0.80	-	-	-	-	
APRI	F≥2 0.73	71	70	61	78	Adams 11 [250]
	F≥3 0.79	72	77	47	91	
	F4 0.75	77	71	22	97	
	F≥2 0.70	-	-	-	-	Shah 09 [72]
	F≥3 0.73	-	-	-	-	
BARD	F≥2 0.61	44	70	50	65	Adams 11 [250]
	F≥3 0.70	60	72	37	87	
	F4 0.75	52	84	25	94	
	F≥2 0.68	-	-	-	-	Shah 09 [253]
	F≥3 0.70	-	-	-	-	
	F≥2 0.94	79	96	88	92	Harrison08[255]
AST/ALT	F≥3 0.83	74	78	44	93	McPherson 10 [256]
	F≥3 0.74	-	-	-	-	Shah 09 [253]

NITs for fibrosis developed using Genomic and Proteomic Studies

In contrast to hypothesis driven candidate biomarker studies, both genomics and proteomics facilitate the discovery of novel biomarkers through hypothesis-free studies. The development of inflammation and fibrosis in an individual is the result of the complex interplay between environmental factors and host factors. Genomic, transcriptomic and proteomic research has helped elucidate the host factors (genes, gene expression and protein expression) that contribute to the development of a disease state such as fibrosis or NASH. Hitherto, multiple genomic and proteomic studies have been performed to identified biomarkers of fibrosis, steatosis and NASH using *in vitro*, animal and human studies.[257, 258] Unfortunately, the results thus far have been relatively disappointing. As compared to their existing hypothesis-driven candidate biomarker counterparts, potential biomarkers identified through genomic and proteomic studies have demonstrated either inferior (or at best similar) diagnostic accuracy. Moreover, these studies thus far have yielded heterogenous findings; this may relate to methodological factors such as sample preparation and patient selection. Proteomic studies are challenging because unlike DNA and mRNA, proteins cannot be amplified. Thus far attempts to identify potential biomarkers using mass spectrometry have been hindered by the presence of abundant urine or serum proteins. In NASH, genomic and proteomic studies have interestingly identified targets relating to a variety of pathways including lipid metabolism, extracellular matrix remodeling, apoptosis and liver regeneration. However, these potential biomarkers have not been able to reliably discriminate between simple steatosis and NASH. In addition to novel biomarkers the components of several biomarker panels (α-2-macroglobulin, haptoglobin and apolipoprotein A1) have also been identified by proteomic studies as being associated with fibrosis.

Imaging based modalities to assess liver fibrosis

Imaging modalities have measured liver stiffness to detect liver fibrosis. The use of liver stiffness as a modality by which liver fibrosis is assessed was derived from the observation that increasing liver fibrosis results in an increase in liver stiffness. Whereas serum markers of fibrosis may not be entirely liver specific, liver stiffness itself is importantly an intrinsic physical property of liver tissue. The most widely studied modality for assessing liver stiffness is transient elastography. Other methods to assess liver stiffness include acoustic radiation force impulse imaging (ARFI), 2D-shear wave elastography (2D-SWE) and magnetic resonance elastography (MRE). On an individual patient basis, it can be more time consuming to assess liver fibrosis using an imaging modality than by using a serum marker. Whereas scanning itself may take only a few minutes the procedure time may be prolonged due to factors such as patient positioning and patient compliance. By contrast, serum markers of fibrosis can be measured at the time of routine blood tests. Machines used to assess liver stiffness are expensive and require formal training (albeit relatively short) prior to patient use for their use. For these reasons, imaging modalities that assess liver stiffness are more suited for use in a secondary care setting.

Transient elastography (Fibroscan)

Liver stiffness measurement using Fibroscan® (Echosens, Paris, France) ultrasound-based transient elastography (TE) was first described in 2003. Measurements are taken with a patient lying reclined with the right upper quadrant exposed. Thereafter, a transducer mounted on a hand-held probe is applied perpendicular to the skin and in between the ribs targeting the right lobe of the liver. Activation of the probe produces

low frequency (50Hz) mild amplitude vibrations; these vibrations generate mechanical waves that propagate through the liver. Liver stiffness (LS) is proportional to the velocity of these propagating waves; wave velocity is measured using by pulse-echo acquisition. LS is a surrogate for fibrosis: patients without hepatic fibrosis will demonstrate the slowest wave velocities whereas patients with cirrhosis will demonstrate the fastest velocities. The LS measurement itself is measured not over the entire liver but instead a 1cm x 4cm volume of liver tissue between 25mm and 65mm below the skin surface. Whereas this is much larger than a liver biopsy specimen (hundred-fold) it is still 1/500th the size of the liver itself. TE measures liver stiffness in kilopascals (kPA) with a range between 2.5-75kPa; normal values are approximately 5 kPa. Criteria for a valid TE assessment are: (1) at least ten attempts at liver stiffness measurement (2) 60% success rate (percentage of successful liver stiffness acquisitions) (3) an interquartile range below 30% of the median stiffness value.[259]

A summary of the studies of TE in the assessment of liver fibrosis in CHC are outlined in table 3.14. In keeping with contemporaneous serum marker research, the early studies of TE focused on fibrosis discrimination in CHC. Thereafter TE has been validated in other liver diseases such as Hepatitis B and NAFLD. A summary of the studies of TE in the assessment of liver fibrosis in NAFLD are outlined in table 3.15.

Several meta-analyses have been published on the use of TE in discriminating fibrosis stages. For the detection of cirrhosis TE exhibits excellent performance. For the detection of moderate fibrosis (F2-4) the diagnostic accuracy of TE has generally been good (AUROC>0.8). Similar findings have been observed in CHC and other liver diseases.

Overall, TE is an effective modality for the detection of liver fibrosis. TE is reproducible with scans normally taking only a few minutes in thin patients. The largest study of TE published thus far has documented a failure rate of 3.1% with unreliable LS measurements obtained in 15.8% of examinations.[260] Obesity was the main reason for technical failure with operator inexperience also a major contributing factor. A significant failure rate continues even when using a probe specifically designed for obese patients.[261]

Other causes for failure include patients with narrow intercostal spaces and the presence of ascites. Moreover, increasing liver fibrosis is not the only cause for increasing liver stiffness. Other factors such as hepatic inflammation, congestive heart failure, alcohol intake and infiltration (iron overload) can give rise to an increase in liver stiffness resulting in erroneous results. Moreover, as TE assesses a small part of the right lobe only, significant geographic variation of parenchymal disease can also under stage or over stage fibrosis. Finally, TE should not be performed within 2 hours of eating as during this period liver stiffness will be transiently raised.

Table 3.14 Diagnostic performance of TE for detecting F2 and F4 in CHC

Author	Year	Patient (n)	AUROC F2	% F2	Threshold F2 (kPa)	Sens. F2	Spec. F2	AUROC F4	% F4	Threshold F4 (kPa)	Sens. F4	Spec. F4
Castera *et al*.[262]	2005	183	0.83	74	7.1	67	89	0.95	25	12.5	87	91
Ziol *et al.*[263]	2005	251	0.79	65	8.6	56	91	0.87	19	14.6	86	96
Arena *et al*.[264]	2008	150	0.91	56	7.8	83	82	0.98	19	14.8	94	92
Lupsor *et al*.[265]	2008	324	0.86	65	7.4	76	84	0.94	21	11.9	87	91
Wang *et al*.[266]	2009	214	0.82	42	9.5	70	83	0.93	19	12	79	85
Degos *et al*.[267]	2010	93	0.75	6	5.2	90	32	0.90	14	12.9	72	89
Zarski *et al*.[268]	2012	382	0.82	47	5.2	97	35	0.93	14	12.9	77	90

Table 3.15 Diagnostic performance of TE for detecting F2 and F4 in NAFLD

Author	Year	Patient (n)	AUROC F2	% F2	Threshold F2 (kPa)	Sens F2	Spec F2	AUROC F4	% F4	Threshold F4 (kPa)	Sens F4	Spec F4
Yoneda *et al*.[269]	2008	97	0.86	50	6.6	88	74	0.99	9	17.0	100	97
Nobili *et al*.[270]	2008	50	0.99	24	7.4	100	92	-	-	-	-	-
Lupsor *et al*.[271]	2010	72	0.79	25	6.8	67	84	-	-	-	-	-
Wong *et al*.[272]	2010	246	0.84	41	7.0	79	76	0.95	10	10.3	92	88
Gaia *et al.*[273]	2011	72	0.80	46	7.0	76	80	0.94	12.5	10.5	78	96
Petta *et al*.[274]	2011	169	0.79	47	7.25	69	70	-	-	-	-	-
Myers *et al*.[275]	2012	75	0.86	-	7.8	84	79	0.88	-	22.3	80	91
Wong *et al*.[276]	2012	193	0.83	45	7.0	79	64	0.83	13	10.3	81	83

Ultrasound Elastography

Ultrasound Elastography is a modality that combines ultrasonography with an evaluation of liver stiffness. As ultrasound elastography employs real time abdominal imaging, confounding anatomic structures (for example overlying bowel or blood vessels) can be avoided thereby targeting liver parenchyma itself for stiffness evaluation. Ultrasound Elastography was developed after the development of transient elastography and many studies have compared the diagnostic performance of these two modalities for fibrosis discrimination. The main form of ultrasound elastography that is currently being used in clinical practice is ARFI. ARFI can be performed on standard ultrasound equipment (Siemens and Phillips). In contrast to TE, the parenchymal area that is assessed by ARFI is smaller in size (10mm by 6mm) but this area can be targeted. The reproducibility of ARFI is good but stiffness values are subject to the same confounding factors as TE. Similarly to TE, ARFI detects cirrhosis more accurately than mild or advanced fibrosis.

Magnetic Resonance Imaging (MRI)

MRI, a modality commonly used for cross sectional liver and abdominal anatomical imaging, has been studied as a NIT for fibrosis staging by using the characteristic water-diffusion abnormalities with cirrhosis and LS measurement. Of the several types of MRI that have been developed as a NIT for fibrosis staging, magnetic resonance elastography (MRE) is the most widely studied. MRE is performed by generating propagating waves within the liver.[277] Thereafter, the subsequent micron level displacements are imaged using a special MRI technique emptying motion-sensitising gradients. These displacements are then processed using an algorithm to generate a

virtual landscape of liver fibrosis. In contrast to both TE and ARFI, MRE provides information on the spectrum of fibrosis in the entire liver. Limitations of MRE include the high cost and its decreased accuracy in the context of iron overload. Moreover, MRE is poorly tolerated by patients who suffer with claustrophobia. Studies thus far have yielded conflicting results as to whether MRE exhibits better performance than TE for the detection of liver fibrosis.

Other Diagnostic Targets of NIT In CLD

NIT for Hepatic Venous Pressure Gradient

Both TE and serum markers of liver fibrosis have been identified as effective NITs for the cross-sectional detection of clinically significant (HVPG>10mmHg) and severe portal hypertension (HVPG>12mmHg). A meta-analysis of TE in the detection of clinically significant portal hypertension amongst patients with cirrhosis identified excellent diagnostic performance (AUROC 0.93).[278] Therefore, when applied in an appropriate clinical context, these NITs could be used to identify which patients do not require endoscopic screening for oesophageal and gastric varices. Using these findings, recent Baveno VI criteria have also defined patients with compensated cirrhosis in whom endoscopy can be avoided as using NIT results (those with a liver stiffness (LS) by transient elastography <20 kPa and a platelet count >150,000/mm^3).[197]

Both indirect and direct serum markers of fibrosis have also demonstrated good performance in the detection of clinically significant portal hypertension including ELF [279] and Fibrotest.[280]

NIT to ascertain prognosis in CLD

The morbidity and mortality associated with CLDs is directly related to severe portal hypertension, HCC and liver failure. In contrast to histologic staging with its inherent limitations, liver related events (LRE) are a robust clinical endpoint that can be used as an outcome variable in NIT studies. Analogous to their relationship with HVPG, the relationship between NITs for liver fibrosis and prognosis have been explored. Cross sectional studies have identified that that NIT values at a single time point are

associated with the future development of LRE. Indeed, several NITs for fibrosis have demonstrated better performance than histologic staging itself for ascertaining prognosis including TE[281], Fibrotest[282] and ELF.[283, 284] An explanation for the apparent superiority of these NITs over conventional histology in conferring prognosis may lie in the observation that NIT values continue to rise in the face of increasing portal pressure and inflammation. By contrast, conventional histologic staging reaches a plateau at the stage of cirrhosis regardless of ongoing pathophysiologic changes. Recent studies have also identified that a dynamic increase in liver stiffness (≥1.5kPA per year[285] or ≥1 kPA per year[286]) occurring on a background of at least moderately raised stiffness measurements is a poor prognostic indicator; these findings have not been reproduced as yet using serum markers.

Using NITs to assess response to antifibrotic therapy

NIT values (predominantly TE and Fibrotest) have been studied prior to and after antiviral therapy for CHC and CHB.[287] These studies have identified that Fibrotest and LSM values decreased in virologic responders in comparison to non-responders. Three studies (2 in CHC and 1 in CHB) have also examined both biopsy specimens and Fibrotest both before and after antiviral treatment. In these studies, there was a non-significant trend towards concordance between the two modalities after categorising patients into those with and without virologic response. However, the evolution of NIT values during putative antifibrotic therapy over more than two time points has not been demonstrated thus far (with or without using histology as a reference standard). Furthermore, for studies investigating putative antifibrotic therapy in cirrhotic patients, liver related outcomes are also desirable outcome measures given the inherent ongoing risk of decompensation. Thus far, changes in NIT scores arising

during antifibrotic therapy have not yet been studied for their association with liver related outcomes.

Feeding and Inflammation as Confounding Factors for the assessment of fibrosis using NITs

To maximise biomarker accuracy, subjects may need to be fasted when NIT readings are taken. For example, ELF test scores can rise transiently for up to 2 hours after ingesting food; this is primarily due to a rise in hyaluronic acid. As a result, it is recommended that patients should be fasted for 2 hours prior to ELF testing (and maker panels employing hyaluronic acid) [288]. These caveats also apply to the use of transient elastography. Concurrent hepatic inflammation has the greatest impact on indirect markers scores but also affects direct serum markers such as ELF[289] and transient elastography. Consequently, it is essential that NIT scores are interpreted in the context of a patient's overall clinical state.

Combining NITs

Hitherto, numerous attempts have been made to combine diagnostic tests in the form of diagnostic algorithms. Notably, extensive work has been conducted in the field of staging fibrosis in CHC.[290-292] The aims of such studies are common; they include a reduction in the number of cases classified as indeterminate thus minimizing the need for liver biopsy whilst attempting to preserve or increase diagnostic accuracy. NITs for liver fibrosis including both serum markers and imaging modalities have been employed both in series and in parallel algorithms. However, many of these algorithms have been created without consideration of the intrinsic diagnostic attributes of the biomarkers such a good diagnostic performance as assessed by AUROC but also robust performance at their diagnostic thresholds. Recently, NITs parallel algorithms were studied in NAFLD combining indirect markers with transient elastography.[293] In this study, the application of parallel algorithms increased both PPV and NPV.

NITS for the detection of Steatosis and NASH in NAFLD

Following the widespread success of the non-invasive detection of liver fibrosis, focus has also turned to the diagnostic targets of both steatosis and NASH in NAFLD. Given the current obesity epidemic, the identification of individuals with NAFLD and NASH amongst the at-risk population is challenging without the use of sensitive and specific NITs.

NITs for Steatosis

A summary of NIT studies for the non-invasive diagnosis of steatosis is presented in table 3.16. The reference standard used to derive a NIT for steatosis has generally been transabdominal ultrasound with a single study has using histologic staging and another using magnetic resonance spectrometry. Overall, NIT for steatosis have a demonstrated AUROC of approximately 0.8. Diagnostic thresholds have been proposed for NIT enabling sensitivity of between 80-90%. However, the use of ultrasound as a reference standard for the presence of steatosis is questionable. A meta-analysis identified a AUROC of 0.93 for the performance of ultrasound in detecting moderate to severe steatosis.[294] Ultrasound is, however, operator dependent and it is less sensitive when steatosis infiltration is lower than 30%. Other imaging modalities (CT, MR, Elastography) have also been employed as NITs for steatosis.

Table 3.16 Summary of NIT Studies for the non-invasive diagnosis of Steatosis

Marker	Reference for Steatosis	Parameters	Number of Patients	Patients' characteristics	AUROC
SteatoTest[295] 2005	Histologic Steatosis >5%	ALT, A2M, ApoA1, haptoglobin, bilirubin, γGT, cholesterol, triglycerides, glucose, age, gender, BMI	310	Training Cohort Mixed aetiology	0.79
			171	Validation: HCV pre-therapy	0.80
			201	Validation: HCV- SVR	0.86
			62	Validation: Alcohol-related liver disease	0.72
Fatty liver index (FLI)[296] 2006	Steatosis assessed by ultrasound	BMI, waist circumference, triglycerides and γGT	496	Patients with and without elevated ALT and GGT	0.84
Hepatic Steatosis Index (HSI) [297] 2010	Steatosis assessed by ultrasound	AST, ALT, BMI, Diabetes	10724	Derived during routine patient check up	0.81
Lipid Accumulation Product (LAP)[298] 2010	Steatosis assessed by ultrasound	Waist circumference, triglycerides, gender	588	Patients without suspected CLD	0.79
Index of NASH (ION)[299] 2014	Steatosis assessed by ultrasound	♂: waist to hip ratio, triglycerides, ALT, HOMA ♀: triglycerides, ALT, HOMA	4458	Derived from NHANES population, Validated in NAFLD patients	0.77
NAFLD Fat Score[300] 2009	Steatosis assessed by MR Spectroscopy	Metabolic syndrome, diabetes, insulin, ALT, AST	470	Patients with metabolic syndrome	0.87

NITs for Non-alcoholic Steatohepatitis

The reference standard for NIT test derivation for the diagnostic target of NASH has either been a histopathologic diagnosis of NASH or a NAS score of ≥5. The NAS is a semi-continuous score enabling changes in disease severity to be documented. However, as described earlier a NAS score ≥5 is not fully concordant with a histologic diagnosis of NASH. As a result some patients with NAS score ≥5 do not have NASH and may ultimately have an excellent prognosis. Accordingly, NITs for steatohepatitis should be studied for their association with both NAS≥5 and a histologic diagnosis of NASH. A summary of the studies for NITs for NASH are presented in table 3.17.

- **Cytokeratin 18 (CK-18)**

The most studied NIT for NASH is CK-18. Cytokeratins (CKs) are normal constituents of epithelial cell cytoskeleton. CK-18 an intermediate filament representing about 5% of total protein in the liver. Studies have identified that cellular release of CK-18 fragments occurs as a consequence of caspase digestion during the intermediate stage of apoptosis and thereby can serve as markers of apoptosis.[301] As NASH is associated with an increased apoptosis it was proposed that serum CK18 fragments that may therefore distinguish NASH from simple steatosis. A validation study described good performance of caspase-cleaved CK-18 in discriminating NASH from simple steatosis (SS) with AUROC of 0.83 in a cohort of patients with advanced fibrosis.[302] Thereafter, numerous validation studies have been performed which have demonstrated a wide variation in the diagnostic performance of CK-18 for detecting NASH. A meta-analysis of 11 studies that investigated CK-18 as a NIT of NASH identified that AUROC ranged 0.71-0.93, sensitivity was 66%, and specificity was 82%.[303] A subsequent recent study of 424 patients identified poorer diagnostic performance of CK-18 for the detection of NASH with the respective AUROC,

sensitivity and specificity only 0.65, 58% and 68%.[304] These findings suggest that the diagnostic performance of CK-18 is insufficient for enable its use as a standalone NIT for NASH but it certainly could be employed as a constituent component of a biomarker panel.

- **Fas ligand**

The Fas ligand is a hepatic transmembrane protein that induces apoptosis. Increased values of Fas ligand have been demonstrated in NASH.[305] In a 2011 study, a prediction model composed of Fas ligand and CK-18 demonstrated excellent diagnostic performance in the derivation cohort (0.93) but only fair performance a validation cohort (AUROC 0.79).[306]

- **NASHTest**

NASHTest is a combination of 13 biochemical and clinical variables (age, sex, height, weight, triglycerides, cholesterol, α2-macroglobulin, apolipoprotein A1, haptoglobin, γGT, ALT, AST, bilirubin) to predict the presence or absence of NASH.[307] Whereas the overall diagnostic accuracy of NASHtest has been moderate (AUROC 0.79 in both training and validation cohorts), its use as a NIT for identifying NASH is limited by poor sensitivity (39% in training cohort and 29% in validation cohort); these findings have been confirmed in numerous further validation studies.

- **NAFIC Score**

In a 2011 study of 177 patients, serum ferritin, fasting insulin and type IV collagen 7S were selected as independent predictors of NASH, by logistic regression analysis. These three variable were combined to form the NAFIC score; the AUROC was found to be good (0.85) in the derivation cohort but fair (0.78) in a validation cohort.[131]

- **NASH Clinical score for morbid obesity**

A 2008 study investigated the association of demographic, clinical and laboratory variables for their association with a diagnosis of NASH in morbidly obese patients.[308] Using logistic regression, 6 factors were combined to form a scoring system for NASH: hypertension, type 2 diabetes, sleep apnoea, AST > 27 IU/L ALT> 27 IU/L, and non-Black race. The AUROC for the score was found to be 0.80 but the score has not been validated.

- **Palekar index[309]**

The Palekar index was derived in a study of 80 patients of which 41 had a histologic diagnosis of NASH. Using logistic regression, six variables were used to construct a predictive score for NASH (Age ≥ 50 yrs, female gender, AST ≥ 45 U/L, AST/ALT ratio ≥ 0.8, BMI ≥ 30 Kg/m2, hyaluronate ≥ 55 mcg/l) The Palekar index was able to discriminate between simple steatosis and NASH with AUROC 0.76. The index has not been validated.

- **Other studies**

A number of other studies have also attempted to discriminate between SS and NASH using serum adipokines[310], ferritin[311], markers of oxidative stress.[312, 313] While these putative biomarkers are statistically associated with a diagnosis of NASH their diagnostic accuracy is insufficient for use in clinical practice.

Table 3.17 Summary of studies investigating the use of biomarkers for the non-invasive diagnosis of NASH

NIT	Parameters	Endpoint	Patients in study	AUROC Derivation	AUROC Validation
CK-18 [302, 303,127]	Serum fragment of cytokeratin-18	NASH (Brunt)	139 NAFLD, 150 age matched controls	0.83	0.71-0.93 0.65
Fas ligand and CK-18[306]	Fas ligand and CK-18	NAS ≥ 5 & NASH (Kleiner)	95 NAFLD 82 Validation	0.93	0.79
Nash Test[307]	Age, gender, BMI, triglycerides, cholesterol, α-2-macroglobulin, GGT, AST, ALT, haptoglobin, apolipoprotein A1, total bilirubin	NAS ≥ 5 & NASH (Kleiner)	257 NAFLD 383 control group	0.79	0.79
NAFIC score[314]	Ferritin ≥ 200 ng/ml (female) or ≥ 300 ng/ml (male), immunoreactive insulin ≥ 10.0 IU/ml, and type IV collagen 7S ≥ 5.0 ng/ml	NASH (Kleiner)	177 NAFLD Drivation 442 NAFLD Validation	0.85	0.78
NASH Clinical score for morbid obesity[308]	hypertension, diabetes, AST ≥ 27 IU/L, ALT ≥ 27 IU/L, obstructive sleep apnea and nonblack race	NAS ≥ 5 & NASH Kleiner	200 obese	0.80	-
Palekar index[309]	Age ≥ 50 yrs, female gender, AST ≥ 45 U/L, AST/ALT ratio ≥ 0.8, BMI ≥ 30 Kg/m^2 , hyaluronate ≥ 55 mcg/l	NASH (Brunt)	80 NAFLD (39 SS* and 41 NASH)	0.76	

CHAPTER 4

DERIVATION AND VALIDATION OF A SERUM MARKER FOR THE DETECTION AND ASSESSMENT OF NON-ALCOHOLIC STEATOHEPATITIS IN PATIENTS WITH NAFLD

ABSTRACT

Background: Liver biopsy is the reference standard for the detection of NASH within NAFLD. Hitherto a reliable biomarker for NASH has not been identified.

Aims: The aim of this study was to identify a biomarker of NASH in patients without significant fibrosis.

Methods: 172 patients from 2 centres with biopsy proven NAFLD were included in this study. 84 patients from a single centre were included as a derivation cohort and 88 patients from a second centre were included as a validation cohort. Serum samples were tested for candidate markers of fibrosis and inflammation alongside haematological and biochemical markers.

Results: Amongst patients without advanced fibrosis, PIIINP was the only marker found to be associated with a histological diagnosis of NASH in both cohorts. PIIINP also correlated with the total NAS score and its constituent components ($p<0.001$). AUROC for PIIINP in discriminating between NASH and simple steatosis was 0.77-0.82 in patients with F0-2 fibrosis and 0.82-0.84 in patients with F0-3 fibrosis. PIIINP was elevated in patients with advanced fibrosis, the overwhelming majority of whom had NASH. When incorporating patients with all degrees of fibrosis from both cohorts, PIIINP was able to discriminate between patients with simple steatosis and those with NASH or advanced fibrosis with AUROC 0.85-0.87.

Conclusions: PIIINP discriminates between simple steatosis and NASH or advanced fibrosis. The use of a single biomarker in this context will be of clinical utility in detecting the minority of patients with NAFLD who have NASH or advanced fibrosis related to NASH.

INTRODUCTION

NAFLD is the hepatic manifestation of the metabolic syndrome and is now one of the leading causes of liver disease worldwide.[315, 316] The clinical spectrum of NAFLD encompasses SS, NASH and progressive liver fibrosis.[317, 318] It is well recognised that the development of fibrotic liver disease in NAFLD is attributed to the progression of SS to NASH. Increasing literature also suggests that patients with histological NASH follow a progressive course which may result in cirrhosis, portal hypertension, liver failure and hepatocellular carcinoma.[319, 320] Prior to the development of advancing fibrosis, NASH is thus the most important prognostic feature of liver damage in NAFLD. The early identification of NASH is important for stratifying patients for therapeutic intervention prior to the development of significant hepatic fibrosis. Histological staging of a liver biopsy is the reference standard used for stratifying patients into those currently with benign phenotype (SS) and those at immediate risk of progressive disease (NASH with developing fibrosis). With the rising prevalence of obesity, up to 30% of the population is at risk of NAFLD of whom 3-5% will be have NASH, and only 1-2% will have progressive liver fibrosis.[4] Liver biopsy cannot be used as the primary method to detect and quantify NASH and fibrosis in the wider population at risk of NAFLD due to practical and ethical reasons because it is invasive, hazardous, resource intensive and costly.[321]

The non-invasive assessment of patients with NAFLD at risk of progressive liver disease has focussed on one of two approaches: either to identify fibrosis at an early stage or to identify NASH prior to the development of significant fibrosis. With regard to the detection of developing liver fibrosis, numerous non-invasive tests have been developed for the detection of fibrosis in NAFLD. However, whilst their performance in the detection of advanced fibrosis ranges from good to excellent, their performance in

the detection of mild liver fibrosis is in general modest.[322] For example, the Enhanced Liver Fibrosis (ELF) test[221] (HA, PIIINP, TIMP1) has been validated in a secondary care setting in the detection of fibrosis in NAFLD.[323] Whilst the detection of severe fibrosis (>F3) in this study was excellent (AUROC 0.90) the performance in the detection of any fibrosis was more modest (AUROC 0.76).

Non-invasive tests have also been developed for the diagnosis of NASH in NAFLD but these models have generally had less success than their fibrosis counterparts.[307, 309, 324] The aim of this study was to identify a biomarker of NASH which would be able to detect NASH prior to the development of significant hepatic fibrosis. Candidate biomarkers of hepatic inflammation (YKL-40, TIMP-1), apoptosis (CK-18) and liver fibrosis (HA, PIIINP, Collagen IV) were selected on the basis of biological plausibility and previous association with NAFLD and analysed alongside standard laboratory and clinical measurements, considered in part because of their ubiquity in clinical practice, low cost and ease of use.[211, 216, 325-329]

METHODS

The patients included in this study were recruited consecutively from outpatient clinics in 2 hepatology centres in the United Kingdom. The derivation and validation cohorts were comprised of patients recruited from the Queen's Medical Centre, Nottingham and the Freeman Hospital, Newcastle-Upon-Tyne respectively. In both cohorts NAFLD was diagnosed using the following criteria: (1) elevated aminotransferases (AST or ALT); (2) appropriate exclusion of 'other' causes of liver disease including alcohol, drugs, autoimmune or viral hepatitis, or cholestatic or metabolic/genetic liver disease. The 'other' causes of liver disease were excluded using specific clinical, biochemical, radiographic and histological criteria. All patients had a weekly ethanol consumption of less than 140g in women and 210g in men. The patients included in this study all had liver biopsies at the centres between 1998 and 2006 and histology was consistent with NAFLD. The serum samples used in the study were taken within 3 months of biopsy. Height and weight was recorded on all patients and body mass index (BMI) calculated. Serum sample samples were obtained for routine biochemistry (including ALT, AST, GGT, bilirubin, albumin and alkaline phosphatase), full blood count, measurements of insulin resistance (fasting glucose and insulin), ferritin and cholesterol. Serum samples were analysed for levels of TIMP-1, HA, PIIINP, Collagen IV and YKL-40 at an independent reference laboratory (iQur Limited, Southampton, UK). Stored sera taken from patients in the derivation cohort (Nottingham) were also tested for caspase- cleaved CK-18 (M30 Apoptosense ELISA assay, Peviva, Sweden).

Liver biopsy

Liver biopsies were assessed by a single senior liver histopathologist at each centre (Nottingham PK, Newcastle AB) and scored for histological grade and fibrosis stage using the classification system described by Kleiner et al.[190] The grade of NASH was assessed using the NAS scored from (0-8) which incorporates the scores of steatosis (0-3), ballooning (0-2) and lobular inflammation (0-3).

In addition to using the Kleiner scoring system, both histopathologists classified the biopsies in a dichotomous manner as either ‘definite NASH’ or ‘borderline NASH / simple steatosis’ using agreed criteria.

Analysis

Patients with F3 & F4 fibrosis in both cohorts were excluded from the initial analysis as the aim of the study was to assess the performance of tests in the detection of patients with NAFLD who have NASH prior to the development of significant liver fibrosis. Furthermore, it is well described that patients with NAFLD and advanced fibrosis may no longer exhibit the necroinflammatory activity or steatosis that was present earlier in their disease course.[330] Nineteen and 17 patients (total 36) were excluded from the initial analyses of the derivation and validation cohort respectively due to the presence of advanced fibrosis.

Statistical Analysis

Data were analysed using SPSS v. 20.0 (SPSS Inc., Chicago, IL) and STATA v. 11.0 (SAS Inc). Patients without advanced fibrosis (F0-2) were separated into 2 categories based on their final histological diagnosis of either ‘definite NASH’ or ‘borderline NASH / simple steatosis’ and clinical and laboratory variables were compared between categories. Categorical variables including sex and presence of diabetes were

compared between the two groups and the Chi-square test was applied to test significance. Parametric continuous variables and non-parametric continuous variables were compared between the two groups using 2-sided students t-test and the Mann-Whitney U Test respectively.

Binary logistic regression was used to assess the association of variables with a histological diagnosis of NASH. Correlation coefficients (Spearman's rho) were used to determine the strength of the relationship of the multiple variables to the NAS score and its individual components in patients without advanced fibrosis. ROC curves together with 95% confidence intervals were plotted to determine the performance of the identified variable(s). AUROC curves were compared using the method of Delong.[331] Sensitivity and specificity, predictive values, likelihood ratios (LR+ & LR-) and diagnostic odds ratios (OR) were calculated at thresholds derived from ROC curves. The identified parameter(s) in the initial analyses were then compared with the NASH scores and fibrosis stages of patients in the entire cohorts.

RESULTS

Derivation and validation cohorts

The baseline characteristics of both cohorts are shown in tables 4.1 and 4.2 and figure 4.1. Whilst the distribution of gender, age and diabetes was similar within both cohorts, patients in the derivation cohort were found to have significantly lower BMI ($p<0.001$), more severe fibrosis ($p=0.022$), a greater prevalence of NASH and higher NAS scores ($p<0.001$). The majority of patients in the validation cohort had either minimal or no fibrosis. In contrast, the distribution of fibrosis in the derivation cohort was more analogous to a secondary care population with all stages of fibrosis well represented.

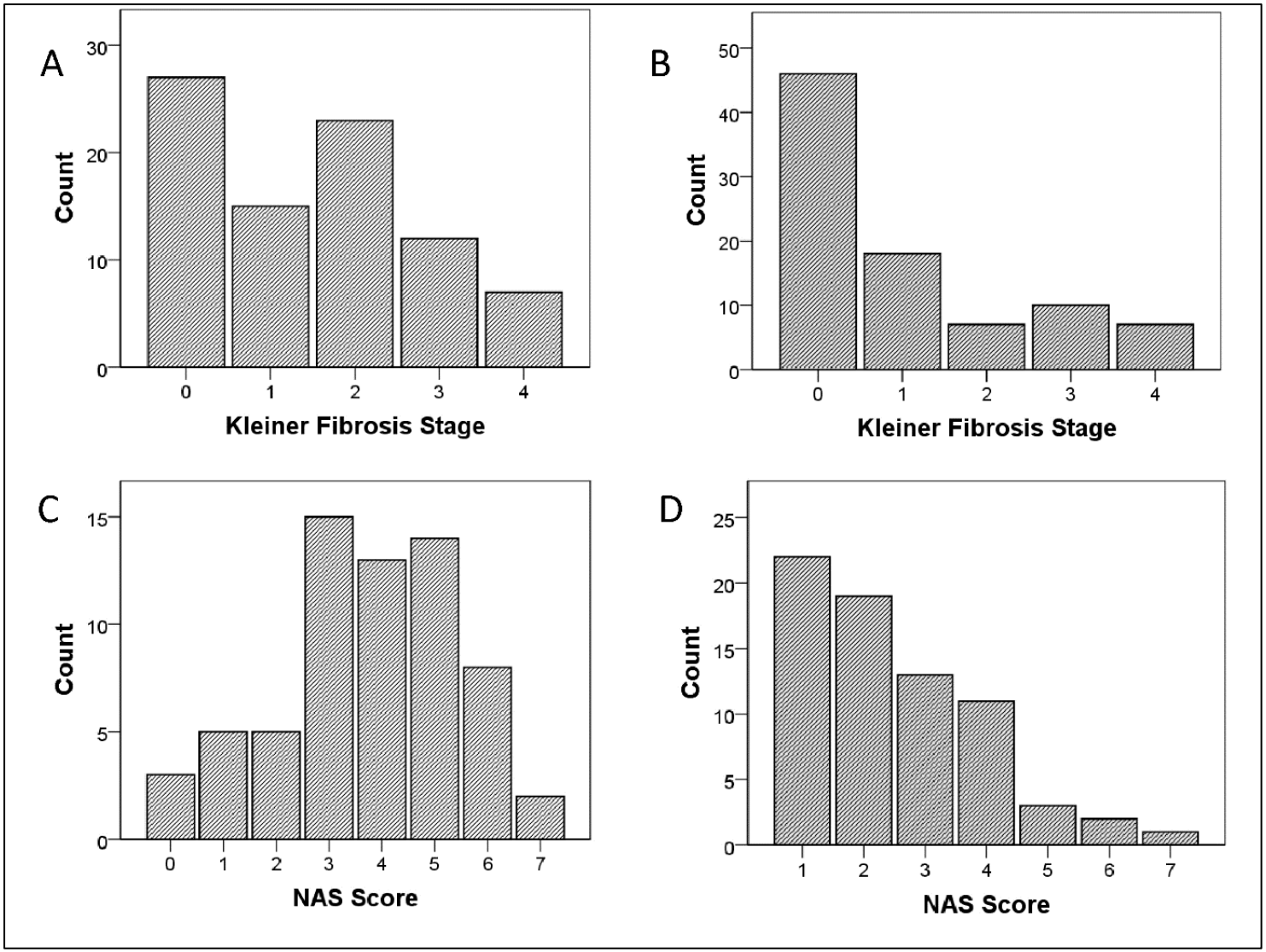


Figure 4.1 Histograms of the fibrosis stages and NAS score in both cohorts

Distribution of fibrosis stages in the derivation (Panel A) and validation (Panel B) cohorts.

Distribution of NAS score in patients without advanced fibrosis (F0-F2) in derivation (Panel C) and validation (Panel D) cohorts.

Table 4.1 Derivation Cohort: Demographic and Baseline data of patients without advanced fibrosis (F0-2)

Variable	Simple Steatosis/ Borderline NASH n=11	NASH n=54	*P value*	Overall n=65	Univariate Regression OR	95% CI *(P value)*	Multiple Regression OR	95% CI *(P value)*
Male (n,%)	7 (64%)	14 (70%)	*NS*	45 (69.2%)				
Age (years)	48.8 ± 9.8	47.9 ± 11.5	*NS*	48.0 ± 11.1				
BMI (kg/m^2)	30.55 ± 6.17	29.69 ± 4.50	*NS*	29.83 ± 4.56				
Diabetes	7 (64%)	32 (59%)	*NS*	39 (60%)				
Trig. (mmol/L)	1.95 ± 0.76	1.99 ± 1.19	*NS*	1.98 ± 1.13				
Glucose (mmol/L)	6.07 ± 2.00	5.67 ± 1.34	*NS*	5.74 ± 1.47				
HOMA-IR	2.83 (IQR 2.11)	2.86 (IQR 2.85)	*NS*	2.86 (IQR 2.53)				
ALT (IU/L)	53 (IQR 31)	64.0 (IQR 50.8)	*NS*	61.0 (IQR 38.5)				
GGT (IU/L)	131 ± 111	120 ± 111	*NS*	116 ± 82				
Ferritin (ng/ml)	**55 (IQR 63.5)**	**179 (IQR 200)**	***0.029***	**147 (IQR 212)**	**1.01**	**1.00-1.02 *(0.04)***	**1.02**	**1.00-1.03 *(0.026)***
Platelets (x10^9/L)	235.0 ± 35.4	251.4 ± 70.8	*NS*	248.8 ± 66.6				
CK18 (ng/ml)	149.3 ± 60.3	208.4 ± 162.4	*NS*	156	**1.01**	**0.99-1.01 *(0.295)***		
HA (ng/ml)	30.9 (IQR 33.7)	42.3 (IQR 26.8)	*NS*	21.4 (IQR 28.1)				
TIMP-1 (ng/ml)	**619 ± 54**	**755 ± 133**	***0.002***	**733 ± 134**	**1.01**	**1.00-1.03 *(0.005)***		
PIIINP (ng/ml)	**5.36 ± 2.05**	**8.92 ± 2.93**	***0.004***	**7.03 ± 2.90**	**2.22**	**1.26-3.91 *(0.006)***	**3.00**	**1.40-6.50 *(0.005)***
YKL-40 (ng/ml)	81.8 (IQR 67.6)	92.0 (IQR 77.1)	*NS*	84.9 (IQR 82.5)				
Coll IV (ng/ml)	**145 ± 16**	**177 ± 35**	***0.005***	**172 ± 35**	**1.04**	**1.00-1.07 *(0.009)***		
NAS Score	1.2 ± 1.1	4.3 ± 1.3		3.8 ± 1.7				

Significant Associations (P value<0.05) are displayed in bold

Table 4.2 Validation Cohort- Demographic and Baseline data of patients without advanced fibrosis (F0-2)

Variable	Simple Steatosis/ Borderline NASH n=51	NASH n=20	*P value*	Overall n=71	Univariate Regression OR	95% CI *(P value)*	Multiple Regression OR	95% CI *(P value)*
Male (n,%)	40 (78%)	12 (60%)	*NS*	52 (73%)				
Age (years)	44.7 ± 11.3	43.8 ± 14.3	*NS*	44.3 ± 12.1				
BMI (kg/m^2)	33.3 ± 5.9	36.4 ± 5.1	*NS*	34.20 ± 5.85				
Diabetes	28 (54%)	10 (50%)	*NS*	33 (47%)				
Trig. (mmol/L)	2.81 ± 1.99	2.62 ± 1.52	*NS*	2.76 ± 1.87				
Glucose (mmol/L)	5.53 ± 1.09	5.67 ± 0.87	*NS*	5.58 ± 1.02				
HOMA-IR	3.38 (IQR 4.91)	5.36 (IQR 9.48)	*NS*	3.58 (IQR 4.83)				
ALT (IU/L)	58.0 (IQR 47.0)	69.0 (IQR 34.0)	*NS*	66.0 (IQR 46.0)				
GGT (IU/L)	93.4 ± 53.5	68.8 ± 49.5	*NS*	86 ± 53			**0.97**	**0.95-0.99 (0.02)**
Ferritin (ng/ml)	125 (IQR 150)	180 (IQR 171)	*NS*	143 (IQR 152)				
Platelets (x10^9/L)	264.1 ± 53.4	252.9 ± 60.6	*NS*	261.0 ± 55.6				
CK18 (ng/ml)	NA	NA	*NA*	NA				
HA (ng/ml)	25.5 ± 19.9	30.5 ± 19.2	*NS*	26.9 ± 19.7				
TIMP-1 (ng/ml)	709 ± 121	730 ± 120	*NS*	715 ± 121				
PIIINP (ng/ml)	**6.4 ± 2.73**	**10.2 ± 6.58**	***0.001***	**7.52 ± 4.48**	**1.29**	**1.07-1.56 (0.009)**	**1.44**	**1.14-1.82 (0.002)**
YKL-40 (ng/ml)	44.4 (IQR 56.1)	73.2 (IQR 92.1)	*NS*	54.0 (IQR 57.1)				
Coll IV (ng/ml)	147 (IQR 43)	161 (IQR 39)	*NS*	149 (IQR 43)				
NAS Score	1.82 ± 0.87	4.20 ± 1.20						

Significant Associations (P value<0.05) are displayed in bold

PIIINP is the only marker associated with a histological diagnosis of NASH in both cohorts after multivariate analysis

Within the derivation cohort, univariate analysis identified a significant association between a histological diagnosis of NASH and serum ferritin, TIMP-1, PIIINP and Collagen IV. However, only PIIINP and ferritin were significantly associated with a histological diagnosis of NASH after accounting for the remaining variables as potential confounders. Within the validation cohort only PIIIINP was significantly associated with a histological diagnosis of NASH on univariate analysis. However, multivariate analysis identified that both PIIINP and GGT were associated with a histological diagnosis of NASH.

PIIINP correlates with NAS Score and its constituent components

Given the positive association of PIIINP with steatohepatitis in both cohorts, the relationship of PIIIINP with both the NAS score and its constituent components was explored. Within both cohorts PIIINP was found to correlate significantly with the NAS score, the degree of steatosis, ballooning and lobular inflammation (tables 4.3 and 4.4).

Table 4.3 Derivation Cohort: Univariate Correlation (Spearman's Rho) of variables with NAS score and its components

Variable	NAS Score		Steatosis		Ballooning		Lobular Inflammation	
	Univariate Correlation (r)	*P value*	Univariate Correlation (r)	*P value*	Univariate Correlation (r) with Ballooning	*P value*	Univariate Correlation (r) with Lobular Inflammation	*P value*
Male	-0.119	*NS*	-0.085	*NS*	-0.120	*NS*	0.082	*NS*
Age	0.094	*NS*	-0.071	*NS*	0.221	*NS*	0.051	*NS*
BMI	0.001	*NS*	0.033	*NS*	-0.051	*NS*	-0.021	*NS*
Cholesterol	**-0.294**	**0.018**	**-0.307**	***0.014***	-0.170	*NS*	-0.195	*NS*
LDL	**-0.300**	**0.019**	**-0.359**	***0.004***	-0.159	*NS*	-0.159	*NS*
Triglycerides	-0.010	*NS*	-0.007	*NS*	-0.006	*NS*	-0.006	*NS*
Glucose	0.006	*NS*	-0.102	*NS*	0.003	*NS*	0.003	*NS*
HOMA-IR	0.204	*NS*	0.127	*NS*	0.146	*NS*	0.259	*NS*
ALT	0.228	*NS*	**0.301**	***0.015***	0.127	*NS*	0.134	*NS*
GGT	**-0.342**	***0.005***	**-0.281**	***0.023***	-0.231	*NS*	**-0.265**	***0.033***
Bilirubin	-0.160	*NS*	-0.199	*NS*	-0.079	*NS*	-0.052	*NS*
Ferritin	0.181	*NS*	0.204	*NS*	0.259	*NS*	0.259	*NS*
Platelets	0.302	*NS*	0.021	*NS*	-0.029	*NS*	-0.029	*NS*
CK18	0.179	*NS*	0.050	*NS*	0.229	*NS*	0.231	*NS*
HA	0.186	*NS*	0.110	*NS*	0.144	*NS*	0.088	*NS*
TIMP-1	**0.590**	***$<10^{-6}$***	**0.397**	***0.001***	**0.542**	***$<10^{-5}$***	**0.579**	***$<10^{-6}$***
PIIINP	**0.628**	***$<10^{-7}$***	**0.455**	***0.0001***	**0.552**	***$<10^{-5}$***	**0.497**	***$<10^{-5}$***
YKL-40	0.125	*NS*	-0.009	*NS*	0.241	*NS*	0.153	*NS*
Collagen IV	**0.318**	***0.010***	0.079	*NS*	**0.418**	***0.001***	0.334	*NS*

Significant Associations (P value<0.05) are displayed in bold

148

Table 4.4 Validation Cohort: Univariate Correlation (Spearman's Rho) of variables with NAS score and its components

Variable	NAS Score		Steatosis		Ballooning		Lobular Inflammation	
	Univariate Correlation (r)	*P value*	Univariate Correlation (r)	*P value*	Univariate Correlation (r)	*P value*	Univariate Correlation (r)	*P value*
Male	-0.031	*NS*	0.067	*NS*	-0.095	*NS*	-0.047	*NS*
Age	-0.034	*NS*	-0.084	*NS*	-0.026	*NS*	0.000	*NS*
BMI	**0.294**	***0.015***	0.200	*NS*	0.334	*NS*	0.221	*NS*
Triglycerides	-0.052	*NS*	0.042	*NS*	-0.092	*NS*	-0.054	*NS*
Glucose	0.186	*NS*	0.100	*NS*	0.186	*NS*	0.176	*NS*
HOMA-IR	0.248	*NS*	0.262	*NS*	0.209	*NS*	0.118	*NS*
ALT	**0.311**	***0.008***	**0.242**	***0.042***	0.272	*NS*	0.224	*NS*
GGT	**-0.256**	***0.032***	-0.175	*NS*	-0.179	*NS*	-0.213	*NS*
Ferritin	**0.292**	***0.014***	0.154	*NS*	**0.358**	***0.002***	0.215	*NS*
Platelets	-0.035	*NS*	0.035	*NS*	-0.103	*NS*	-0.001	*NS*
HA	0.132	*NS*	-0.020	*NS*	0.221	*NS*	0.162	*NS*
TIMP-1	**0.306**	***0.010***	**0.384**	***0.001***	0.165	*NS*	0.121	*NS*
PIIINP	**0.606**	***<10^{-8}***	**0.407**	***0.0004***	**0.543**	***<10^{-6}***	**0.482**	***0.000 02***
YKL-40	0.134	*NS*	-0.013	*NS*	0.181	*NS*	0.217	*NS*
Collagen IV	**0.323**	***0.006***	**0.365**	***0.002***	0.199	*NS*	0.169	*NS*

Significant Associations (P value <0.05) are displayed in bold

PIIINP performs well in discriminating both between NASH and simple steatosis and also between differing grades of NASH in patients with F0-2 & F0-3 fibrosis

ROC curves were plotted for PIIINP levels in detecting NASH and discriminating between differing degrees of NASH (NAS score, degree of ballooning and degree of lobular inflammation) (table 4.5).

Amongst patients without advanced fibrosis (F0-2) the performance of PIIINP in discriminating NASH from simple steatosis was 0.83 and 0.77 in the derivation and validation cohorts respectively. When considering patients without cirrhosis (F0-3, the performance of PIIINP in discriminating NASH from simple steatosis was 0.84 and 0.82 in the derivation and validation cohorts respectively.

The performance of PIIINP in discriminating between differing degrees of NASH in both cohorts, both amongst patients within F0-3 was good and in particular the ability to discriminate lobular inflammation was very good (tables 4.6).

Table 4.5 PIIINP: Performance with respect to discriminating both histological NASH from 'non-NASH' and differing grades of NASH in patients with F0-2 in both cohorts.

NASH Grade	Cohort	Numbers in each group	AUROC	95% CI	*P value*	Standard Error
Histological Diagnosis 'Non-NASH' v NASH	Derivation	n=11 v n=54	0.83	0.71-0.94	*0.001*	0.058
	Validation	n=51 v n=20	0.78	0.67-0.89	*<0.001*	0.057
NAS: 0-2 v 3-8	Validation	n=51 v n=20	0.78	0.67-0.89	*<0.001*	0.057
	Validation	n=41 v n=30	0.80	0.70-0.90	*<0.001*	0.051
NAS: 0-3 v 4-8	Derivation	n=28 v n=37	0.86	0.77-0.95	*<0.001*	0.046
	Validation	n=54 v n=17	0.79	0.67-0.91	*<0.001*	0.060
NAS: 0-4 v 5-8	Derivation	n=41 v n=24	0.88	0.80-0.96	*<0.001*	0.041
	Validation	n=65 v n=6	0.80	0.61-0.98	*0.017*	0.092
NAS: 0-5 v 6-8	Derivation	n=55 v n=10	0.83	0.70-0.96	*0.001*	0.066
	Validation	n=68 v n=3	0.83	0.70-0.97	*0.052*	0.052
Lobular Inflammation: 0 v 1-3	Derivation	n=50 v n=15	0.77	0.63-0.91	*0.002*	0.072
	Validation	n=46 v n=25	0.85	0.76-0.94	*<0.001*	0.046
Lobular Inflammation: 0-1 v 2-3	Derivation	n=59 v n=6	0.89	0.74-1.00	*0.002*	0.064
	Validation	n=67 v n=4	0.86	0.75-0.97	*0.055*	0.055
Ballooning: 0-1 v 2	Derivation	n=48 v n=17	0.80	0.68-0.91	*<0.001*	0.058
	Validation	n=69 v n=2	0.76	0.64-0.88	*0.211*	0.061

Table 4.6 The ability of PIIINP to discriminate between differing grades of NASH in the F0-F3 portions of both cohorts

Discrimination of either Histological NASH or NASH Grade	Cohort	Numbers in each group	AUROC	95% CI	*P value*	Standard Error
Histological Diagnosis: 'Non-NASH' v NASH	Derivation	n=11 v n=66	0.84	0.74-0.94	*<0.001*	0.052
	Validation	n=51 v n=30	0.82	0.73-0.91	*<0.001*	0.047
NAS: 0-4 v 5-8	Derivation	n=34 v n=43	0.81	0.71-0.90	*<0.001*	0.050
	Validation	n=67 v n=14	0.82	0.69-0.94	*<0.001*	0.064
Lobular Inflammation: 0-1 v 2-3	Derivation	n=64 v n=13	0.82	0.70-0.94	*<0.001*	0.064
	Validation	n=72 v n=9	0.86	0.62-0.92	*0.009*	0.075

PIIINP levels in patients with advanced fibrosis (F3-4)

In both cohorts, the majority of patients (90-94%) with advanced fibrosis had a histological diagnosis of NASH. PIIINP levels reflected a hierarchy of liver disease severity ranging from simple steatosis with no/mild fibrosis, steatohepatitis with no/mild fibrosis to advanced fibrosis (figures 4.2 - 4. 4). This observation was made regardless of which scoring system for steatohepatitis was employed (histological NASH, NAS ≥5, lobular inflammation ≥2).

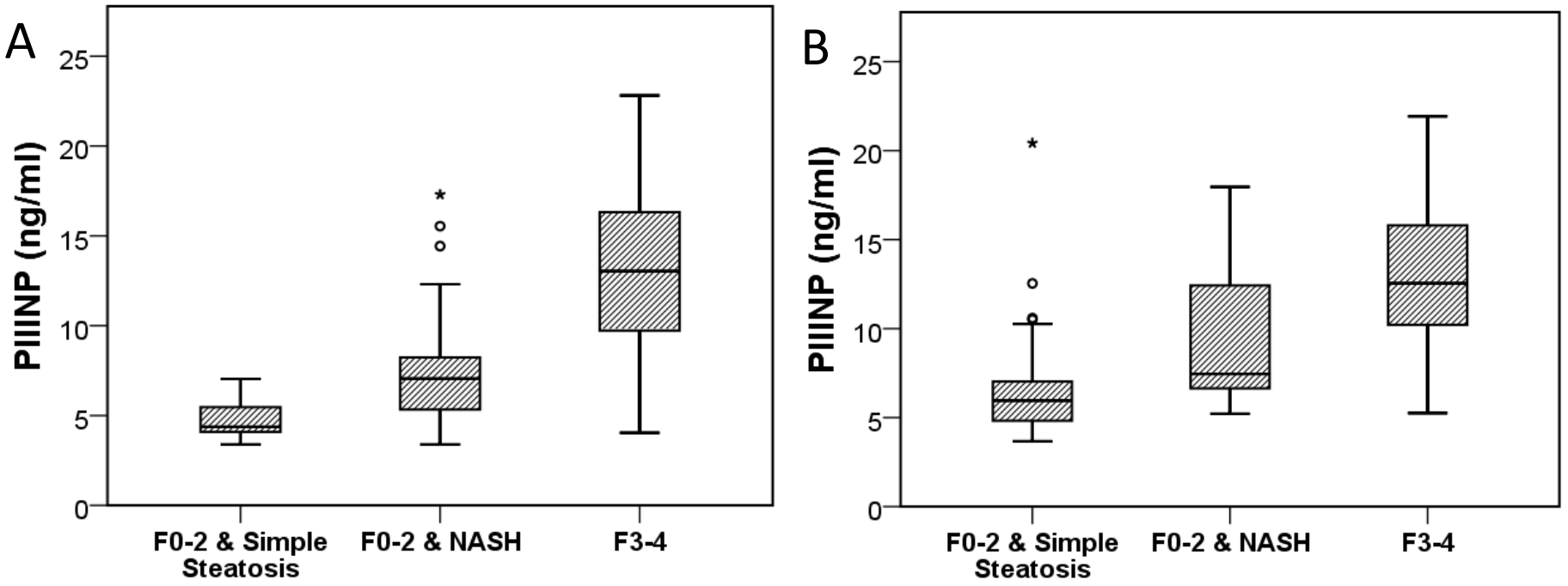


Figure 4.2 Boxplots of PIIINP concentration (ng/ml) in both the derivation and validation cohorts with respect to patients with Simple Steatosis, NASH and advanced fibrosis.

Panel A (Derivation Cohort) & B (Validation Cohort):
Patients stratified into F0-2 with simple steatosis, F0-2 with NASH, and advanced fibrosis (F3-4).

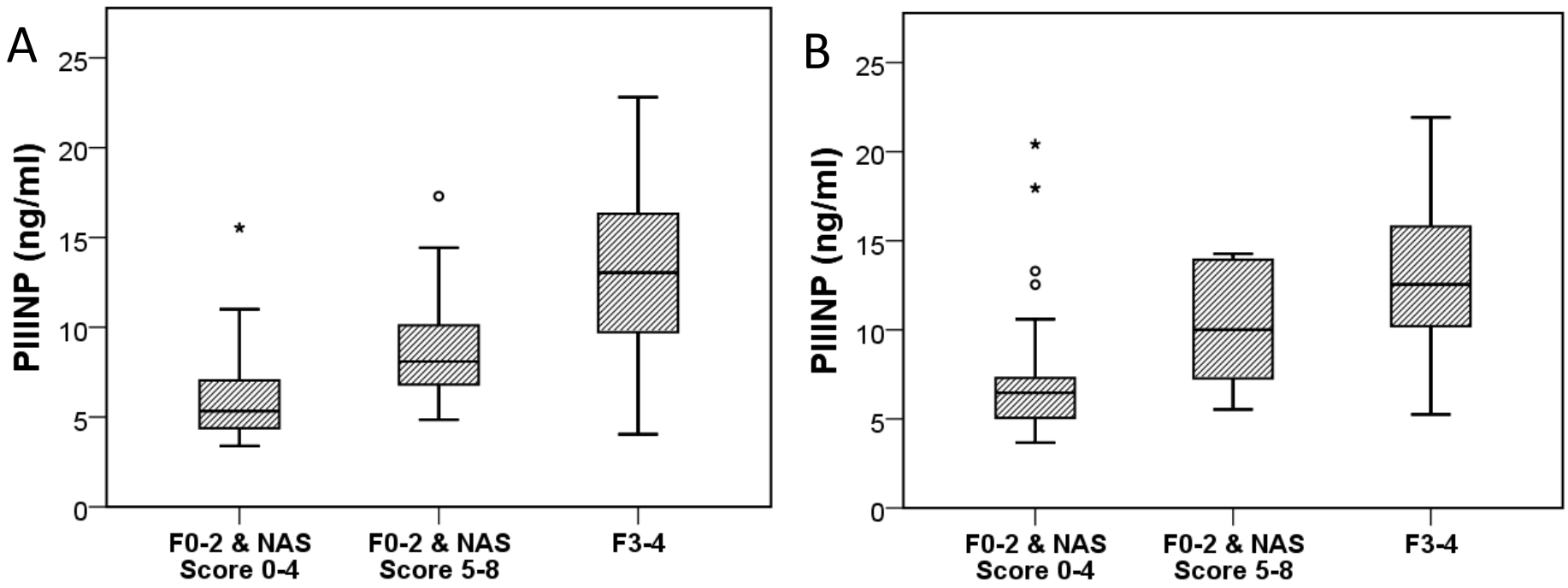


Figure 4.3 Boxplots of PIIINP concentration (ng/ml) in both the derivation and validation cohorts with respect to patients with Simple Steatosis, NASH and advanced fibrosis.

Panels A (Derivation Cohort) & B (Validation Cohort):
Patients stratified into F0-2 with NAS 0-4, F0-2 with NAS 5-8, and advanced fibrosis (F3-4).

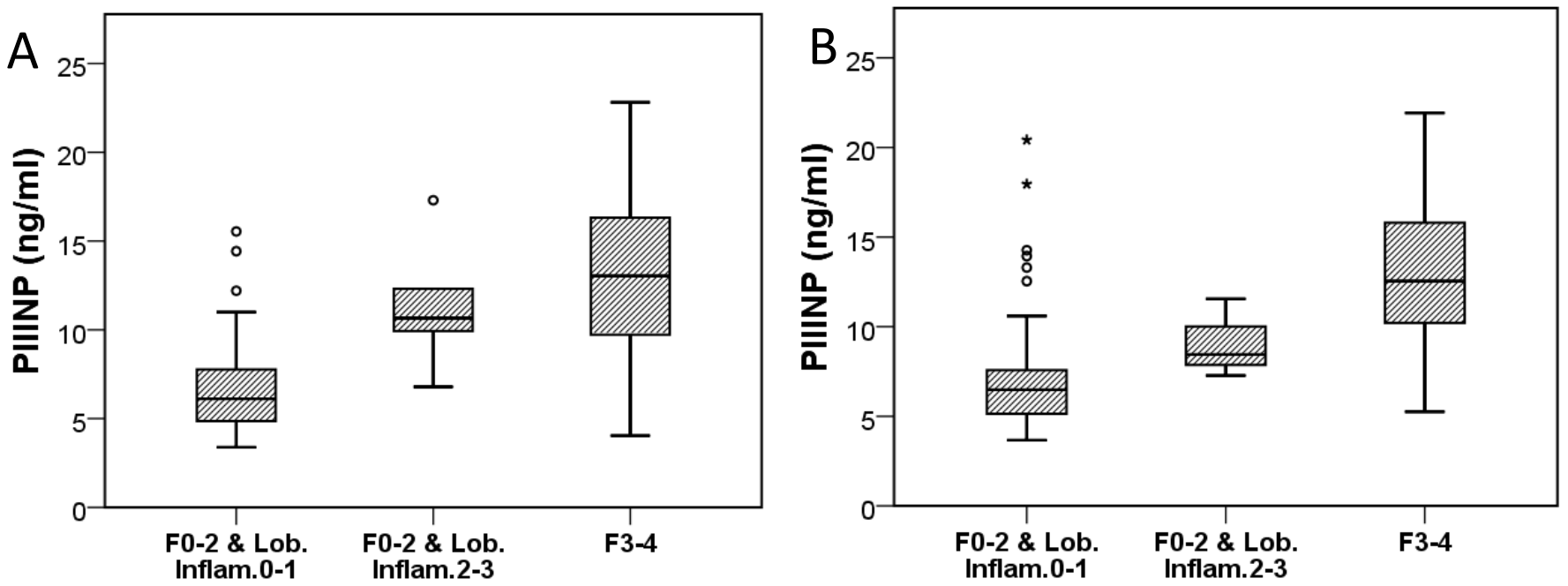


Figure 4.4 Boxplots of PIIINP concentration (ng/ml) in both the derivation and validation cohorts with respect to patients with Simple Steatosis, NASH and advanced fibrosis.

Panels A (Derivation Cohort) & B (Validation Cohort):
Patients stratified into F0-2 without moderate or severe lobular inflammation, F0-2 with moderate or severe lobular inflammation and advanced fibrosis (F3-4)

Clinical utility of PIIINP in the diagnosis of NASH and the exclusion of advanced fibrosis

The performance of PIIINP in the detection of NASH alone, advanced fibrosis alone, or NASH and advanced fibrosis is presented in table 4.7. Despite the two cohorts having a very different prevalence of steatohepatitis the performance of PIIINP in discriminating SS from either NASH or advanced fibrosis was uniformly good (AUROC 0.85-0.87).

Table 4.7 Overlapping utility- the ability of PIIINP to discriminate between either simple steatosis and NASH or Advanced Fibrosis (F3-4) in both cohorts

NASH Grade	Cohort	Numbers in each group	AUROC	95% CI	*P value*	Standard Error
Non-NASH and F0-2 v NASH or F3-4	Derivation	n=11 v n=73	0.86	0.77-0.95	*<0.001*	0.047
	Validation	n=51 v n=37	0.85	0.77-0.93	*<0.001*	0.041
NAS 0-4 and F0-2 v NAS 5-8 or F3-4	Derivation	n=14 v n=43	0.86	0.78-0.94	*<0.001*	0.041
	Validation	n=65 v n=23	0.87	0.78-0.96	*<0.001*	0.045
Lob. Inf.0-1 and F0-2 v Lob. Inf. 2-3 or F3-4	Derivation	n=59 v n=25	0.86	0.77-0.96	*<0.001*	0.049
	Validation	n=68 v n=20	0.87	0.78-0.95	*<0.001*	0.044

As previously stated, PIIINP values associated with NASH in patients without advanced fibrosis were lower than those associated with advanced fibrosis. The clinical utility of this observation is that a clinically relevant diagnostic threshold selected for the detection of NASH (as derived from patients without advanced fibrosis) will have a higher sensitivity (and therefore higher negative predictive value) for a detection of advanced fibrosis.

In this way, PIIINP levels above a threshold set for the diagnosis of NASH could be used as a screening tool in a primary care setting. Whilst a positive result could be used to ‘rule in’ NASH, a negative result could be used to ‘rule out’ NASH and/or advanced fibrosis. Diagnostic thresholds for PIIINP and its performance in the diagnosis of histological steatohepatitis and differing degrees of NASH (NAS score, ballooning and lobular inflammation) are shown in table 4.8. The thresholds corresponding to 80% sensitivity and specificity have been displayed together with a third threshold which maximises the positive predictive value for either histological NASH or the relevant grade of NASH. Also displayed are the negative predictive values for advanced fibrosis and cirrhosis when these thresholds were applied to the entire cohort (F0-4).

Table 4.8 Performance of PIIINP in discriminating either histological NASH or differing grades of NASH and utility in excluding advanced fibrosis

Thresholds of PIIINP concentration (ng/ml) applied to patients with F0-2 fibrosis.

Histologic NASH or grade of NASH	Prev.	Centre	PIIINP Threshold (ng/ml)	Sens.	Spec.	PPV	NPV	LR (+ve)	LR (-ve)	DOR	NPV* in entire cohort for fibrosis	
											<F3-4	<F4
Histologic NASH	83%	Derivation	5.2	80	73	94	43	3	0.27	11.0	90	100
			6.0	65	80	94	32	3.3	0.44	7.4	94	100
			11.0	13	100	100	19	N/A	0.87	N/A	89	99
	28%	Validation	6.4	80	56	41	88	1.8	0.36	5.1	96	100
			7.2	60	80	53	84	3	0.50	6.00	98	100
			11.0	30	97	80	78	10	0.72	13.9	93	99
NAS ≥5	37%	Derivation	6.6	80	68	60	85	2.5	0.29	8.6	95	100
			7.2	77	80	69	86	3.9	0.29	13.3	95	100
			11.0	21	97	80	68	7	0.81	8.6	89	99
	9%	Validation	6.6	80	57	16	97	1.9	0.35	5.3	97	100
			7.2	67	80	25	96	3.4	0.41	8.2	98	100
			11.0	50	94	45	95	8.3	0.53	15.7	93	99
Severe Ballooning	26%	Derivation	6.7	80	64	44	90	2.2	0.31	7.2	95	100
			7.7	59	80	50	85	3.0	0.51	5.8	92	98
			11.0	29	97	77	80	9.7	0.73	13.2	89	99
	3%	Validation	7.2	100	71	16	100	6.25	0	N/A	98	100
			7.9	50	80	7	98	2.5	0.63	4.0	98	100
			11.0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	93	99
Lob Inflamm. ≥1	23%	Derivation	5.2	80	60	37	91	2.0	0.33	6.1	90	100
			7.0	52	80	44	85	2.6	0.60	4.3	95	100
			11.0	12	99	100	80	N/A	0.88	N/A	89	99
	35%	Validation	6.6	80	74	62	87	3.1	0.27	11.4	97	100
			7.0	76	80	67	86	3.8	0.30	12.7	98	100
			11.0	32	100	100	73	N/A	0.68	N/A	93	99
Lob. Inflamm. ≥2	9%	Derivation	7.2	83	67	20	98	2.5	0.25	10.1	95	100
			8.0	83	80	29	98	4.2	0.21	19.8	92	98
			11.0	50	95	50	95	10	0.53	15.7	89	99
	6%	Validation	7.2	80	73	16	98	3.0	0.27	11.0	98	100
			7.6	75	80	19	98	3.8	0.31	12.1	98	100
			11.0	50	91	26	97	5.6	0.55	10.1	93	99

* Prevalence of F3-F4 & F4 fibrosis in the derivation and validation cohorts is 23% & 8% and 19% & 8% respectively.

Clinically relevant PIIINP thresholds

The application of PIIINP at a threshold of 11.0 ng/ml resulted in a positive predictive value of between 80-100% for a histological diagnosis of NASH, and 100% for the presence of lobular hepatitis in the F0-2 group of both cohorts. When applied to patients with all degrees of fibrosis in both cohorts, a threshold of 11.0ng/ml resulted in a positive predictive value of between of between 74-100% for a diagnosis of either steatohepatitis or advanced fibrosis (table 4.9). At this threshold the negative predictive value for a diagnosis of advanced fibrosis alone ranged between 89-93%.

The use of the PIIINP at a 6.6 ng/ml threshold resulted in a negative predictive value for a severe grade of steatohepatitis (NAS 5-8, severe ballooning, severe lobular inflammation) in the F0-2 portion of both cohorts of between 85-100%. When applied to patients with all degrees of fibrosis in both cohorts a PIIINP level of 6.6 ng/ml had a negative predictive value for either NASH or advanced fibrosis of between 80-95% and a negative predictive value for advanced fibrosis of between 95-97% and 100% for cirrhosis.

Table 4.9 Overlapping clinical utility

The performance of PIIINP in discriminating between either simple steatosis and NASH or Advanced fibrosis (F3-4) in both cohorts using the thresholds described for a diagnosis of differing grades of NASH in the F0-2 portion of both cohorts.

Thresholds derived from discrimination of:	Centre	Prev. NASH/ F3-4	PIIINP Threshold (ng/ml)	Performance of PIIINP in discriminating between either simple steatosis or NASH/F3-4						
				Sens.	Spec.	PPV	NPV	LR (+ve)	LR (-ve)	DOR
NASH	Derivation	87%	5.2[a]	81	72	95	36	2.9	0.26	11.0
			6.0[b]	70	82	96	29	3.9	0.37	10.6
			11.0[c]	75	100	100	37	N/A	0.25	N/A
	Validation	42%	6.4[a]	89	57	60	88	2.1	0.19	10.7
			7.2[b]	76	78	71	82	3.5	0.31	11.2
			11.0[c]	46	94	85	71	7.7	0.57	13.3
NAS ≥5	Derivation	52%	6.6[a]	84	68	74	80	2.6	0.24	11.2
			7.2[b]	74	78	79	73	3.4	0.33	10.1
			11.0[c]	40	97	94	60	13.3	0.62	21.6
	Validation	26%	6.6[a]	91	55	42	95	2.0	0.16	12.4
			7.2[b]	91	72	53	96	3.3	0.13	26.0
			11.0[c]	65	92	74	88	8.1	0.38	21.4
Lob Inflamm. ≥2	Derivation	30%	7.2[a]	84	64	50	90	2.3	0.25	9.3
			8.0[b]	80	80	63	90	4.0	0.25	16.0
			11.0[c]	60	95	84	85	12.0	0.42	28.5
	Validation	23%	7.2[a]	95	70	48	98	3.2	0.07	44.3
			7.6[b]	90	72	49	96	3.2	0.14	23.1
			11.0[c]	65	89	64	90	5.9	0.39	15.0

Key to thresholds: a=80%sensitivity, b=80% specificity, c=threshold which maximises PPV

PIIINP as a marker of fibrosis or NASH?

PIIINP has previously been studied as a marker of fibrosis and is one of the 3 component proteins of the ELF test. Given that NASH is associated with the development of fibrosis we assessed whether the performance of PIIINP as a marker of NASH was simply due to its ability to detect patients with NASH who were accruing fibrosis. The ability of PIIINP to discriminate differing degrees of fibrosis in patients without advanced fibrosis (F0-2) in the derivation cohort was poor (AUROC 0.63 F0 v F1-2, AUROC 0.65 F0-1 v F2) (figure 4.5). It is evident that in patients with Kleiner fibrosis stages 0-2, PIIINP discriminated poorly between fibrosis stages, and median PIIINP levels were similar between groups (figure 4.5A). However, PIIINP levels were significantly different between patients stratified by a histological diagnosis, NAS score (0-4 and 5-8) and lobular inflammation at each fibrosis stage (figure 4.5B). Furthermore, when analysing patients with no and minimal fibrosis only (F0-1) in the derivation cohort (table 4.10), the ability of PIIINP to discriminate between NASH and simple steatosis was consistent (AUROC 0.83) and the ability of PIIINP to discriminate differing degrees of NASH was even better (NAS 0-4 v 5-8 AUROC 0.90, Ballooning 0-1 v 2 AUROC 0.92, Lobular Inflammation 0-1 v 2-3 AUROC 0.96).

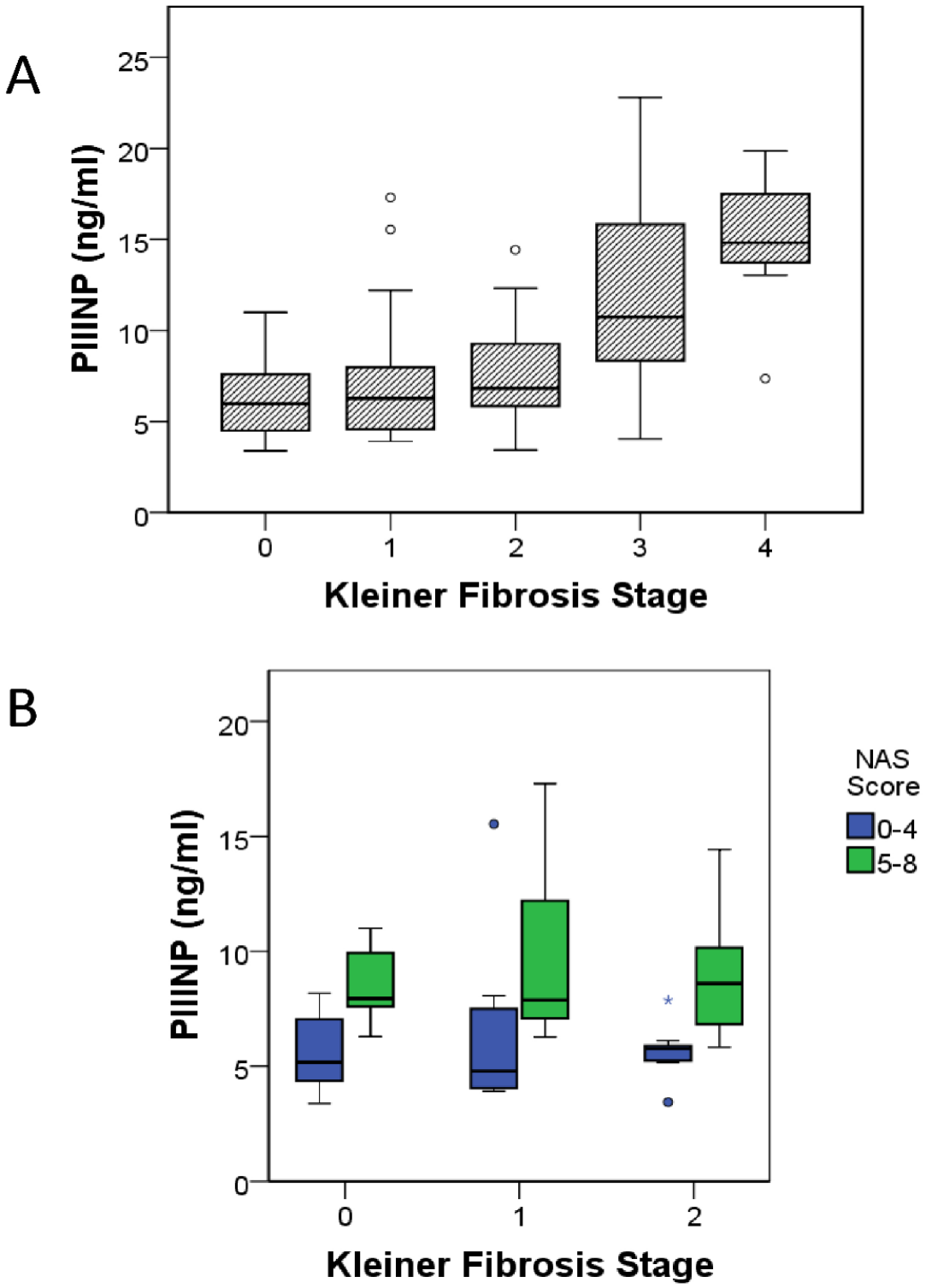


Figure 4.5 PIIINP levels with respect to fibrosis stage

Panel A: Boxplots of serum PIIINP concentration (ng/ml) with respect to fibrosis stage in the derivation cohort.
Panel B: Boxplots of serum PIIINP concentration (ng/ml) in patients without advanced fibrosis (F0-F2) with respect to fibrosis stage when stratified by NAS Score (NAS 0-4 and NAS 5-8).

Table 4.10 Performance of PIIINP in discriminating between differing grades of NASH in patients with either no or minimal fibrosis (F0-1) in the derivation cohort

Degree of NASH	AUROC	95% CI	P-value	Std Error
Simple Steatosis/ Borderline NASH (n=32) v Definite NASH (n=10)	0.83	0.70-0.96	0.002	0.066
NAS 0-4 (n=32) v NAS 5-8 (n=10)	0.78	0.63-0.93	0.009	0.075
NAS 0-5 (n=38) v NAS 6-8 (n=4)	0.90	0.00-1.00	0.009	0.071
B0-1 (n=36) v B2 (n=6)	0.92	0.00-1.00	0.001	0.050
L0 (n=30) v L1-3 (n=12)	0.74	0.58-0.91	0.015	0.085
L0-1 (n=40) v L2-3 (n=2)	0.96	0.00-1.00	0.029	0.037

DISCUSSION

In this chapter, serum levels of PIIINP, a biological component involved in fibrogenesis, were found to discriminate between patients with SS alone and those with NASH amongst a population of patients with NAFLD. Amongst NAFLD patients with fibrosis ranging from none to moderate (F0-2) PIIINP levels differed significantly between patients with or without a histological diagnosis of NASH (AUROC 0.78-0.83), a NAS score of 5-8 (0.80-0.88), and with or without severe lobular hepatitis (AUROC 0.86-0.89), thus permitting the detection of patients with pathological lesions from those with less concerning manifestations of NAFLD. Considering patients with all degrees of fibrosis, PIIINP was able to discriminate between patients with SS and those with either NASH or advanced fibrosis (AUROC 0.85-0.87). This is unsurprising as progressive fibrosis in NAFLD is thought to be the consequence of longstanding NASH. Elevated serum PIIINP has been previously been shown to be associated with elevated pro-inflammatory cytokines in liver disease and this pro-inflammatory state is likely to be the driving force for progressive liver fibrosis. [214] This hypothesis is also supported by genomic and proteomic studies in NASH that have identified targets relating extracellular matrix remodelling.[258] Furthermore, serum PIIINP has been extensively studied in inflammatory joint disease where elevated PIIINP has been associated with active inflammation representing increased tissue turnover.[332] One hypothesises that a similar process occurs in active steatohepatitis.

Hitherto, almost all published studies of markers for NASH have evaluated combinations of laboratory parameters and clinical scores. The only other single marker to be validated in the detection of NASH is CK-18, a marker of apoptosis. A validation study described good performance of caspase-cleaved CK-18 in discriminating NASH from SS with AUROC of 0.83 in a cohort of patients with

advanced fibrosis. [302] Biopsies in this study were scored as ‘not NASH’, ‘borderline NASH’ and NASH. In the present study CK-18 was not identified as a discriminatory marker of NASH.

In addition to using a categorical diagnosis of NASH, the NAS score and its constituent components was used to grade the degree of steatohepatitis. Despite its criticisms, the use of NAS score the in the grading of NASH in contrast to a binary diagnosis of NASH (NASH or ‘not NASH’) permits the evaluation of the relative contributions of the components of the NAS score to a specific variable or panel of tests. However, whilst many investigators have taken NAS≥5 to be synonymous with a diagnosis of NASH, a recent study suggested that NAS at this score is associated with a histological diagnosis of NASH in approximately 70% of cases. [77] Higher NAS scores, whilst more specific for a histological diagnosis of NASH are less sensitive necessitating the use of a NAS score of 5 in clinical practice. In addition, the NAS score has been criticised because it places equal weighting on lobular hepatitis and less weight on ballooning than steatosis. In this study, it was noted that patients in the more obese validation cohort had more steatosis and less lobular inflammation than patients with the same NAS score in the derivation cohort. Regarding caspase-cleaved CK-18, within the derivation cohort the AUROC for CK-18 for its ability to discriminate between patients with a histological diagnosis of steatohepatitis was 0.56, and in patients without severe lobular inflammation or ballooning were both 0.71. Another approach to the use of CK-18 as a biomarker of NASH has been to measure the total level of CK-18. In another study, total CK-18 levels performed better than caspase-cleaved CK-18 with AUROC for discriminating NASH from SS of 0.81.[324] In this study, the AUROC for caspase-cleaved CK-18 in its ability to discriminate NASH from SS was 0.71 which is similar to that observed in the present study. In contrast, however, both total CK-18 and CK-18

fragments performed similarly in their ability to discriminate NASH from SS in another study of 101 patients with morbid obesity with AUROC 0f 0.81 and 0.83 respectively.[333] Despite both cohorts containing patients that had been recruited from specialist liver centres, a differing spectrum of NAFLD related liver disease was identified in each cohort. The derivation cohort was composed of a patient population more analogous to what one might expect to see in secondary care with patients with minimal fibrosis and simple steatosis being underrepresented. In contrast, the validation cohort was composed of a patient population similar to that encountered in primary care with the majority of patients having no or mild fibrosis. As a result, the two populations studied have provided a useful platform on which to derive and validate biomarker test performance. In addition, to simulate biomarker performance in a primary care setting where the advantage of non-invasive assessment of NAFLD is likely to be of greatest benefit, the initial analysis was restricted to patients who did not have advanced fibrosis. Critical evaluations of test performance in primary care are required to determine the utility of PIIINP in that setting.

Whilst one acknowledges that these results need to be reproduced in larger independent series, the relative consistency of the thresholds of PIIINP for detecting differing grades of steatohepatitis in each cohort is striking in view of their differing disease severity. PIIINP at thresholds of 6.6, 7.2 and 11.0 ng/ml were found in the F0-2 portion of both cohorts to have 80% sensitivity, 80% specificity and the highest PPV (45-80%) respectively for a NAS score of 5-8. When the same thresholds were applied to the entire cohort, their performance in the detection of NASH or fibrosis remained high. In theory, these thresholds could be applied in both primary and secondary care settings. The lower threshold of 6.6 ng/ml with 80% sensitivity could be used to ‘rule out’ patients with NASH in primary care. Conversely the higher thresholds of 7.6 or 11

ng/ml could be used to ‘rule in’ NASH or advanced fibrosis respectively in secondary care where more advanced disease is more prevalent. Nevertheless, one is mindful that the performance of these proposed thresholds may vary significantly in other patient populations due to the relatively small sample sizes of the two cohorts.

What is also intriguing is the relatively linear relationship between PIIINP and the grade of NASH. As a result, there is potential for PIIINP to be used as a marker of NASH both in treatment studies and longitudinal studies. Also, of interest will be whether PIIINP can also be used in the non-invasive assessment of inflammation in diseases such as alcoholic hepatitis and viral hepatitis.

In summary, this study has reported two important observations in two cohorts with differing severities of NAFLD related liver disease that are relevant to both primary and secondary care settings. Firstly, within a cohort of patients with NAFLD without advanced fibrosis, PIIINP levels permitted the discrimination of most patients with simple steatosis from those with steatohepatitis. Secondly, within a cohort of patients with NAFLD comprised of all stages of fibrosis, PIIINP levels permitted the discrimination of most patients with NASH or advanced fibrosis from those with simple steatosis. Further studies are required in both primary and secondary care to confirm these findings including studies to assess the relationship between changes in steatohepatitis and PIIINP in response to interventions.

CHAPTER 5

FURTHER VALIDATION OF PIIINP FOR THE DETECTION AND ASSESSMENT OF NON-ALCOHOLIC STEATOHEPATITIS IN PATIENTS WITH NON-ALCOHOLIC FATTY LIVER DISEASE

ABSTRACT

Aims: To further validate PIIINP as a biomarker of NASH in a cohort of patients with biopsy proven NAFLD; to evaluate its performance at the proposed diagnostic thresholds:
5.2 ng/ml (80% Sensitivity), 7.2 ng/ml (80% Specificity) and 11.0 ng/ml (>97% Specificity)

Methods: Seventy-one patients with NAFLD and no evidence of other liver disease were included in this study. Liver biopsies were performed on all patients and analysed by an expert liver histopathologist. All liver biopsies were of suitable size for analysis (>12mm and >5 portal tracts). Liver fibrosis was assessed using the five stage Scheuer classification modified to stage NAFLD. In addition, liver biopsies were classified in a dichotomous manner into those with and without histological NASH and those with or without advanced fibrosis. Serum samples were taken at the time of liver biopsy.

Results: Sixty patients (85%) did not have advanced fibrosis (4 with S0, 18 with S1, and one with S2). Fourteen patients without advanced fibrosis (23%) had NASH. All 11 patients with advanced fibrosis (9 with S3, 2 with S4) had NASH. PIIINP performed well in its ability to discriminate between those with and without histological NASH both in patients with non-advanced fibrosis (S0-2, n=46 vs n=14, AUROC 0.81) and patients with all degrees of fibrosis (S0-4, n=46 vs n=25, AUROC 0.87). By comparison, the AUROC of ALT in discriminating between those with and without histological NASH ranged between 0.43-0.45. The proposed thresholds (5.2, 7.2 and 11ng/ml) of PIIINP performed well in their ability to diagnose NASH in this study population.

Conclusions: These findings further validate the ability of PIIINP to discriminate between patients with and without NASH. Moreover, comparable diagnostic performance has been identified at the proposed thresholds of 5.2, 7.2 and 11ng/ml in this secondary care cohort. Further studies of PIIINP in this context are warranted in larger independent cohorts including those in primary care.

INTRODUCTION

Due to the ubiquity of NAFLD, liver biopsy is not practicable for use as the primary modality for disease detection within the population at risk. By contrast, serum biomarkers are a far more attractive modality by which liver disease can be detected in a primary care setting. However, prior to the implementation of a diagnostic strategy employing NITs in clinical practice, it is essential that NITs are comprehensively validated prior to their use including validation in an appropriate clinical setting.

Hitherto, there have been numerous examples of NITs that have exhibited more modest performance during validation thus allowing both researchers and clinicians view the potential diagnostic aspirations of the NIT in question with more realism. In the context of NAFLD, the use of CK-18 as a NIT for the detection of NASH is such an example. The excitement generated by the original study of CK-18 as a NIT for the detection of NASH within NAFLD [328] has been dampened by numerous subsequent validation studies during which the diagnostic performance of CK-18 in this context have been far more modest.[334] These results suggest that the diagnostic performance of CK-18 as a NIT of NASH (total levels and cleaved) were initially overestimated. Factors that may herald an overestimation of diagnostic performance include factors relating to the patient population, factors relating to the NIT such as sample processing and the reference standard itself.

In this chapter PIIINP is further validated as a biomarker of NASH in a cohort of patients with biopsy proven NAFLD recruited in a multicentre international study following referral from primary care with abnormal liver function tests.

METHODS

Patients

The patients included in this validation study were recruited in an international multicentre cohort study of 1021 patients in which levels of serum markers were compared with liver histologic fibrosis staging.[221] All 1021 patients were recruited following referral to secondary care hepatology clinics with suspected chronic liver disease (abnormal LFTs >6 months). Of the 1021 patients, clinical details or biochemistry was incomplete for 45 patients and for 55 patients the liver biopsy specimen was of an inadequate size (<12mm or <5 portal tracts). Of the 921 patients included in the study, 92 had a suspected clinical diagnosis of NAFLD. However, after a complete clinicopathologic evaluation 21 patients were identified to have other contributing aetiologies for CLD thus leaving 71 patients with a solitary clinical diagnosis of NAFLD.

Liver Biopsy

Liver biopsy was performed on all patients with histology scored by a senior histopathologist according to the 5-point Scheuer staging system (S0-S4). All liver biopsy specimens were of adequate size (>12mm or >5 portal tracts). Intra-observer agreement was assessed (κ=0.934 SE 0.12 *p=0.001*). For the patients with NAFLD, the statement criteria for the Scheuer staging system were modified to appropriately stage NAFLD staging (e.g. perivenular and pericellular fibrosis replaced portal and periportal fibrosis). The grade of steatohepatitis was categorised in a dichotomous manner into those with (definite NASH) and without (simple steatosis or borderline NASH) histological NASH. Advanced fibrosis was defined as Scheuer stages S3 and S4.

Serum Samples

Serum samples were obtained on all patients which were analysed for direct serum markers of fibrosis including PIIINP using immunoassays that were developed to run on the Bayer IMMUNO 1 platform.

Analysis

As before, patients with advanced liver fibrosis (S3 and S4) were excluded from the initial analyses to assess the performance of PIIIINP in identifying NASH in patients prior to the development of advanced liver fibrosis. Thereafter, the performance of PIIINP in staging the severity of NAFLD amongst patients with all degrees of fibrosis was assessed. Given that standard liver function tests are commonly employed in primary care to assess the severity of NAFLD, the performance of ALT in discriminating between patients with and without steatohepatitis and/or advanced fibrosis was also assessed.

Statistical Analysis

Diagnostic performance of both PIIINP and ALT in discriminating between different stages of NAFLD was assessed using receiver operating characteristic curves with the AUROC and 95% confidence intervals of AUROC calculated. Sensitivity and specificity, predictive values, likelihood ratios (LR+ and LR-), and diagnostic odds ratios were calculated at thresholds derived from ROC curves.

Thereafter, the diagnostic performance of PIIINP at the 3 proposed diagnostic thresholds identified in chapter 4 were validated (5.2 ng/ml (80% Sensitivity), 7.2 ng/ml (80% Specificity) and 11.0 ng/ml (>97% Specificity)).

RESULTS

Patient demographics

Of the 71 patients with NAFLD, 60 had non-advanced fibrosis (S0-2) and 11 had advanced fibrosis (S3-4). The baseline characteristics of all the patients and the patients without advanced fibrosis are shown in Tables 5.1 and 5.2 respectively. All patients with advanced fibrosis (S3-4) had a histologic diagnosis of NASH in contrast to only 23% amongst patients without advanced fibrosis (S0-2).

Table 5.1 Demographic and baseline data of patients with all degrees of fibrosis (S0-4)

Variable	Simple Steatosis/ Borderline NASH n=46 (65%)	NASH n=25 (35%)	*P value*	Overall n=71
Male (n,%)	39 (85%)	15 (60%)	*0.031*	53 (76%)
S3-4	0	25 (100%)	***<0.001***	25 (35%)
Age (years)	41.1 ± 9.8	50.5 ± 12.5	*NS*	44.4 ± 11.6
ALT (IU/L)	66.3 ± 38.8	69.5 ± 65.5	*0.011*	67.4 ± 48.9
Platelets (x10^9/L)	239 ± 60	241 ± 96	*NS*	239 ± 74
HA (ng/ml)	25 ± 24.6	95 ± 136	***<0.0001***	50 ± 89
TIMP-1 (ng/ml)	611 ± 234	831 ± 435	***0.001***	688 ±334
PIIINP (ng/ml)	3.80 ± 1.48	8.41 ± 4.70	***<0.0001***	5.42 ± 3.71

Data are presented as mean ± standard deviation
Significant Associations (P value<0.05) are displayed in bold

Table 5.2 Demographic and baseline data of patients without advanced fibrosis (S0-2)

Variable	Simple Steatosis/ Borderline NASH n=46 (77%)	NASH n=14 (23%)	*P value*	Overall n=60
Male (n,%)	39 (85%)	11 (79%)	*NS*	50 (83%)
Age (years)	41.1 ± 9.8	49.4 ± 14.5	***0.017***	43.0 ± 11.5
ALT (IU/L)	66.3 ± 38.8	63.2 ± 55.0	*NS*	28.6 ± 27.8
Platelets (x10^9/L)	239 ± 61	273 ± 103	*NS*	247 ± 73.5
HA (ng/ml)	25.1 ± 24.6	40.3 ± 34.9	*NS*	28.6 ± 27.8
TIMP-1 (ng/ml)	611 ± 234	826 ± 405	***0.015***	881 ± 293
PIIINP (ng/ml)	3.80 ± 1.48	5.71 ± 1.52	***<0.001***	4.2 ± 1.7

Data are presented as mean ± standard deviation
Significant Associations (P value<0.05) are displayed in bold

Serum biomarker levels

Serum levels of HA, TIMP-1 and PIIIINP were all significantly raised amongst patients with NASH within patients with all degrees of fibrosis. Amongst patients without advanced fibrosis, however, HA levels did not differ significantly between patients with and without NASH: this was in contrast to both PIIINP (p<0.0001) and TIMP-1 (p=0.001). PIIINP performed well in its ability to discriminate between patients with NAFLD with and without NASH (table 5.3) amongst patients both without advanced fibrosis (AUROC 0.81) and with advanced fibrosis (AUROC 0.87). By contrast, serum ALT levels were unable to discriminate between patients with NAFLD, with and without NASH (table 5.4, figure 5.1), amongst patients both without advanced fibrosis (AUROC 0.43) and with advanced fibrosis (AUROC 0.47).

Table 5.3 Diagnostic Performance of PIIINP in the detection of a severe manifestation of NAFLD

NAFLD Stage	Cohort	Numbers in each group	AUROC	95% CI	*P value*	Standard Error
'Non-NASH' v NASH	S0-2	n=46 v n=14	0.81	0.69-0.94	*<0.001*	0.62
	S0-4	n=46 v n=25	0.87	0.79-0.96	*<0.001*	0.04
'Non-NASH' v either NASH or advanced fibrosis	S0-4	n=46 v n=25	0.87	0.79-0.96	*<0.001*	0.04

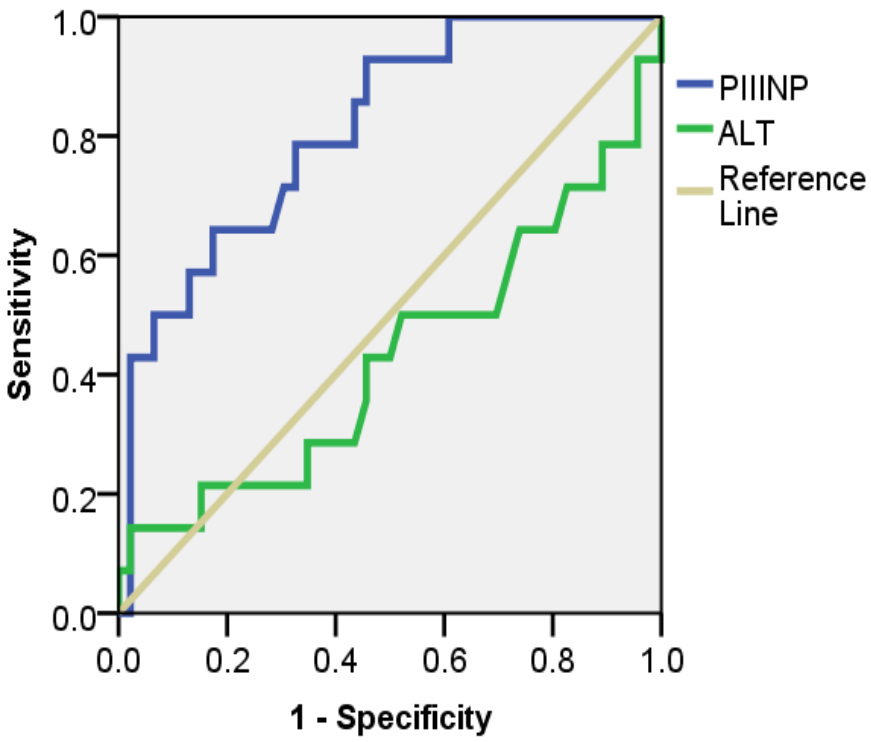
Performance (AUROC) with respect to detecting histological NASH in patients with F0-2, F0-4 and detecting either NASH or advanced fibrosis in patients with S0-4

Table 5.4 Performance of alanine aminotransferase in the detection of a severe manifestation of NAFLD

NASH Stage	Cohort	Numbers in each group	AUROC	95% CI	*P value*	Standard Error
Histologic Diagnosis: 'Non-NASH' v NASH	S0-2	n=46 v n=14	0.43	0.24-0.62	*NS*	0.10
	S0-4	n=46 v n=25	0.47	0.32-0.63	*NS*	0.08
Histologic Diagnosis: 'Non-NASH' v either NASH or advanced fibrosis	S0-4	n=46 v n=25	0.47	0.32-0.63	*NS*	0.08

Performance with respect to detecting histological NASH in patients with S0-2, S0-4 and detecting either NASH or advanced fibrosis in patients with S0-4

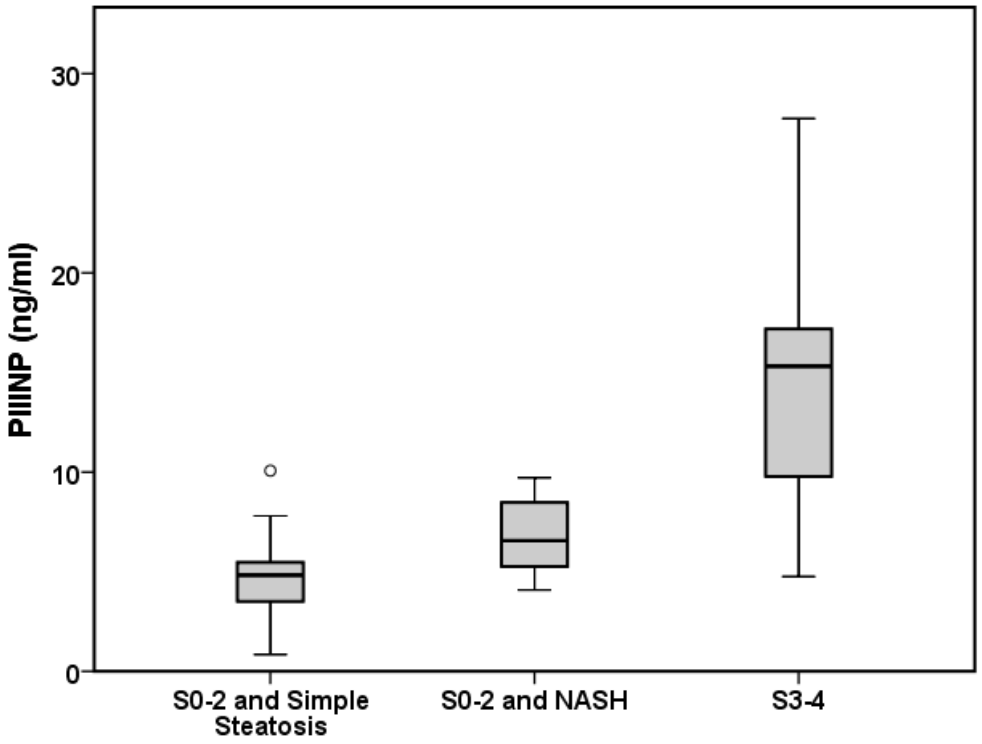
Figure 5.1 ROC curves for both PIIINP and ALT in their ability to discriminate between NASH and Simple Steatosis in patients with F0-2



PIIINP levels represent a hierarchy of liver disease

As identified in chapter 4, PIIINP levels once again were found reflect a hierarchy of liver disease severity ranging from SS with no or mild fibrosis, steatohepatitis with no or mild fibrosis and advanced fibrosis (figure 5.2).

Figure 5.2 Boxplot of PIIINP concentration with respect to patients with Simple Steatosis and NASH amongst patients advanced fibrosis (S0-2) and advanced fibrosis (S3-4)



Performance of PIIINP at the proposed diagnostic thresholds (table 5.5) of 5.2 ng/ml (80% Sensitivity), 7.2 ng/ml (80% Specificity) and 11.0 ng/ml (>97% Specificity)

- **Patients without advanced fibrosis (S0-2)**

  Employing PIIINP at a threshold of 5.2ng/ml resulted in an NPV of 91% (79% sensitivity, 63% specificity). The application of PIIINP at a threshold of 7.2ng/ml resulted in a PPV of 72% (43% sensitivity and 95% specificity). At a threshold of 11.0ng/ml, PIIINP was 100% specific (PPV 100%).

- **Patients with all degrees of fibrosis (S0-4)**

  Employing PIIINP at a threshold of 5.2ng/ml resulted in an NPV of 93% (84% sensitivity, 65% specificity). The application of PIIINP at a threshold of 7.2ng/ml resulted in a PPV of 86% (64% sensitivity and 97% specificity); at 11.0ng/ml the test exhibited 100% specificity and 28% sensitivity (PPV 100% and NPV 72%).

Table 5.5 Performance of PIIINP in discriminating either histological NASH or 'NASH with advanced fibrosis' at diagnostic thresholds(ng/ml) in patients with either S0-2 or S0-4 fibrosis

Histologic Stage	Prev.	Population	PIIINP Threshold (ng/ml)	Sens.	Spec.	PPV	NPV	LR (+ve)	LR (-ve)	DOR
Histological NASH	23%	S0-2	5.2	79	63	39%	91%	2.1	0.33	6.4
			6.0	57	83	50%	87%	3.4	0.52	6.5
			7.2	43	95	72%	85%	8.6	0.60	14.3
			11.0	0	100	N/A	N/A	N/A	1.00	N/A
Histological NASH	35%	S0-4	5.2	84	65	56%	89%	2.4	0.25	9.8
			6.0	76	82	70%	86%	4.2	0.29	14.4
			7.2	64	97	92%	83%	21.3	0.37	57.5
			11.0	28	100	100%	72%	∞	0.72	∞
Histologic NASH or Advanced Fibrosis	35%	S0-4	5.2	84	65	56%	89%	2.4	0.25	9.8
			6.0	76	82	70%	86%	4.2	0.29	14.4
			7.2	64	97	92%	83%	21.3	0.37	57.5
			11.0	28	100	100%	72%	∞	0.72	∞

DISCUSSION

In this study PIIINP has been further validated as a biomarker of NASH and advanced fibrosis in NAFLD. These data confirm earlier findings that PIIINP can discriminate between simple steatosis and the more severe forms of NAFLD namely NASH and advanced fibrosis. Furthermore, the diagnostic thresholds identified in chapter 4 have demonstrated robust performance in either ruling out or ruling in either NASH and/or advanced fibrosis.

Combining the 71 patients included in this validation study with the 88 patients in which PIIINP was validated in chapter 4 creates a validation sample of 159 patients in which PIIINP is able to discriminate well between patients with and without NASH (amongst patients with F0-2 AUROC 0.82 95% CI 0.74-0.90, and F0-4 AUROC 0.88 95% CI 0.82-0.93) Whereas these numbers are relatively modest, a particular strength of these validation studies is the high quality of liver biopsy interpretation which is confirmed by a high level of intra-observer and inter-observer agreement.[221] Moreover, a diagnosis of NAFLD was made only after strict clinicopathologic evaluation during which other co-morbidities such as granulomatous disease and iron overload were excluded. Liver biopsies in this study were classified as those with and without NASH. As previously outlined, a histologic diagnosis of NASH is a much more powerful prognostic indicator than the NAS score itself or its constituent components. As such, studies defining the prognostic ability of PIIINP to determine liver related events and outcomes are also required.

Regardless, there is a much higher prevalence of the severe manifestations of NAFLD in this study cohort than in the wider population. In this study group, 20% of patients were found to have NASH without advanced fibrosis and a further 15% had both NASH and advanced fibrosis. These figures compare to 3-5% and 1-2% in the general

population respectively. This observation was not unexpected as secondary care validation cohorts typically are enriched with cases of more severe disease. Diagnostic studies performed amongst a secondary care population may overestimate the diagnostic performance of a NIT that will be implemented in a primary care setting where the prevalence of advanced disease is less. Regardless, the patients included in this study were referred after presenting to secondary care with abnormal liver function tests persisting for at least 6 months and negative liver serology for other co-morbidities, thus representing a patient population for which there remains uncertainty regarding the criteria for referral in the absence of a defined referral strategy. As such, whilst these results are encouraging further validation studies will now need to be conducted in larger patient groups of clinical relevance such as primary care, diabetic clinics, obesity clinics.

Nevertheless, the overlapping utility of a single test that can detect both NASH and/or advanced fibrosis is certainly appealing in primary care. A negative result after employing a diagnostic threshold of high sensitivity could be used to appropriately rule out a severe manifestation of NAFLD. With regard to a diagnosis of NASH, a positive result could either prompt a referral to secondary care or signal the need to apply an additional NIT. Certainly, within secondary care a combination of PIIINP with other NITs such as CK-18 may be beneficial given that the PPV of NASH will be limited by the low pre-test probability of NASH. With regard to a diagnosis of either advancing or advanced liver fibrosis, a positive PIIINP result could be further characterised by the application of other NITs such as transient elastography or an additional serum marker.

These results also confirm that standard liver function tests such as ALT are unable to quantify the severity of NAFLD. In this study, the performance of ALT in detecting

both NASH and /or advanced fibrosis was worse than chance itself. These results highlight the urgent need to implement a referral system in primary care that can appropriately identify NAFLD patients that have either NASH or advanced fibrosis that does not rely on abnormality of transaminases.

Currently, efficacious antifibrotic therapies are not available for the treatment of liver fibrosis in CLD in general or within NAFLD itself. As such, the current focus is on disease detection. In contrast to the stage of advanced fibrosis, the presence of NASH in a liver without advanced fibrosis is a stage at which it is more likely that disease progression can be arrested by interventions such as dietary modification and exercise. Moreover, as NASH itself is a potential therapeutic target of antifibrotic therapy, candidates for such therapy will need to have ongoing NASH regardless of their fibrosis stage.

In summary, these results further validate the use of PIIINP as a biomarker of NASH and/or advanced fibrosis in NAFLD. Given that a NIT for NASH will primarily be employed in primary care, larger independent studies are required to confirm the performance of PIIINP in this context.

CHAPTER 6

COMPARISON OF SERUM BIOMARKERS IN DISCRIMINATING BETWEEN FIBROSIS STAGES IN NAFLD: DIRECT SERUM MARKERS ARE MORE ACCURATE THAN INDIRECT SERUM MARKERS

ABSTRACT

Background: The identification of fibrosis in patients with NAFLD is important for ascertaining prognosis and stratifying patients for emerging therapeutic interventions. Use of both direct marker panels (liver matrix components) and indirect marker panels (simple biochemical tests) have been described for the detection of fibrosis in NAFLD.

Aims: To compare the performance of direct and indirect serum marker panels in the detection of fibrosis in NAFLD as compared liver biopsy.

Methods: From 2 centres, 177 patients were recruited and underwent percutaneous liver biopsy. Fibrosis staging was assessed using Kleiner criteria by 2 senior liver-histopathologists. Serum obtained within 3 months of liver biopsy was used to calculate 6 indirect marker panels of fibrosis (APRI, BAAT, BARD, Cirrhosis discriminate score, NAFLD fibrosis index and FIB4). These panels were compared with the direct serum marker panels, PIIINP, HA and ELF Test (HA, TIMP1, PIIINP). Diagnostic accuracy was assessed using receiver operating characteristic curves for the detection of mild fibrosis (F1-4), moderate fibrosis (F2-4), advanced fibrosis (F3-4) and cirrhosis. The areas under receiver operating characteristic curves (AUROC) were compared using the method of DeLong. The Obuchowski Measure was used to assess overall discriminatory power.

Results: The distribution of fibrosis stages in the cohort was as follows: F0 39.5% (n=70), F1 19.2% (n=34), F2 17.5% (n=31), F3 13.6% (n=24), F4 10.2% (n=18). None of the biomarkers tested exhibited good performance for the detection of mild fibrosis.

Both ELF (AUROC 0.82) and HA (AUROC 0.81) were the only biomarkers demonstrating good performance for detecting moderate fibrosis. Both direct and indirect markers demonstrated good performance for the identification of advanced (F3-4) fibrosis: ELF (AUROC 0.89), HA (AUROC 0.88), PIIINP (AUROC 0.83), NAFLD Fibrosis score (AUROC 0.83) and FIB4 (AUROC 0.83). For a diagnosis of cirrhosis, ELF (AUROC 0.94), HA (AUROC 0.91), PIIINP (AUROC 0.90) all demonstrated excellent diagnostic performance; good performance was exhibited by APRI (AUROC 0.81), FIB4 (AUROC 0.85) and NAFLD Fibrosis Score (0.86). ELF had the best performance overall (Obuchowski measure 0.91), which was significantly better than HA in its ability to discriminate both minimal fibrosis ($p = 0.02$) and cirrhosis ($p = 0.05$).

Conclusions: Amongst patients with NAFLD, direct serum marker panels of fibrosis have superior discriminatory performance compared to indirect marker panels. Regardless, none of the markers tested demonstrated good performance for detecting mild fibrosis. ELF was the best direct biomarker studied, offering excellent diagnostic performance for cirrhosis and good performance for both moderate and advanced fibrosis. FIB4 and NAFLD fibrosis score were the best indirect biomarkers studied and exhibited good performance for the detection of advanced fibrosis and cirrhosis.

INTRODUCTION

The hepatic manifestation of the metabolic syndrome, NAFLD, is the most common cause of chronic liver disease worldwide.[4, 7, 8, 335] Whereas the clinical spectrum of NAFLD encompasses simple steatosis (SS) and NASH with progressive fibrosis, only 1-2% of patients with NAFLD will have progressive fibrosis.[4] which may progress to cirrhosis and its complications.[4, 320] Due to the inherent limitations of liver biopsy as a tool to stage liver disease in the population at risk of NAFLD, numerous serum biomarkers have been developed to detect progressive liver fibrosis. Serum biomarkers of fibrosis can be categorized in several ways including into direct serum markers (measuring parameters directly related both the fibrogenic and fibrolytic processes involved in liver matrix turnover) and indirect serum markers (combinations of serum parameters which are related to liver function including AST and ALT).[3] The diagnostic targets of these biomarkers have included mild (F1), moderate (F2) and advanced (F3-4) liver fibrosis.[322]

Hitherto many tests have been described for the assessment of fibrosis in NAFLD but there have been very studies have been conducted that directly compare their performance head to head. [336] Due to factors such as spectrum bias[201], any comparison of the performance of biomarkers derived from patient populations with differing characteristics can be subject to type I error.

In chapters 4 and 5 it was demonstrated that PIIINP performs well at detecting NASH or advanced fibrosis but its discriminatory performance for detecting mild and moderate fibrosis is suboptimal.

To complement the introduction of a biomarker of NASH, accurate biomarkers of fibrosis are needed in clinical practice that can accurately discriminate between

different fibrosis stages. In addition, accurate biomarkers are required in clinical and research practice that can monitor fibrosis not only in response to therapeutic interventions (diet and exercise, bariatric and pharmacologic interventions) for fibrosis but also during the natural history of NAFLD.

In this chapter, to address these 2 needs the performance of both direct and indirect biomarkers in the detection of liver fibrosis has been studied within a population of patients recruited from secondary care liver clinics.

METHODS

Patient population

The patients included in the study were recruited consecutively from outpatient clinics in two liver units in the United Kingdom (Queen's Medical Centre, Nottingham and the Freeman Hospital, Newcastle-Upon-Tyne) and were diagnosed with NAFLD using the criteria outlined in chapter 4.

Histological assessment

Liver biopsies were assessed by a single senior liver histopathologist at each centre (PK in Nottingham PK, AB in Newcastle) and scored for histological grade and fibrosis stage using the classification system described by Kleiner et al.[190] In this system, fibrosis was scored using a 5 point scoring system (1= mild /moderate zone 3 perisinusoidal fibrosis, or portal fibrosis only; 2 = zone 3 and portal/periportal fibrosis; 3 = bridging fibrosis; 4 = cirrhosis).

Biomarkers used in this study

The serum samples used in the study were taken within 3 months of biopsy and were analysed for routine biochemistry (including ALT, AST, gamma glutamyl transferase [GGT], bilirubin, albumin and alkaline phosphatase), full blood count, measurements of insulin resistance (fasting glucose and insulin), ferritin and cholesterol. Serum samples were stored at -70ºC prior to transfer to the central laboratory, where serum samples were analyzed for levels of HA, TIMP-1 and PIIINP using the proprietary assays developed for the ELF test by Siemens Healthcare Diagnostics Inc (Tarrytown, New York, USA). The assays are magnetic particle separation immunoassays and were performed on the ADVIA Centaur® immunoassay system (Siemens Medical Solutions Diagnostics Inc, Tarrytown, New York, USA).

Markers Evaluated in this study

The diagnostic performance of nine biomarkers of liver fibrosis were compared in this study: (ELF[221], Hyaluronic acid[337], PIIINP[206], BAAT[338], BARD[255], Cirrhosis Discriminant Score[339], NAFLD Fibrosis score[252], APRI[340], FIB4[341]).

The discriminant scores of these biomarkers were calculated using published formulae (full description in appendix).

Statistical Analysis

Statistical analyses were performed using SPSS for Windows (version 20, SPSS Inc, Chicago, IL) and R for Windows (version 2.15.1, The R Foundation for Statistical Computing). Patient demographic and clinical laboratory characteristics were descriptively summarized and reported as mean ± standard deviation (SD) and range. All tests were two-sided and statistical significance was assessed at the 0.05 threshold. The diagnostic performance of the biomarkers was compared to liver biopsy and was assessed using ROC curves. The AUROC and 95% confidence intervals of AUROC were calculated. Good performance for a test within study cohort was defined as an AUROC > 0.8 and excellent diagnostic performance defined as AUROC > 0.9.[199]

AUROC were compared using the method of Delong.[331] Sensitivity, specificity and predictive values were calculated at thresholds derived from ROC curves.

The Obuchowski[202] method of correcting for spectrum effect was applied in a similar fashion to previously published literature. The Obuchowski measure (ordROC) gives an average of the N(N-1)/2 AUROC pairwise comparisons between N categories of gold standard outcome. Thus, using the Metavir scale with its N (=5) categories of fibrosis staging (F0-4) there are 10 pair wise comparisons between any 2 of the N categories. As previously described, a penalty function proportional to the difference

in Metavir units between fibrosis stages was employed (table 6.1).[342] The penalty function that punishes the degree of misclassification of fibrosis stages, was 0.25, 0.5, 0.75, and 1 when the difference between Metavir stages was 1, 2, 3 and 4 stages respectively. It is important to recognize that the numerical value generated through calculation of the Obuchowski measure is not directly comparable to the AUROC. In addition, values of the Obuchowski score have not been defined in the literature to describe levels of performance (such as good or excellent).

Table 6.1 Matrix used to apply the Obuchowski Penalty Correction

0	**0.25**	**0.5**	**0.75**	**1**
0.25	**0**	**0.25**	**0.5**	**0.75**
0.5	**0.25**	**0**	**0.25**	**0.5**
0.75	**0.5**	**0.25**	**0**	**0.25**
1	**0.75**	**0.5**	**0.25**	**0**

RESULTS

Patient characteristics and baseline histology

The baseline characteristics of the study cohort are described in table 6.2. The mean age of the subjects was 48.8 years. Sixty five percent of patients were male and 35% were diabetic. Seventy seven percent of patients had a histologic diagnosis of NASH. All stages of fibrosis were represented with the spectrum of fibrosis typical of a secondary care cohort. The prevalence of mild (F1-4), moderate fibrosis (F2-4), advanced fibrosis (F3-4) and cirrhosis (F4) was 61%, 42%, 24% and 10% respectively. After stratifying the baseline characteristics of the subjects of the study according to their histologic fibrosis stage, it was evident that the presence of advanced fibrosis was significantly associated with older subjects ($p<0.001$), diabetes ($p-=0.011$), the presence of NASH ($p<0.001$), platelet count ($p<0.001$) and serum albumin ($p<0.001$).

Boxplots of the 9 biomarkers with respect to fibrosis stage in the study population are presented as figures 6.1-6.3.

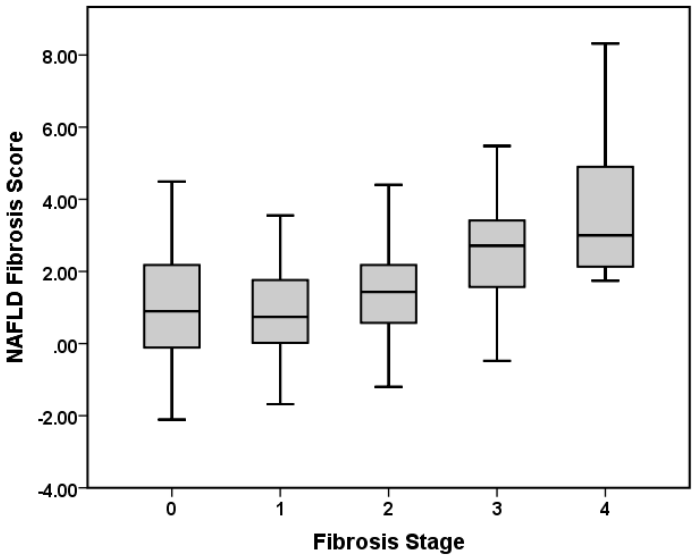
Table 6.2 Demographics of the patient cohort

	Non- advanced Fibrosis (F0-2)	Advanced Fibrosis (F3-4)	*P-Value*	Entire Cohort (F0-4)
Number	**135 (76%)**	**42 (24%)**	**-**	**177**
Age (years)	**46.1 ± 12.3**	**57.3 ± 9.5**	***<0.001***	**48.8 ± 12.6**
Male subjects	**n=95 (83%)**	**n=20 (48%)**	***0.007***	**115 (65%)**
Diabetes (yes)	**n=73 (54%)**	**n=33 (79%)**	***0.011***	**n=61 (35%)**
NASH	**n=96 (54%)**	**n=40 (98%)**	***<0.001***	**n=136 (77%)**
Waist (cm)	**107 ± 12.1**	**109 ± 16.2**	***0.319***	**108 ± 13.2**
BMI (kg/m^2)	**32.0 ± 5.8**	**32.8 ± 5.4**	***0.416***	**32.2 ± 5.7**
Fasting Glucose (mmol/L)	**5.8 ± 1.7**	**7.8 ± 4.3**	***<0.001***	**6.3 ± 2.7**
Triglycerides (nmol/L)	**2.3 ± 1.6**	**2.7 ± 2.3**	***0.202***	**2.4 ± 1.8**
Platelets ($x10^9/L$)	**254 ± 63**	**192 ± 76**	***<0.001***	**239 ± 70**
ALT (IU/L)	**80 ± 61**	**69 ± 39**	***0.271***	**77.5 ± 57.2**
AST (IU/L)	**45 ± 19**	**51 ± 18**	***0.091***	**46.7 ± 18.9**
GGT (IU/L)	**104 ± 88**	**172 ± 188**	***0.001***	**120 ± 123**
Albumin (g/L)	**44 ± 3.4**	**42 ± 4**	***<0.001***	**44.3 ± 3.8**
Fibrosis Stage				
F0				**n=70, 39%**
F1				**n=34, 19%**
F2				**n=31, 18%**
F3				**n=24, 14%**
F4				**n=18, 10%**

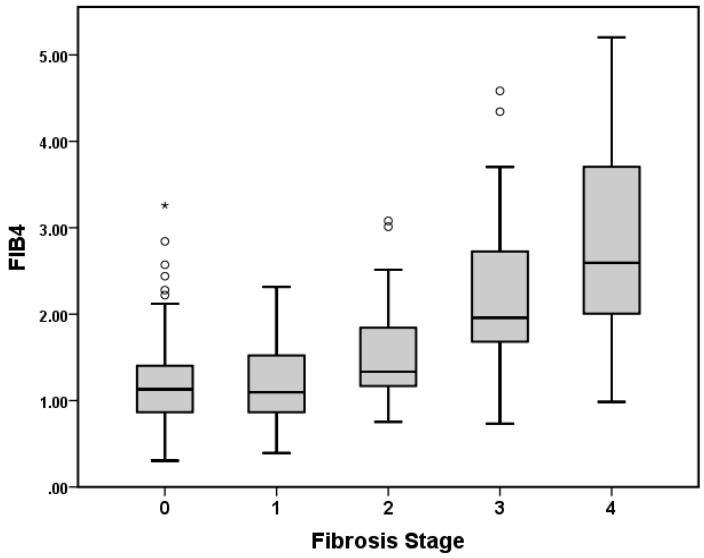
Figure 6.1 Boxplots of indirect serum markers with respect to fibrosis stage in study population

(Panel A: NAFLD Fibrosis Score, Panel B: FIB4, Panel C APRI.

A



B



C

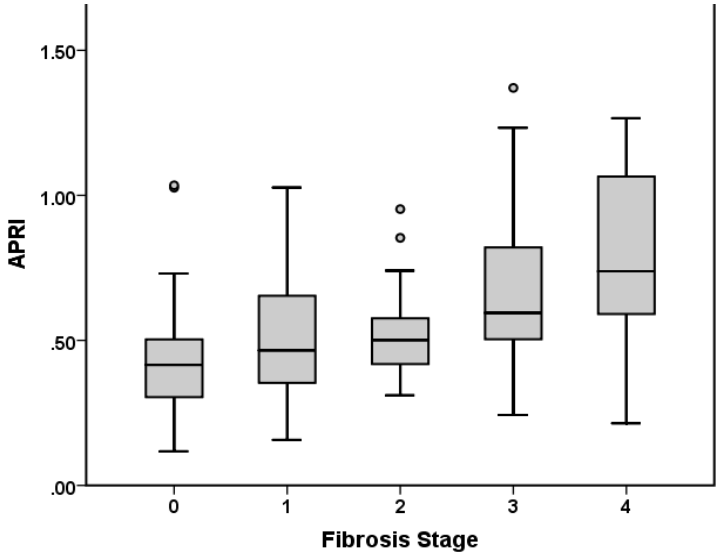
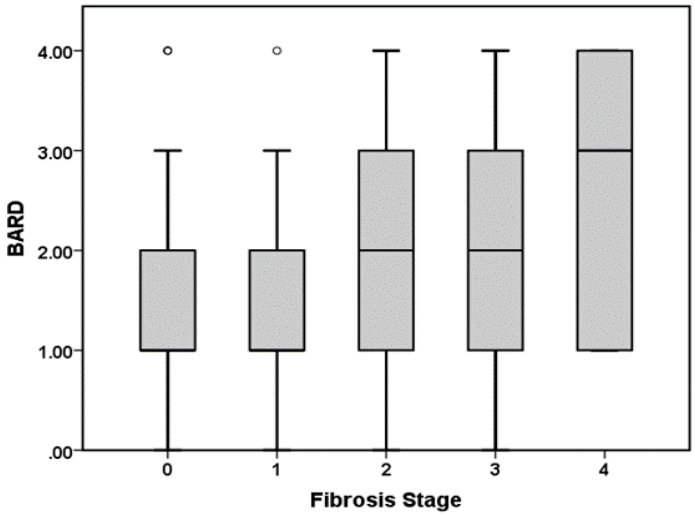


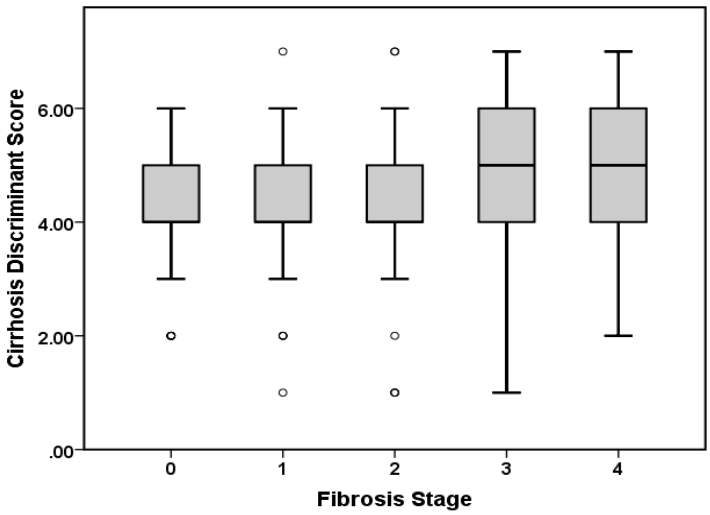
Figure 6.2 Boxplots of indirect serum markers with respect to fibrosis stage in study population

Panel A: BARD, Panel B: Cirrhosis Discriminant Score, Panel C BAAT

A



B



C

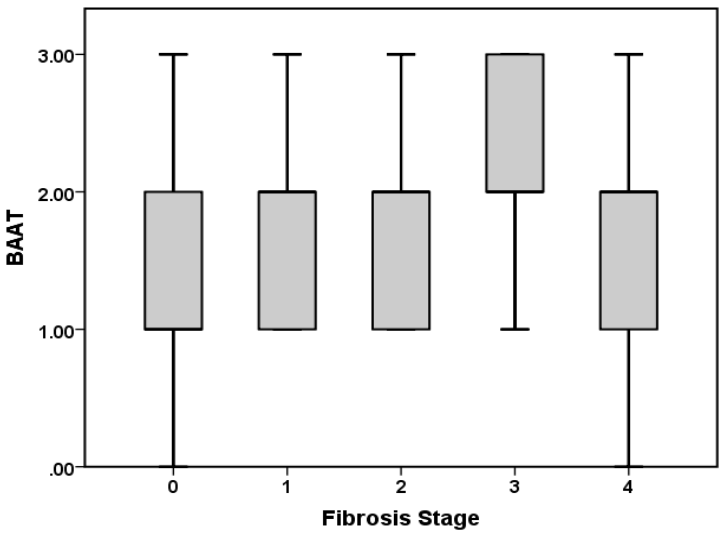
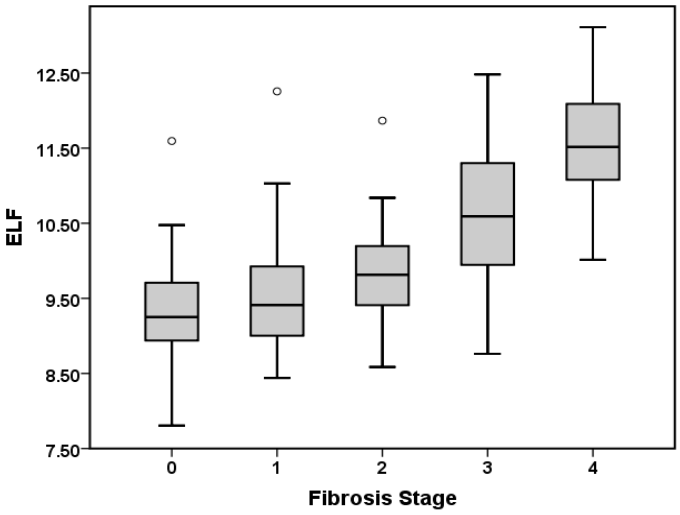


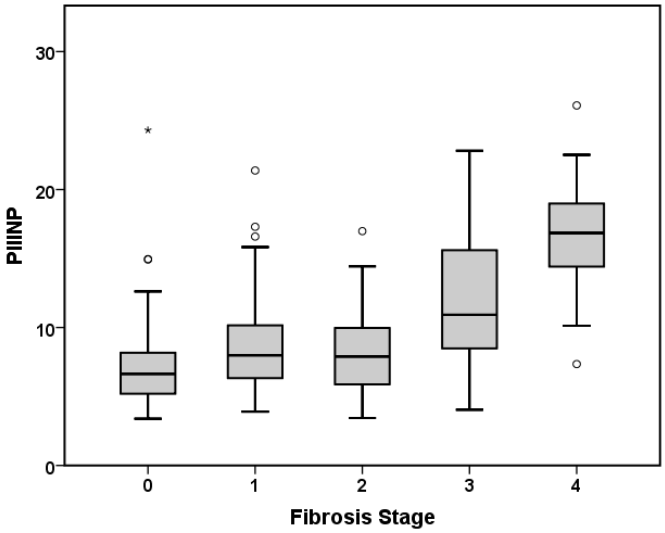
Figure 6.3 Boxplots of direct serum markers with respect to fibrosis stage in study population

Panel A: ELF, Panel B: PIIINP, Panel C HA).

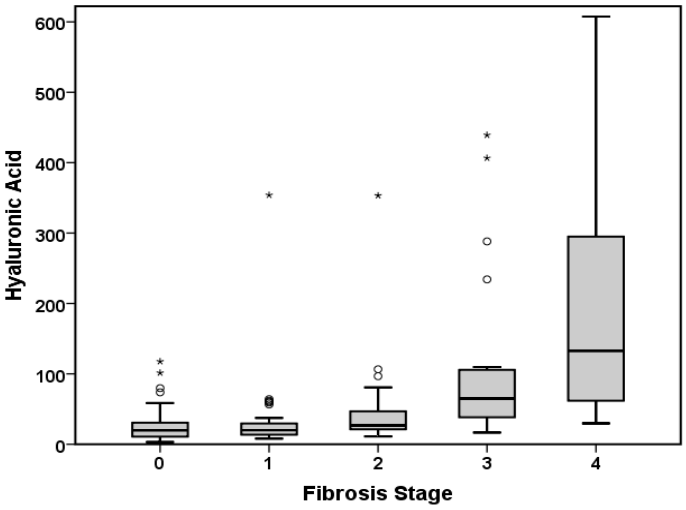
A



B



C



Comparison of the diagnostic performance of all 9 biomarkers

Ability to Discriminate Mild Fibrosis: F0 Vs. F1-4 (tables 6.3 and 6.4)

The diagnostic performance of the 9 biomarkers for the detection of mild liver fibrosis ranged from poor to moderate. APRI was the best performing indirect biomarker (AUROC 0.71). Whereas all direct tests exhibited moderate performance for the detection of mild fibrosis, the best performance was exhibited by ELF (AUROC 0.76); this was significantly greater than both the Cirrhosis Discriminant Score and BARD.

Ability to Discriminate Moderate Fibrosis: F0-1 Vs. F2-4 (tables 6.3 and 6.5)

Moderate diagnostic performance was exhibited by 3 of indirect biomarkers for the detection of moderate fibrosis (APRI 0.73, FIB4 0.78, NAFLD Fibrosis Score 0.76). The AUROC of NAFLD Fibrosis Score and FIB4 were significantly greater than that of Cirrhosis Discriminant Score, BARD and BAAT. Amongst the direct biomarkers only HA (AUROC 0.80) and ELF (AUROC 0.81) exhibited good performance for the detection of moderate fibrosis. The best performing direct biomarker, ELF, generated a significantly higher AUROC than PIIINP ($p<0.001$) but not HA ($p=0.54$).

Ability to Discriminate Advanced Fibrosis: F0-2 Vs. F3-4 (tables 6.3 and 6.6)

Two of indirect biomarkers (FIB4 and NAFLD fibrosis score) exhibited good diagnostic performance for the detection of advanced fibrosis (AUROC 0.83). By contrast, all the direct biomarkers tested demonstrated good diagnostic performance for the detection of advanced fibrosis (AUROC 0.83-0.89). ELF was the best performing biomarker (AUROC 0.89) and demonstrated significantly better diagnostic performance than 6 of the biomarkers including PIIINP ($p<0.001$), but not FIB4 ($p=0.12$), NAFLD Fibrosis Score ($p=0.07$) and HA ($p=0.25$).

Ability to Discriminate Cirrhosis: F0-3 Vs. F4 (tables 6.3 and 6.7)

Three of indirect biomarkers exhibited good performance for the detection of cirrhosis (APRI, FIB4 and NAFLD Fibrosis score). Excellent diagnostic performance for the detection of cirrhosis was exhibited by all 3 direct biomarkers tested (AUROC 0.90 - 0.94). ELF generated the highest AUROC (0.94) which was significantly higher than all the biomarkers other than FIB4 ($p=0.07$) and PIIINP ($p=0.12$). Furthermore, the AUROC generated by HA and PIIINP were not significantly different to those generated by the 3 best performing indirect biomarkers (APRI, FIB4 and NAFLD Fibrosis score).

Overall Performance (Obuchowski Measure)

The ability to discriminate between all degrees of fibrosis was assessed by the Obuchowski measure (table 6.3). ELF was the best test at discriminating between all fibrosis stages ELF (ordROC 0.91). FIB4 was the indirect biomarker with the highest overall discriminatory performance (ordAUROC 0.89) which was higher than PIIINP (ordAUROC 0.88) but not HA (ordAUROC 0.90).

Table 6.3 Comparison of 8 serum markers of fibrosis in their ability to discriminate (AUROC) between mild, moderate, advanced fibrosis, cirrhosis and all degrees of fibrosis (Obuchowski Measure) in the study population (n=177)

Serum Marker	F0 v F1-4 Mild Fibrosis 70 (vs) 107 AUROC *95% Confidence interval*	F0-1 vs F2-4 Moderate Fibrosis 104 (vs) 73 AUROC *95% Confidence interval*	F0-2 vs F3-4 Advanced Fibrosis 135 (vs) 42 AUROC *95% Confidence interval*	F0-3 vs F4 Cirrhosis 159 (vs) 18 AUROC *95% Confidence interval*	Obuchowski Measure OrdROC *Standard Error*
APRI	**0.71** *(0.63-0.78)*	**0.73** *(0.65-0.80)*	**0.78** *(0.69-0.87)*	**0.81** *(0.71-0.92)*	**0.87** *(0.02)*
FIB4	**0.69** *(0.61-0.77)*	**0.78** *(0.72-0.85)*	**0.83** *(0.76-0.91)*	**0.85** *(0.75-0.94)*	**0.89** *(0.01)*
NAFLD Fibrosis Score	**0.67** *(0.59-0.75)*	**0.76** *(0.68-0.83)*	**0.83** *(0.75-0.90)*	**0.86** *(0.78-0.93)*	**0.88** *(0.01)*
Cirrhosis DS	**0.58** *(0.50-0.67)*	**0.59** *(0.50-0.68)*	**0.65** *(0.55-0.75)*	**0.63** *(0.49-0.78)*	**0.80** *(0.02)*
BARD	**0.60** *(0.52-0.68)*	**0.66** *(0.58-0.74)*	**0.68** *(0.58-0.78)*	**0.73** *(0.61-0.86)*	**0.83** *(0.02)*
BAAT	**0.64** *(0.56-0.73)*	**0.62** *(0.53-0.70)*	**0.67** *(0.57-0.77)*	**0.52** *(0.37-0.66)*	**0.81** *(0.02)*
HA	**0.72** *(0.64-0.79)*	**0.80** *(0.74-0.87)*	**0.88** *(0.82-0.94)*	**0.91** *(0.86-0.97)*	**0.90** *(0.01)*
ELF	**0.76** *(0.69-0.83)*	**0.81** *(0.75-0.88)*	**0.89** *(0.84-0.95)*	**0.94** *(0.90-0.98)*	**0.91** *(0.01)*
PIIINP	**0.72** *(0.65-0.80)*	**0.72** *(0.64-0.80)*	**0.83** *(0.80-0.91)*	**0.90** *(0.83-0.96)*	**0.88** *(0.01)*

Table 6.4 Pairwise comparison of the diagnostic performance (AUROC) generated by the 9 biomarkers for the detection of at least mild fibrosis (F1-4) in the study population

The method of Delong is used to compare AUROC for each respective biomarker with the p value for each pairwise comparison AUROC displayed in the table. P values less than 0.05 are displayed in bold.

Test	**FIB4** AUROC 0.69	**APRI** AUROC 0.71	**NFS** AUROC 0.67	**CDS** AUROC 0.58	**BARD** AUROC 0.60	**BAAT** AUROC 0.64	**HA** AUROC 0.72	**PIIINP** AUROC 0.72
APRI - AUROC 0.71	0.61							
NFS - AUROC 0.67	0.51	0.39						
CDS (AUROC 0.58)	**<0.001**	**<0.01**	0.07					
BARD (AUROC 0.60)	0.04	**0.03**	**0.04**	0.80				
BAAT (AUROC 0.64)	0.36	0.23	0.61	0.27	0.46			
HA (AUROC 0.72)	0.55	0.89	0.30	**0.01**	**0.02**	0.18		
PIIINP (AUROC 0.72)	0.54	0.73	0.34	**<0.01**	**0.02**	0.11	0.86	
ELF (AUROC 0.76)	0.11	0.28	**0.05**	**<0.001**	**<0.001**	**0.02**	**0.02**	0.23

Table 6.5 Pairwise comparison of the diagnostic performance (AUROC) generated by the 9 biomarkers for the detection of at least moderate fibrosis (F2-4) in the study population

The method of Delong is used to compare AUROC for each respective biomarker with the p value for each pairwise comparison displayed in the table. P values less than 0.05 are displayed in bold.

Test	**FIB4** AUROC 0.78	**APRI** AUROC 0.73	**NFS** AUROC 0.76	**CDS** AUROC 0.59	**BARD** AUROC 0.66	**BAAT** AUROC 0.62	**HA** AUROC 0.78	**PIIINP** AUROC 0.72
APRI AUROC 0.73	**0.06**							
NFS AUROC 0.76	0.30	0.45						
CDS AUROC 0.59	**<0.001**	**<0.001**	**<0.01**					
BARD AUROC 0.66	**<0.001**	0.21	**0.01**	0.27				
BAAT AUROC 0.62	**<0.001**	**0.04**	**<0.01**	0.60	0.47			
HA AUROC 0.80	0.54	**0.08**	0.17	**<0.001**	**<0.001**	**<0.001**		
PIIINP AUROC 0.72	0.21	0.88	0.48	**0.01**	0.26	**0.04**	**0.03**	
ELF AUROC 0.81	0.44	**0.05**	0.14	**<0.001**	**<0.001**	**<0.001**	0.57	**<0.001**

Table 6.6 Pairwise comparison of the diagnostic performance (AUROC) generated by the 9 biomarkers for the detection of at least advanced fibrosis (F3-4) in the study population

The method of Delong is used to compare AUROC for each respective biomarker with the p value for each pairwise comparison displayed in the table. P values less than 0.05 are displayed in bold.

Test	**FIB4** AUROC 0.83	**APRI** AUROC 0.78	**NFS** AUROC 0.83	**CDS** AUROC 0.65	**BARD** AUROC 0.68	**BAAT** AUROC 0.67	**HA** AUROC 0.78	**PIIINP** AUROC 0.83
APRI AUROC 0.78	0.03							
NFS AUROC 0.83	0.78	0.18						
CDS AUROC 0.65	**<0.001**	**<0.001**	**<0.001**					
BARD AUROC 0.68	**<0.001**	0.13	**<0.001**	0.66				
BAAT AUROC 0.67	**<0.001**	0.10	**<0.01**	0.73	0.88			
HA AUROC 0.78	0.22	**0.03**	0.13	**<0.001**	**<0.001**	**<0.001**		
PIIINP AUROC 0.83	0.99	0.28	0.90	**<0.001**	**0.01**	**0.01**	0.18	
ELF AUROC 0.89	0.12	**0.01**	0.07	**<0.001**	**<0.001**	**<0.001**	0.25	**<0.001**

Table 7.7 Pairwise comparison of the diagnostic performance (AUROC) generated by the 9 biomarkers for the detection of cirrhosis (F4) in the study population

The method of Delong is used to compare AUROC for each respective biomarker with the p value for each pairwise comparison displayed in the table. P values less than 0.05 are displayed in bold.

Test	**FIB4** AUROC 0.85	**APRI** AUROC 0.81	**NFS** AUROC 0.86	**CDS** AUROC 0.63	**BARD** AUROC 0.73	**BAAT** AUROC 0.52	**HA** AUROC 0.78	**PIIINP** AUROC 0.90
APRI AUROC 0.81	0.29							
NFS AUROC 0.86	0.76	0.41						
CDS AUROC 0.63	**<0.001**	**<0.001**	**<0.001**					
BARD AUROC 0.73	0.16	0.39	**0.03**	0.39				
BAAT AUROC 0.52	**<0.001**	**<0.01**	**<0.001**	0.25	**0.04**			
HA AUROC 0.78	0.21	0.11	0.18	**<0.001**	**<0.01**	**<0.001**		
PIIINP AUROC 0.90	0.37	0.19	0.39	**<0.001**	**<0.001**	**<0.001**	0.75	
ELF AUROC 0.94	0.07	**0.04**	**0.04**	**<0.001**	**<0.001**	**<0.001**	**0.05**	0.12

Summary of diagnostic performance

ELF was the best performing biomarker overall. Amongst the direct biomarkers, ELF generated a significantly higher AUROC than its constituent component HA for the detection of cirrhosis ($p=0.05$) and mild fibrosis ($p=0.02$). In addition, ELF generated a higher AUROC than PIIINP for the detection of moderate ($p<0.001$) and advanced fibrosis ($p<0.001$).

Amongst the indirect biomarkers, FIB4 and NAFLD Fibrosis score were the best performing biomarkers exhibiting good performance for the detection of advanced fibrosis and cirrhosis. Overall, Cirrhosis Discriminant Function, BARD and BAAT performed poorly.

Clinical Applicability of Biomarkers of Fibrosis given performance of PIIINP in detecting NASH

- **Cross sectional Detection of Mild fibrosis (F1-F4)**

In this study population, none of the biomarkers tested exhibited either good or excellent performance for the discrimination of mild fibrosis. Moreover, this level of performance was lower than that identified for PIIINP in detecting histologic NASH (AUROC 0.82-0.84).

- **Cross sectional Detection of Moderate Fibrosis (F2-F4)**

Of the nine biomarkers tested only ELF (AUROC 0.81) and HA (AUROC 0.80) exhibited good performance for the detection of mild fibrosis in the study population. Again, this level of performance was lower than that generated by PIIINP for the detection of histologic NASH.

- **Cross sectional Detection of Advanced fibrosis (F3-F4)**

Two of the 6 indirect biomarkers, FIB4 and NAFLD Fibrosis score (both AUROC 0.83) and 3 of the direct biomarkers (AUROC 0.83-0.89) exhibited good performance for the detection of patients with advanced fibrosis. Of note, the AUROC of ELF in the detection of advanced fibrosis was significantly higher than that of PIIINP ($p<0.001$).

- **Cross sectional Detection of Cirrhosis (F4)**

Three indirect biomarkers (APRI 0.81, FIB4 0.85, NAFLD Fibrosis Score 0.86) exhibited good diagnostic performance for the detection of cirrhosis. By contrast all 3 direct biomarkers exhibited excellent performance for the detection of cirrhosis (HA 0.91, ELF 0.94, PIIINP 0.90)

PIIINP levels are influenced by the presence of NASH

Of note, levels of HA (ordROC 0.90), ELF (ordROC0.91) and FIB4 (ordROC0.89) offered superior discrimination between all degrees of fibrosis than PIIINP (ordROC 0.88). This is in keeping with the observation that prior to the development of advanced fibrosis, PIIINP levels are influenced by the presence of NASH.

DISCUSSION

In this study, the performance of 9 serum markers of fibrosis in a population of patients with NAFLD was compared by studying their ability to discriminate between patients with and without mild fibrosis, moderate fibrosis, advanced fibrosis and cirrhosis. Moreover, their ability to discriminate between all fibrosis stages (ordROC) have been studied.

Within the population, none of the biomarkers tested exhibited good performance for the detection of mild fibrosis (F1-4). At best, fair diagnostic performance achieved by one of the indirect biomarkers (APRI AUROC 0.71) and all 3 of the direct markers (AUROC 0.72-0.76) for the detection of mild fibrosis. Two of the 3 direct biomarkers (HA: AUROC 0.80 and ELF: AUROC 0.81) demonstrated good performance for the detection of moderate fibrosis (F2-4), with 3 of the 6 indirect tests demonstrating fair performance (APRI, FIB4 NFS: AUROC 0.73-0.78). All 3 direct biomarkers exhibited good diagnostic performance (AUROC 0.83-0.89) for the detection of advanced fibrosis as did 2 of the 6 indirect biomarkers (FIB4 and NFS AUROC 0.83). Excellent diagnostic performance (AUROC 0.90-0.94) was demonstrated by all 3 direct biomarkers for the detection of cirrhosis. By comparison, 3 of the 6 indirect biomarkers demonstrated good performance (APRI, FIB4 and APRI, AUROC 0.81-0.86) for the detection of cirrhosis. Of note, in this study, TIMP-1 performed less well than the 2 other constituent components of ELF for discriminating between fibrosis stages (detection of mild fibrosis AUROC 0.69, moderate fibrosis AUROC 0.70, advanced fibrosis AUROC 0.77 and cirrhosis AUROC 0.86).

In contrast to AUROC which is an assessment of the cross-sectional fibrosis stage detection, the Obuchowski measure is a measure of the global ability of a biomarker to discriminate between all fibrosis stages. Given the discriminatory performances of

the 9 biomarkers described above, it is not a surprise that ELF attained the highest level of Obuchowski measure (ordROC 0.91) with HA attaining the second highest value (ordROC 0.90). Thereafter, FIB4 was the third best performing biomarker (ordROC 0.89) followed by NAFLD Fibrosis Score (ordROC 0.88), PIIINP (ordROC 0.88) and APRI (ordROC 0.87). The other 3 indirect biomarkers tested all performed more poorly at global fibrosis discrimination (ordROC 0.80-0.83). Whereas BAAT, BARD and Cirrhosis Discriminant Score all performed poorly in this study, the comparisons between the other 3 indirect biomarkers and all 3 of the direct biomarkers tested were more favorable with many of these comparisons statistically insignificant.

These results need to be interpreted in the context of the findings made in chapters 4 and 5. It was identified that serum PIIINP levels can be used to discriminate between NASH and simple steatosis amongst patients with NAFLD. In this chapter, these results indicate that amongst patients without advanced fibrosis, ELF discriminates between fibrosis stages better than PIIINP. However, this level of performance is inferior to the performance of PIIINP at discriminating between patients with and without NASH. Moreover, the diagnostic performance of ELF is superior to that of PIIINP for the detection of advanced fibrosis and cirrhosis and for the discrimination between all fibrosis stages (ordROC). These results emphasise that PIIINP alone is not an ideal biomarker of liver fibrosis as prior to the development of advanced fibrosis, PIIINP levels are heavily influenced by NASH.

It would therefore seem sensible to combine these approaches when attempting to stage the severity of NALFD using serum biomarkers. The presence or absence of steatohepatitis and/or advanced fibrosis could be detected by the application of PIIINP ideally using a diagnostic threshold with high sensitivity such as 5.2 ng/ml (80% sensitive for NASH and 95% sensitive for advanced fibrosis). A PIIINP result below

this threshold will exclude the overwhelming majority of patients with NASH and/or advanced fibrosis. Conversely a PIIINP result above this diagnostic threshold would indicate the need to apply a biomarker of fibrosis that will more reliably stage the severity of liver fibrosis due to NAFLD and identify the presence or absence of advanced fibrosis and cirrhosis.

Progression of NAFLD to cirrhosis is highly important as transition to cirrhosis heralds the advent of complications related to portal hypertension and hepatocellular carcinoma. Considering the failings of liver biopsy and its inherent disadvantages compared to NIT, in this study ELF exhibited almost perfect performance for the detection of cirrhosis thus obviating the need for liver biopsy to make this diagnosis.[200] It is also important, however, to identify patients with fibrosis due to NAFLD prior to the development of cirrhosis. Given the performance of biomarkers demonstrated in this study, it would be feasible select a test that has sufficiently good enough diagnostic performance to detect patients with fibrosis prior to the development of cirrhosis.

In contrast to the cross-sectional detection fibrosis as assessed by AUROC, the Obuchowski measure is an assessment of the global discriminatory function of a biomarker of fibrosis that incorporates a penalty function for incorrectly staged fibrosis. These results confirm that ELF performs as a biomarker that discriminates between all fibrosis stages with good effect.[342] These results certainly allow one to consider ELF as a biomarker of fibrosis that could be used to monitor changes in fibrosis severity rather than simply a biomarker of cross-sectional discrimination.[343] However evaluation of the performance of ELF in this context will require prospective studies comparing ELF with histology. In this way, ELF could be used to monitor both the natural history of disease and also monitor any antifibrotic response may occur in response to therapeutic interventions for liver fibrosis. Moreover, given that ELF has

been found to predict liver related mortality in a cohort of patients with mixed aetiology chronic liver disease it will be interesting if these observations can also be replicated in patients with chronic liver disease due to NAFLD.

The National Institute for Health and Care Excellence (NICE) has endorsed the use of ELF at a diagnostic threshold of 10.51 for the detection of advanced fibrosis in NAFLD.[29] Specifically, these guidelines recommend that a diagnosis of advanced fibrosis can be made or excluded if the ELF score is above and below 10.51 respectively. Within this study cohort, the application of ELF at a threshold of 10.51 resulted in a PPV of 76%, NPV of 92%, successfully detecting 74% of cases of advanced fibrosis, while also incorrectly diagnosing more than a quarter of the prevalent pool of cases with advanced fibrosis as having 'non-advanced' fibrosis (false negative). An alternative to applying ELF at a single threshold is to employ ELF at 2 diagnostic thresholds (combined cut-off) to increase diagnostic accuracy. Whereas a combined cut-off approach results in the generation of indeterminate cases amongst which the severity of fibrosis is uncertain, these indeterminate cases can be retested by an alternative test. In this study, it has been demonstrated that multiple diagnostic tests have good performance for detecting advanced fibrosis in NAFLD. Given these results, a combined cut-off approach using two of these NIT could be a viable option to improve diagnostic performance. In addition, given the higher cost of direct biomarkers, employing an indirect biomarker as the first test in a combined cut-off approach will also result in a reduction of the number of direct tests required.

PIIINP is also currently widely employed in clinical practice to detect methotrexate induced liver fibrosis. Given the ubiquity of NAFLD, these results also demonstrate that rising PIIINP levels do not necessarily imply drug induced fibrosis.

In summary, in this study both indirect and direct fibrosis biomarkers in NAFLD have been validated. The best performing test overall was ELF which was significantly better than the second best performing test, HA, for the detection of mild fibrosis and cirrhosis. Whereas ELF exhibited excellent performance for the detection of cirrhosis, it demonstrated good performance for the detection of moderate and advanced fibrosis but only fair performance demonstrated for the detection of mild fibrosis. Three of the indirect tests (BAAT, BARD and Cirrhosis discriminant score) all performed at a level at which they could not be implemented in clinical practice. The other indirect tests, however, demonstrated much better diagnostic performance for the detection of advanced fibrosis and cirrhosis. This is certainly encouraging as many clinicians do not have access to direct tests due to their associated cost. Given the level of performance of PIIINP for the detection of NASH reported in chapters 4 and 5, these results suggest that a diagnostic target of NASH using PIIINP should be preferred over a diagnostic target of mild or moderate fibrosis when attempting to initially stage NAFLD using serum biomarkers. Thereafter, having ruled out or ruled in NASH using PIIINP, fibrosis can be quantified by a biomarker with good diagnostic performance for the discrimination of advanced fibrosis. Furthermore, given that multiple biomarkers have demonstrated good performance for the detection of advanced fibrosis, further studies should be performed to explore whether diagnostic performance can be enhanced by combining these tests in clinical practice.

CHAPTER 7

NONINVASIVE MARKERS OF FIBROSIS IN NON-ALCOHOLIC FATTY LIVER DISEASE FOR THE DETECTION OF ADVANCED FIBROSIS (KLEINER F3-4): SERIAL AND PARALLEL COMBINATIONS OF SERUM BIOMARKERS ENHANCE DIAGNOSTIC PERFORMANCE FOR THE DETECTION OF ADVANCED FIBROSIS (KLEINER F3-4)

ABSTRACT

Background: Whereas NICE has recommended ELF (direct biomarker) at a single threshold (10.51) for detecting advanced fibrosis (F3-4) in NAFLD, ELF is not widely available. Inexpensive indirect tests are commonly used in clinical practice using both higher and lower thresholds (combined cut-off approach) to classify patients in low risk (<80% sensitivity), indeterminate risk. and high risk (>98% specificity) Other approaches to improve diagnostic performance include combining biomarkers: (1) 'in parallel' to increase accuracy (2) 'in series' to reduce or eliminate indeterminate cases.

Aims: To investigate the use of combinations of indirect and direct tests for the accurate identification of advanced fibrosis in NAFLD

Methods: The diagnostic performance of 9 biomarkers for liver fibrosis in NAFLD were compared in a cohort of patients with NAFLD (n=177). The best performing 3 biomarkers were selected for use in combination in diagnostic algorithms **(**ELF, FIB4 and NAFLD Fibrosis Score). Algorithms using these tests in both series and parallel were developed and their performance estimated using Bayes theorem and published biomarker threshold performance. Thereafter, algorithm performances were validated and compared.

Results: Estimated and actual diagnostic algorithm performances were similar. In this cohort, the application of ELF at a threshold of 10.51 detected 74% of cases with F3-4 with NPV 92%, PPV 74%, and diagnostic accuracy 88% (0% indeterminate). The highest diagnostic accuracy (96%) was attained by the parallel combination of ELF alongside NAFLD Fibrosis Score but with 41% cases indeterminate. The serial application of FIB4 followed by ELF resulted in a 64% reduction in the requirement for ELF testing with diagnostic accuracy up to 88% (19% indeterminate) and 85% (0% indeterminate). Considering strategies without ELF testing altogether, the parallel and serial application of FIB4 and NAFLD Fibrosis Score attained a diagnostic accuracy of 93% (49% indeterminate) and 82-88% (0-20% indeterminate) respectively.

Conclusions: these findings confirm that strategies to maximise the diagnostic potential of biomarkers of F3-4 in NAFLD can successfully employed in clinical practice.

INTRODUCTION

As previously outlined, liver biopsy is unsuitable as the primary method to detect and quantify fibrosis in CLDs including NAFLD in primary care, the most common cause of chronic liver disease worldwide.[28] As a result, non-invasive serum biomarkers have been developed to detect liver fibrosis. In the previous chapter, the diagnostic performance of both indirect and direct biomarkers was validated for the detection of fibrosis in NAFLD. The diagnostic targets of these biomarkers included mild (F1), moderate (F2) and advanced (F3-4) liver fibrosis and cirrhosis (F3-4).[322] Recent data suggest that both morbidity and mortality related to NAFLD rises considerably after progression to advanced fibrosis. As such, the results presented in chapter 6 suggest that the detection of patients at the stage of advanced fibrosis is a suitable compromise between the capability of current biomarkers to detect fibrosis and the desire to identify patients with progressive disease who are at risk of complications[344] at a point in the evolution of their disease when interventions may reduce morbidity and mortality. Regardless, biomarkers are reported as continuous variables with AUROC used to assess their performance in the detection of a specified categorical stage of fibrosis.[345] In the absence of a 'perfect' diagnostic test which has an AUROC of 1 (100% sensitivity and 100% specificity) the cross-sectional assessment of liver fibrosis relies on the conversion of a continuous biomarker score into a categorical variable by the use of a diagnostic threshold (Figure 7.1). Instead of using a single diagnostic threshold, 2 diagnostic thresholds can be employed using a combined cut-off approach in which a lower threshold with higher sensitivity is employed to 'rule out' disease and an upper threshold with higher specificity employed 'to rule' in disease, with biomarker scores between the 2 thresholds classified as indeterminate (Figure 7.2).

Figure 7.1 Biomarker Application using a single cut off approach

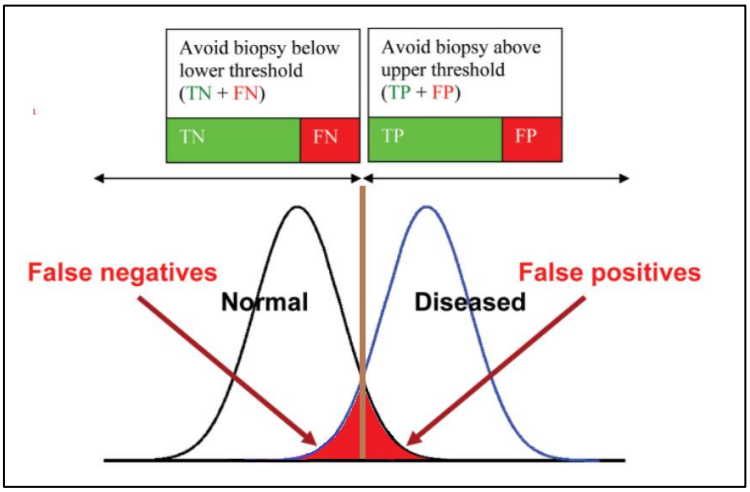
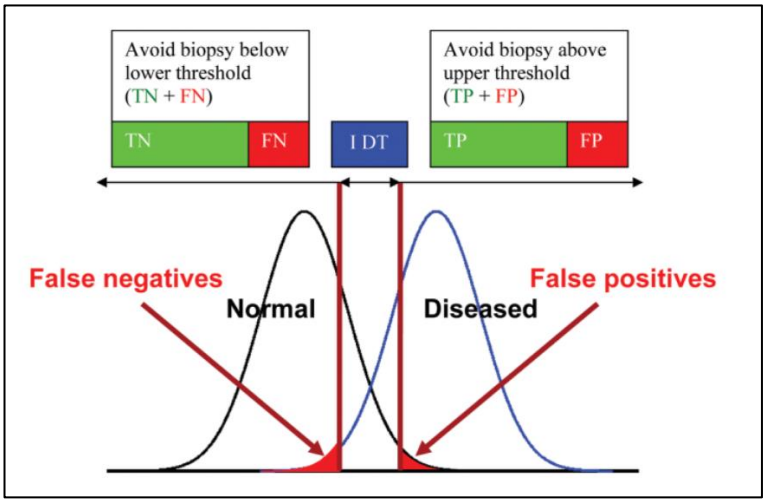


Figure 7.2 Biomarker Application using a combined cut off approach



In comparison to the use of a single diagnostic threshold, a combined cut-off approach enhances diagnostic accuracy but generates a category of indeterminate cases amongst which the severity of fibrosis is uncertain. Conventionally the reference standard, liver biopsy is employed to reassign these indeterminate cases. In clinical practice, however, multiple non-invasive diagnostic tests are available. It is therefore possible to limit the number of cases that require liver biopsy to attain a diagnosis by the application of multiple tests in a way that maximises their diagnostic potential.

NICE has recommended the use of ELF at a single threshold of 10.51 to detect advanced fibrosis (F3-4) in NAFLD.[29] However, ELF is currently not widely available and instead inexpensive indirect tests such as FIB4 are and NAFLD Fibrosis Score are commonly used in clinical practice using a combined cut-off approach.

In this chapter, different diagnostic strategies that combine biomarkers of advanced fibrosis as diagnostic algorithms in NAFLD are explored with the aim of enhancing diagnostic performance. The first aim of this study is to increase diagnostic accuracy. The second aim of this study is to develop algorithms that maintain good diagnostic accuracy but with a reduction in the number, or elimination of cases classified as indeterminate. The first aim will be addressed by the application of a parallel algorithm in which with two biomarkers are employed simultaneously in a combined cut-off approach (a positive or negative algorithm result is conferred by both tests being positive or negative respectively).[346] This approach has the potential to increase specificity and sensitivity amongst cases not classified as indeterminate.[347] The second aim will be addressed by using a series algorithm in which cases classified as indeterminate by a first biomarker are retested by another biomarker. Assuming conditional independence, this approach should reduce the number of cases classified whilst maintaining diagnostic accuracy. Finally given the differential costs of ELF and the indirect biomarkers, strategies will be derived with a reduced requirement for ELF testing whilst preserving diagnostic performance. To facilitate algorithm development, the diagnostic performances of these algorithms will be estimated using Bayes' theorem and published biomarker performance. Given the current NICE recommendations, a comparison of the performance of these diagnostic strategies to the performance of ELF itself at the 10.51 threshold will also be performed.

METHODS

Biomarker Comparison and selection of ELF, FIB4 and NAFLD Fibrosis Score

In chapter 6, the performance of 9 biomarkers (ELF, PIIINP, Hyaluronic acid, BAAT, BARD, Cirrhosis Discriminate Score, NAFLD Fibrosis score, APRI, FIB4) advanced fibrosis (F3-4) were compared in a cohort of patients NAFLD recruited from 2 liver centres.[348] Of the 9 biomarkers tested as continuous variables, ELF, FIB4 and NAFLD Fibrosis score attained an AUROC of greater than 0.8 (ELF 0.89, FIB4 0.83, NAFLD Fibrosis score 0.83). Thereafter, logistic regression was employed to assess statistical independence. Only ELF (OR 4.791, 95% CI 2.549-9.007) and NAFLD Fibrosis Score (OR 1.780, 95% CI 1.163-2.725) were significantly associated with a histologic diagnosis of advanced fibrosis after accounting for the remaining 6 variables for potential confounders. Given that direct biomarkers such as ELF and its constituent components are often not freely available in clinical practice, an analysis of the 6 other indirect biomarkers was also performed. Only FIB4 (OR 2.714, 95% CI 1.221-6.033) and NAFLD Fibrosis Score (OR 1.704, 95% CI 1.053-2.757) were significantly associated with a histologic diagnosis of advanced fibrosis after accounting for the remaining 4 variables as potential confounders.

As a result, 3 biomarkers were used for this study: ELF (direct biomarker) and FIB4 (indirect biomarker) and NAFLD Fibrosis score (indirect biomarker).

Diagnostic Strategy Development

The following definitions were used to facilitate diagnostic strategy development:

- **Single Cut-off***:*

The application of a biomarker at a single diagnostic threshold with biomarker scores above and below the threshold ruling in and ruling out disease respectively.

- **Combined Cut-off:**

The application of a biomarker at two diagnostic thresholds with biomarker scores above the higher threshold ruling in disease, scores below the lower threshold ruling out disease, and scores between the 2 thresholds classified as indeterminate.

- **Fully Assigned Terminal Approach:**

A diagnostic strategy that terminates with a biomarker applied at a single cut-off.

- **Indeterminate Terminal Approach:**

A diagnostic strategy that terminates with a biomarker applied at a combined cut-off.

Biomarker Strategies:

- **Parallel Algorithm**

A diagnostic strategy in which two biomarkers are applied simultaneously: a positive or negative algorithm result is conferred by both tests being positive or negative respectively and discordant results classified as indeterminate. The two biomarkers themselves may be applied using either a single of combined cut-off.

- **Serial Algorithm**

A diagnostic strategy combining two or more biomarkers. A first biomarker is applied in a combined cut off approach. The indeterminate results generated after the application a test are retested by a second biomarker at single cut off (fully assigned approach) or a combined cut off (indeterminate approach).

Biomarker Strategy development and Performance Estimation

- **Identification of Combined Cut-Off Thresholds for the 3 biomarkers**

The original publications of the 3 biomarkers in NAFLD were used to identify both the combined cut-off thresholds and estimated diagnostic performance.[221, 252, 341] The original publication of ELF in NAFLD described a lower threshold of 9.5 (sensitivity 81%, specificity 90%, LR+ 8.1, LR- 0.21, DOR 38.4) and an upper threshold of 10.7 (sensitivity 38%, specificity 98% LR+ 19.0, LR- 0.63, DOR 30.0). The original publication reporting the application of FIB4 in NAFLD described a lower threshold of 1.30 (sensitivity 74%, specificity 71% LR+ 2.6, LR- 0.37, DOR 7.0) and an upper threshold of 2.67 (sensitivity 33%, specificity 98% LR+ 16.5, LR- 0.68, DOR 24.1). Finally, the original publication of the NAFLD Fibrosis score described a lower threshold of -1.455 (sensitivity 82%, specificity 77% LR+ 3.6, LR- 0.23, DOR 15.3) and an upper threshold of 0.676 (sensitivity 51%, specificity 98% LR+ 25.5, LR- 0.50, DOR 51.0).

- **Single Cut off Threshold Identification**

Given the recent NICE guidance, a score of 10.51 was used as a single cut off threshold for ELF. For both FIB4 and NAFLD Fibrosis Score, as a single diagnostic threshold had not been proposed in their respective publications, biomarker scores at the mid-point between the 80% sensitivity and 98% specificity thresholds were selected for use as a single cut off threshold (NAFLD Fibrosis Score -0.779, FIB4 1.99).

Algorithm Performance Estimation using Bayes' Theorem

Algorithm performance was estimated with the assumption that the 3 biomarkers exhibited consistent diagnostic performance at their diagnostic thresholds regardless of whether they were employed in a single biomarker strategy or an algorithmic strategy (conditional independence). To estimate diagnostic performance, the prevalence of advanced fibrosis was modelled on the prevalence of advanced fibrosis identified in the study patient cohort. By the application of Bayes' theorem[204] , the positive and negative predictive values, diagnostic accuracy and the number of cases classified as indeterminate were calculated using formulae inputted into Microsoft Excel. Importantly, a revised prevalence of fibrosis at each diagnostic stage was calculated.

Estimated Biomarker and Algorithm Performance using Bayes Theorem (Table 7.1):

- *Single biomarker, Combined Cut-off, Indeterminate Approach*

  This approach was estimated to perform with diagnostic accuracy of between 89-93%, classifying 16-30% cases as indeterminate.

- *Two biomarker, Parallel algorithm, Indeterminate Approach*

  This approach was estimated to enhance NPV (97-99%), PPV (99%) and overall diagnostic accuracy (98-99%) but with more cases (42%-53%) classified as indeterminate.

- *Two biomarker, Serial algorithm, Indeterminate Approach*

  This approach was estimated to reduce the number of indeterminate cases by one sixth to between 3%-7% whilst maintaining or increasing diagnostic accuracy.

Table 7.1 Estimated Performance of the Biomarker Strategies (derived using Bayes theorem)

Test 1	Test 2	Test 3	Biomarker Strategy: Single, Serial or Parallel	Terminal Approach: Indeterminate or Fully Assigned	PPV%	NPV%	Diagnostic Accuracy (TP+TN)	% Patients requiring ELF Tests	% of Population with Indeterminate Results
ELF	-	-	Single	Indeterminate	86%	94%	93%	100%	16%
FIB4	-	-	Single	Indeterminate	84%	90%	89%	0%	30%
NFS	-	-	Single	Indeterminate	89%	93%	92%	0%	23%
ELF	FIB4	-	Parallel	Indeterminate	99%	98%	98%	100%	47%
ELF	NFS	-	Parallel	Indeterminate	99%	99%	99%	100%	42%
FIB4	NFS	-	Parallel	Indeterminate	99%	97%	98%	0%	53%
ELF	FIB4	-	Series	Indeterminate	84%	94%	93%	100%	5%
ELF	NFS	-	Series	Indeterminate	85%	94%	93%	100%	4%
FIB4	ELF	-	Series	Indeterminate	81%	92%	91%	40%	4%
FIB4	NFS	-	Series	Indeterminate	82%	92%	91%	0%	7%
NFS	ELF	-	Series	Indeterminate	87%	94%	93%	37%	3%
NFS	FIB4	-	Series	Indeterminate	89%	93%	92%	0%	7%

Proposed Diagnostic Strategies

Given these findings, the following diagnostic strategies were selected for validation in the NAFLD cohort (presented in figures 7.3-7.9).

Figure 7.3 Single biomarker employed at single cut-off

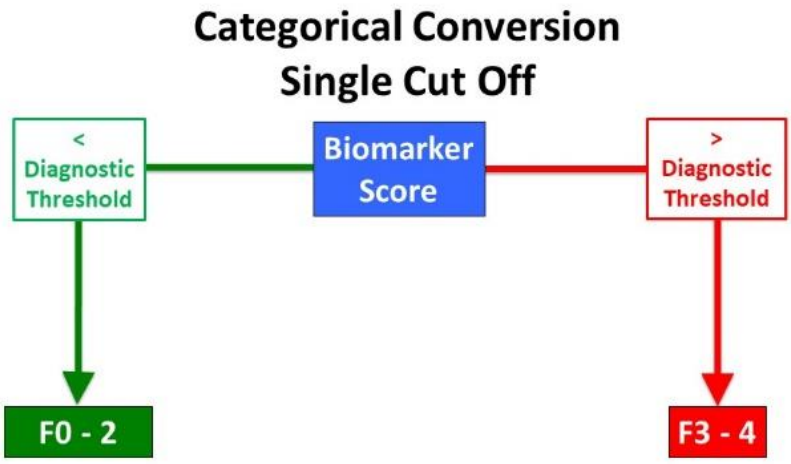


Figure 7.4 Single biomarker employed at combined cut-off

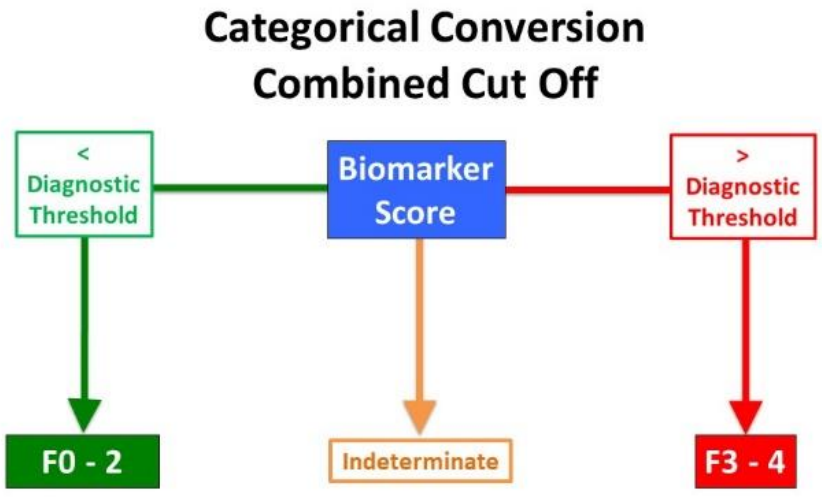


Figure 7.5 Two Biomarker Algorithm: Parallel Strategy with Indeterminate Terminal Approach

ELF is applied using a single cut-off and the indirect tests applied using a combined cut-off

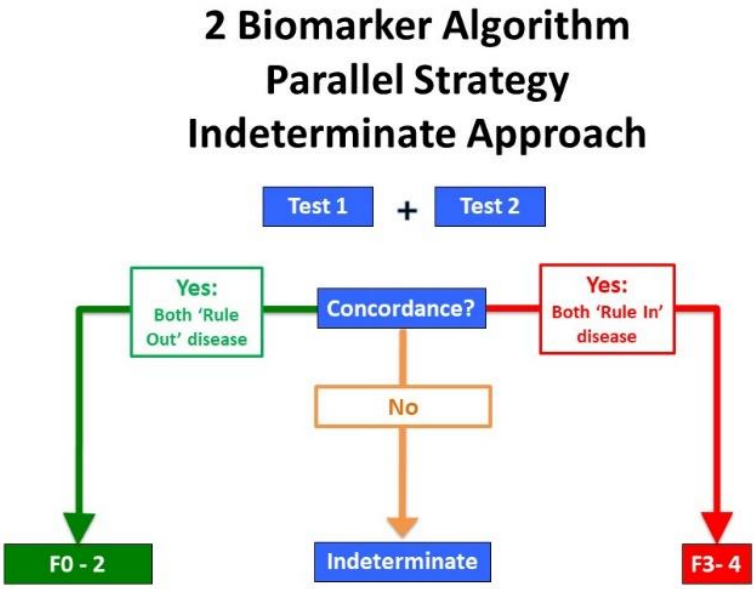


Figure 7.6 Two Biomarker Algorithm: Serial Strategy with Indeterminate Terminal Approach

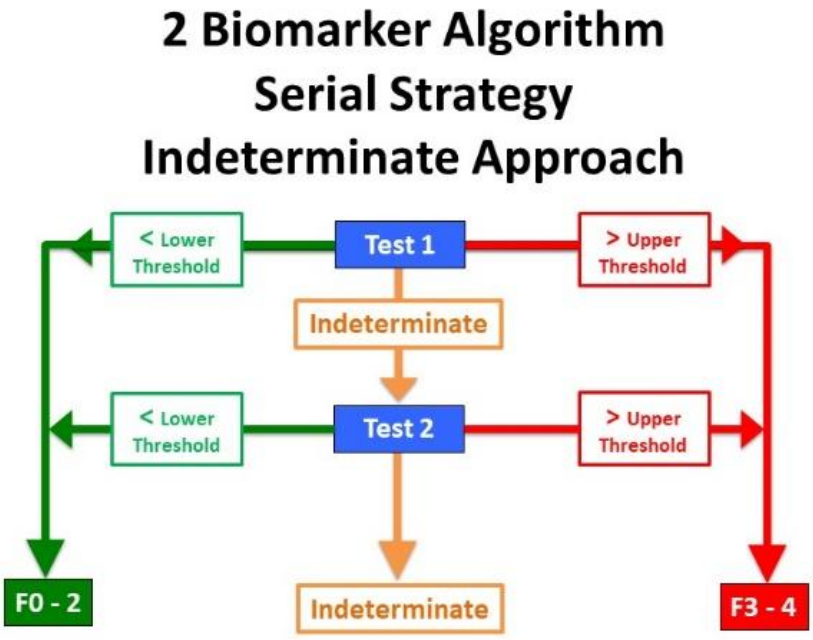


Figure 7.7 Two Biomarker Algorithm: Serial Strategy with Fully Assigned Terminal Approach

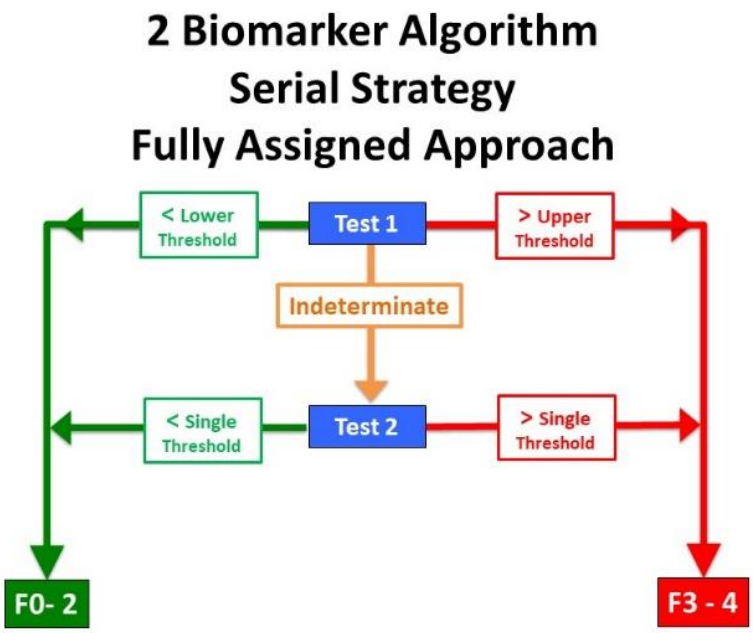


Figure 7.8 Three Biomarker Algorithm: Serial Strategy with Indeterminate Terminal Approach

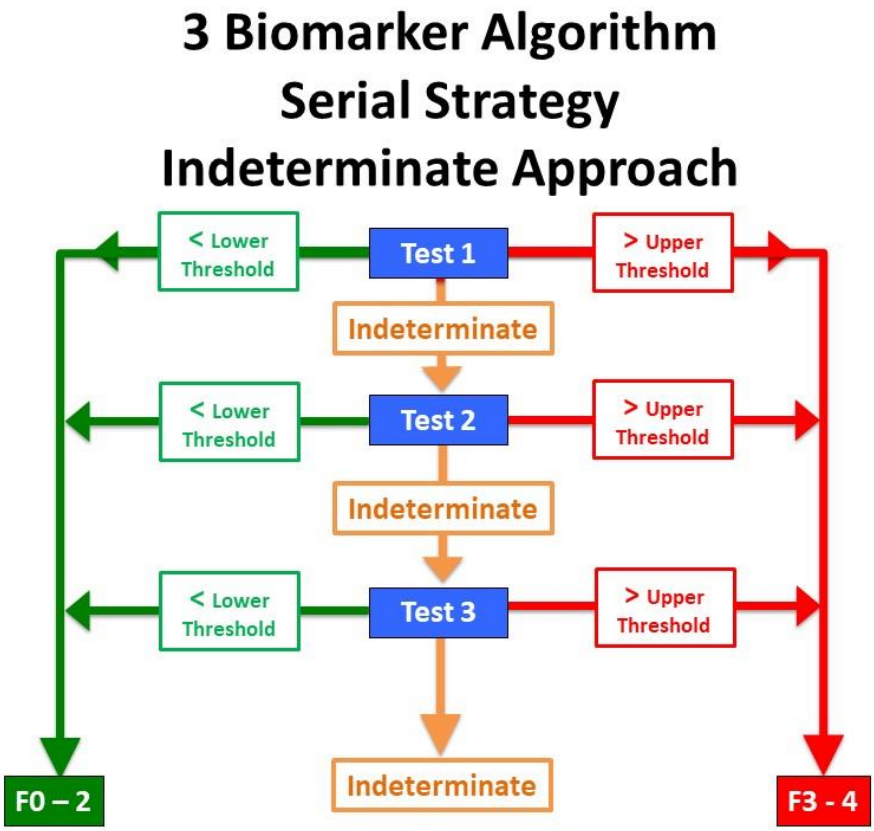
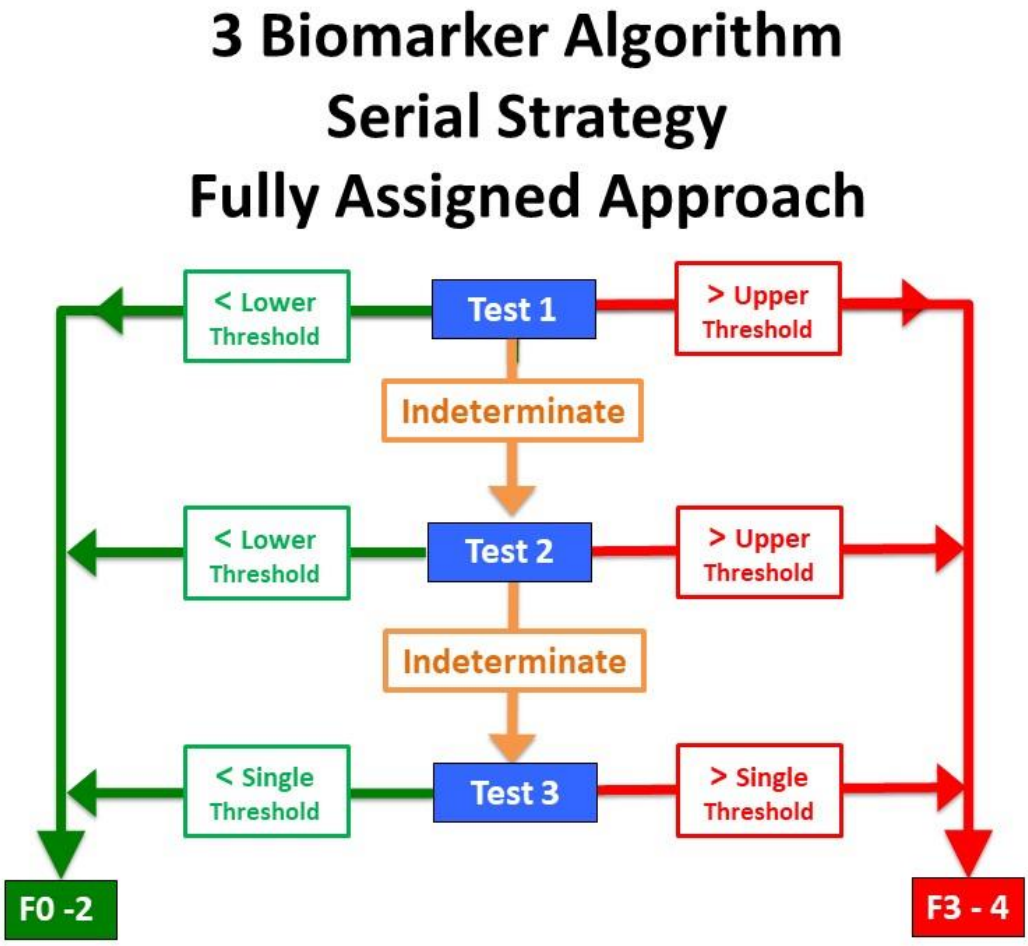


Figure 7.9 Three Biomarker Algorithm: Serial Strategy with Fully Assigned Terminal Approach



Assessment of Diagnostic Accuracy and Statistical Analyses within NAFLD cohort

Data were analysed using SPSS v. 22.0 (SPSS Inc., Chicago, IL) and STATA v. 11.0 (SAS Inc). After the categorisation of biomarker scores by the diagnostic strategies, diagnostic performance was assessed in several ways including the ability of the biomarkers to discriminate between non-advanced (F0-2) and advanced fibrosis (F3-4) as assessed by AUROC, predictive values (positive and negative) and diagnostic accuracy (sum of the percentage of true negatives and positives). Moreover, as the aim of these strategies is to identify patients with advanced fibrosis amongst the population at risk, the percentage of patients correctly identified with advanced fibrosis was calculated 1) amongst patients with a non-indeterminate result (i.e. a positive or negative result) 2) within the overall diagnostic strategy (positive, negative and indeterminate results). For each of the diagnostic strategies, the proportion of patients with an indeterminate score and the proportion of patients who require ELF testing were calculated. Given that spectrum bias can influence biomarker performance, the degree of spectrum bias was assessed by calculation of the DANA.[203] The DANA score can range between 1 and 4 depending on the distribution of fibrosis stages in the population. A DANA of 4 is attained when the population is comprised entirely of patients with F0 and F4 fibrosis. A DANA of 1 is attained when there is central clustering of fibrosis stages with a population composed of patients with F2 and F3 fibrosis only. A uniform distribution of fibrosis stages is defined by a prevalence of 0.20 for each of the 5 stages (uniform prevalence).

RESULTS

Real-world NAFLD Cohort

The baseline characteristics of the study cohort is shown in table 7.2.

The prevalence of advanced fibrosis (24%) in the NAFLD cohort was comparable to those of the populations in the original reference studies of the biomarkers (prevalence of advanced fibrosis in the original studies: FIB4 23%, NAFLD Fibrosis Score 27%, ELF 17%, (table 7.2). The DANA in the study population of the study cohort (2.76) was also comparable to the original studies of NFS and ELF (2.88 and 2.80 respectively) but higher than the original study of FIB4 (2.44) (table 7.3).

Table 7.2 Diagnostic performance of the 3 biomarkers in their index publications and comparison with their performance in the study population (combined and single cut-offs)

Test	Threshold	Sensitivity		Specificity		Likelihood ratio -VE		Likelihood ratio +VE		Diagnostic Odds Ratio	
		Original Publication	Performance in NAFLD cohort	Original Publication	Performance in NAFLD cohort	Original Publication	Performance in NAFLD cohort	Original Publication	Performance in NAFLD cohort	Original Publication	Performance in NAFLD cohort
Combined Cut-Off: Lower Threshold											
ELF	9.5	81%	95%	90%	60%	0.21	0.08	8.1	2.4	38.4	28.5
FIB4	1.30	74%	86%	71%	65%	0.37	0.22	2.6	2.5	7	11.4
NFS	-1.455	82%	72%	77%	71%	0.23	0.39	3.6	2.5	15.3	6.3
Combined Cut-Off: Upper Thresholds											
ELF	10.7	38%	59%	98%	95%	0.63	0.43	19.0	11.8	30.0	27.3
FIB4	2.67	33%	35%	98%	98%	0.68	0.66	16.5	17.5	24.1	26.4
NFS	0.676	51%	24%	98%	98%	0.50	0.78	25.5	12.0	51.0	15.5
Single Threshold											
ELF	10.51	-	74%	-	92%	-	0.28	-	9.3	-	33.0
FIB4	1.99	-	55%	-	86%	-	0.52	-	3.9		7.5
NS	-0.779	-	60%	-	87%	-	0.45	-	4.6		10.2

Table 7.3 Spectrum Analysis of (1) the study population (2) the indeterminate cases generated after applying a biomarker with a combined cut-off approach in study population (3) the populations from which the 3 biomarkers were originally validated

	(1)	(2)			(3)		
	Entire Study n=177	Indeterminate cases after application of ELF at combined cut-off n=55	Indeterminate cases after application of FIB4 at combined cut-off n=64	Indeterminate cases after application of NFS at combined cut-off n=66	FIB4 Original Study	NFS Original Study	ELF Original Study
F0	40%	35%	27%	34%	26%	33%	54%
F1	19%	17%	17%	13%	29%	26%	27%
F2	18%	27%	23%	16%	22%	14%	2%
F3	14%	17%	22%	23%	16%	13%	15%
F4	10%	5%	11%	14%	7%	14%	2%
F0+F1+F2	76%	78%	67%	63%	77%	73%	83%
F3+F4	24%	22%	33%	37%	23%	27%	17%
ADVANCED F34	3.42	3.23	3.33	3.38	3.30	3.52	3.12
NON ADVANCED F012	0.66	0.85	0.86	0.64	0.86	0.64	0.39
DANA F34	2.76	2.37	2.48	2.74	2.44	2.88	2.80

Comparison of the estimated and actual performances of the diagnostic strategies

Compared to their original publications, particularly at their lower thresholds the 3 biomarkers exhibited more comparable sensitivity but lower specificity in the NAFLD cohort. As a result, the process of conversion generated a larger proportion of indeterminate cases (ELF: 32%, FIB4: 36%, NFS: 37%) in the NAFLD cohort than were estimated (ELF: 16%, FIB4: 30%, NFS: 23%). Regardless, the estimated diagnostic performances of these strategies were not dissimilar to the performances that they demonstrated in the NAFLD cohort (Tables 7.1 and 7.4). As expected, all the 3 permutations of the two biomarker parallel algorithm produced an increase in NPV, PPV and diagnostic accuracy with an increase in the number of cases classified as indeterminate. Similarly, the serial indeterminate algorithms maintained diagnostic accuracy whilst reducing the number of indeterminate cases.

Performance of ELF at a threshold of 10.51 (Fully Assigned Approach) and comparison of the multiple diagnostic strategies (table 7.4)

At this threshold, ELF identified 74% of the cases with advanced fibrosis, with PPV 74%, NPV 92%, diagnostic accuracy 88% without generating any indeterminate cases. By contrast, using a fully assigned approach FIB4 and NAFLD Fibrosis Score identified 55% and 60% of patients with advanced fibrosis, attaining diagnostic accuracies of 79% and 81% respectively. Thereafter, the diagnostic performances of these multiple (n=25) strategies have been presented on a single table (with the performance of ELF at a threshold of 10.51 highlighted in yellow) to help facilitate the comparisons described below:

Strategies to increase Positive Predictive Value (table 7.5)

The highest PPV was attained by using parallel strategies with an indeterminate approach; importantly this was achieved with or without the use of ELF. For example, the parallel strategy, indeterminate approach algorithms containing NAFLD Fibrosis Score with either ELF or FIB4 attained a PPV of 100% albeit with 41% and 49% of the population classified as indeterminate respectively.

Strategies to increase Negative Predictive Value (table 7.6)

All strategies that resulted in a NPV of 95% or more used ELF as the spine of the algorithm (serial or parallel) in an indeterminate approach. The highest NPV (98%) was attained by the application of ELF and FIB4 using a parallel strategy, indeterminate approach with 41% of cases classified as indeterminate.

Strategies to increase Diagnostic Accuracy (table 7.7)

The highest diagnostic accuracies were attained by parallel strategies (indeterminate approach) using ELF with either NAFLD Fibrosis Score (96%) or FIB4 (95%); both strategies generated 41% of cases as indeterminate. The third highest diagnostic accuracy (93%) was attained by the application of NAFLD Fibrosis Score and FIB4 also in a parallel strategy, indeterminate approach with 41% of cases classified as indeterminate.

Strategies to increase the detection of advanced fibrosis amongst patients who test either positive or negative (table 7.8)

In contrast to diagnostic accuracy and PPV, this is a measure of disease detection amongst individuals with a 'non-indeterminate' score (true positive cases divided by the sum of true positive and false negative cases). In general, the strategies that attained the best performance in this regard all used ELF as their spine employing an indeterminate approach. For example, the application ELF alone using an indeterminate approach detected 93% of cases advanced fibrosis (32% indeterminate) amongst patients with a non-indeterminate score. A serial strategy using ELF followed by FIB4 using an indeterminate approach detected 87% of patients with advanced fibrosis in this context with only 19% of cases classified as indeterminate. Of note the use of FIB4 and NAFLD Fibrosis score in a parallel strategy indeterminate approach detected 82% of patients with advanced fibrosis with positive or negative result albeit with 49% of cases classified as indeterminate.

Strategies to increase the proportion of F3-4 detected amongst entire population (table 7.9)

This is a measure is of disease detection by a single strategy regardless of whether the approach is fully assigned or indeterminate. The 3 strategies that detected the highest proportion of patients with advanced fibrosis in this population required ELF tests to be performed on all cases: either alone in a fully assigned approach (74% detected) or in fully assigned serial strategy with NAFLD Fibrosis Score (79%) or FIB4(74%).

Strategies to improve diagnostic accuracy whilst testing all patients with ELF (table 7.10)

These results indicate that the diagnostic accuracy of ELF at a threshold of 10.51 in a fully assigned approach can be enhanced by the use of parallel and serial strategies but at the cost of introducing indeterminate cases and thus reducing overall disease detection by the strategy as a whole.

Strategies to improve diagnostic performance whilst reducing the requirement (64-80%) for ELF testing (table 7.11)

Serial strategies that used FIB4 as the spine of the algorithm with ELF as the second test reduced the need for ELF testing by 64%. A serial strategy combining FIB4 and ELF in an indeterminate approach (19% indeterminate) attained diagnostic accuracy of 91% detecting 57% of patients with advanced fibrosis in the entire population. Similarly, a serial strategy combining FIB4 and ELF in a fully assigned approach (0% indeterminate) achieved a diagnostic accuracy of 88% and detecting 67% of patients with advanced fibrosis in the entire population. A serial strategy that terminates with ELF after the application of FIB4 and NAFLD Fibrosis score offered an 80% reduction in the need for ELF testing albeit with a diagnostic accuracy of 86% and only 6% cases classified as indeterminate (indeterminate approach) and 85% (fully assigned approach).

Enhancing performance using strategies that do not use ELF (table 7.12)

Amongst strategies that were composed entirely of FIB4 and NAFLD Fibrosis Score, parallel strategies attained 93% diagnostic accuracy but with 49% of patients classified

as indeterminate. Serial strategies using the indirect tests in an indeterminate approach (20% indeterminate) and fully assigned approach attained diagnostic accuracies of 88% and 82% respectively and detected 57% and 60% of patients with advanced fibrosis respectively.

Enhancing performance using fully assigned (non-indeterminate) strategies (table 7.13)

The best performing fully assigned strategy was the use of ELF alone at a threshold of 10.51 (diagnostic accuracy 88%, detecting 74% of patients with advanced fibrosis in the population). The application of FIB4 followed by NAFLD Fibrosis score in a serial strategy was the best fully assigned strategy that did not contain ELF (diagnostic accuracy 82%, detecting 60% of patients with advanced fibrosis in the population).

Table 7.4 Summary of the diagnostic strategies tested and their diagnostic performance (listed in the order they were described in the chapter)

The fully assigned approach using ELF at the 10.51 threshold is highlighted in yellow.

Algorithm	Test 1	Test 2	Test 3	Biomarker Strategy: Single, Serial or Parallel	Terminal Approach: Indeterminate or Fully Assigned	PPV%	NPV%	Diagnostic Accuracy (TP+TN)	% F3-4 detected amongst patients with positive or negative result	% Patients requiring ELF Tests	% of Population with Indeterminate Results	% of F3-4 detected within entire population by strategy	AUC
-	ELF	-	-	Single	Fully Assigned	74%	92%	88%	74%	100%	0%	74%	0.83
-	ELF	-	-	Single	Indeterminate	77%	97%	92%	93%	100%	32%	62%	0.86
-	FIB4	-	-	Single	Indeterminate	80%	94%	91%	71%	0%	36%	36%	0.80
-	NFS	-	-	Single	Indeterminate	82%	89%	88%	43%	0%	37%	21%	0.74
1	ELF & FIB4		-	Parallel	Indeterminate	82%	98%	95%	88%	100%	41%	33%	0.84
2	ELF & NFS		-	Parallel	Indeterminate	100%	96%	96%	81%	100%	41%	21%	0.83
3	FIB4 & NFS		-	Parallel	Indeterminate	100%	93%	93%	82%	0%	49%	21%	0.76
4	ELF	FIB4	-	Series	Indeterminate	75%	94%	91%	87%	100%	19%	64%	0.88
5	ELF	NFS	-	Series	Indeterminate	74%	95%	90%	81%	100%	16%	62%	0.86
6	FIB4	ELF	-	Series	Indeterminate	80%	92%	91%	77%	36%	19%	57%	0.85
7	FIB4	NFS	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
8	NFS	ELF	-	Series	Indeterminate	77%	90%	88%	63%	32%	16%	48%	0.80
9	NFS	FIB4	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
10	ELF	FIB4	-	Series	Fully Assigned	63%	91%	84%	74%	100%	0%	74%	0.80
11	ELF	NFS	-	Series	Fully Assigned	64%	93%	85%	79%	100%	0%	79%	0.83
12	FIB4	ELF	-	Series	Fully Assigned	78%	90%	88%	67%	36%	0%	67%	0.80
13	FIB4	NFS	-	Series	Fully Assigned	63%	88%	82%	60%	0%	0%	60%	0.74
14	NFS	ELF	-	Series	Fully Assigned	77%	87%	85%	55%	32%	0%	55%	0.75
15	NFS	FIB4	-	Series	Fully Assigned	55%	85%	79%	50%	0%	0%	50%	0.69
16	ELF	FIB4	NFS	Series	Indeterminate	73%	95%	90%	82%	100%	11%	64%	0.87
17	FIB4	ELF	NFS	Series	Indeterminate	77%	93%	90%	73%	36%	11%	57%	0.84
18	FIB4	NFS	ELF	Series	Indeterminate	71%	91%	86%	69%	20%	6%	64%	0.80
19	ELF	FIB4	NFS	Series	Fully Assigned	66%	93%	85%	69%	100%	0%	69%	0.83
20	FIB4	ELF	NFS	Series	Fully Assigned	70%	91%	85%	71%	36%	0%	71%	0.81
21	FIB4	NFS	ELF	Series	Fully Assigned	75%	88%	85%	57%	20%	0%	57%	0.76

Table 7.5 Diagnostic strategies evaluated presented in order of descending PPV

The fully assigned approach using ELF at the 10.51 threshold is highlighted in yellow.

Algorithm	Test 1	Test 2	Test 3	Biomarker Strategy: Single, Serial or Parallel	Terminal Approach: Indeterminate or Fully Assigned	PPV%	NPV%	Diagnostic Accuracy (TP+TN)	% F3-4 detected amongst patients with positive or negative result	% Patients requiring ELF Tests	% of Population with Indeterminate Results	% of F3-4 detected within entire population by strategy	AUC
2	ELF & NFS		-	Parallel	Indeterminate	100%	96%	96%	81%	100%	41%	21%	0.83
3	FIB4 & NFS		-	Parallel	Indeterminate	100%	93%	93%	82%	0%	49%	21%	0.76
-	NFS	-	-	Single	Indeterminate	82%	89%	88%	43%	0%	37%	21%	0.74
1	ELF & FIB4		-	Parallel	Indeterminate	82%	98%	95%	88%	100%	41%	33%	0.84
-	FIB4	-	-	Single	Indeterminate	80%	94%	91%	71%	0%	36%	36%	0.80
6	FIB4	ELF	-	Series	Indeterminate	80%	92%	91%	77%	36%	19%	57%	0.85
12	FIB4	ELF	-	Series	Fully Assigned	78%	90%	88%	67%	36%	0%	67%	0.80
-	ELF	-	-	Single	Indeterminate	77%	97%	92%	93%	100%	32%	62%	0.86
8	NFS	ELF	-	Series	Indeterminate	77%	90%	88%	63%	32%	16%	48%	0.80
14	NFS	ELF	-	Series	Fully Assigned	77%	87%	85%	55%	32%	0%	55%	0.75
17	FIB4	ELF	NFS	Series	Indeterminate	77%	93%	90%	73%	36%	11%	57%	0.84
7	FIB4	NFS	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
9	NFS	FIB4	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
4	ELF	FIB4	-	Series	Indeterminate	75%	94%	91%	87%	100%	19%	64%	0.88
21	FIB4	NFS	ELF	Series	Fully Assigned	75%	88%	85%	57%	20%	0%	57%	0.76
-	ELF	-	-	Single	Fully Assigned	74%	92%	88%	74%	100%	0%	74%	0.83
5	ELF	NFS	-	Series	Indeterminate	74%	95%	90%	81%	100%	16%	62%	0.86
16	ELF	FIB4	NFS	Series	Indeterminate	73%	95%	90%	82%	100%	11%	64%	0.87
18	FIB4	NFS	ELF	Series	Indeterminate	71%	91%	86%	69%	20%	6%	64%	0.80
20	FIB4	ELF	NFS	Series	Fully Assigned	70%	91%	85%	71%	36%	0%	71%	0.81
19	ELF	FIB4	NFS	Series	Fully Assigned	66%	93%	85%	69%	100%	0%	69%	0.83
11	ELF	NFS	-	Series	Fully Assigned	64%	93%	85%	79%	100%	0%	79%	0.83
10	ELF	FIB4	-	Series	Fully Assigned	63%	91%	84%	74%	100%	0%	74%	0.80
13	FIB4	NFS	-	Series	Fully Assigned	63%	88%	82%	60%	0%	0%	60%	0.74
15	NFS	FIB4	-	Series	Fully Assigned	55%	85%	79%	50%	0%	0%	50%	0.69

Table 7.6 Diagnostic strategies evaluated presented in order of descending NPV

The fully assigned approach using ELF at the 10.51 threshold is highlighted in yellow.

Algorithm	Test 1	Test 2	Test 3	Biomarker Strategy: Single, Serial or Parallel	Terminal Approach: Indeterminate or Fully Assigned	PPV%	NPV%	Diagnostic Accuracy (TP+TN)	% F3-4 detected amongst patients with positive or negative result	% Patients requiring ELF Tests	% of Population with Indeterminate Results	% of F3-4 detected within entire population by strategy	AUC
1	ELF & FIB4		-	Parallel	Indeterminate	82%	98%	95%	88%	100%	41%	33%	0.84
-	ELF	-	-	Single	Indeterminate	77%	97%	92%	93%	100%	32%	62%	0.86
2	ELF & NFS		-	Parallel	Indeterminate	100%	96%	96%	81%	100%	41%	21%	0.83
5	ELF	NFS	-	Series	Indeterminate	74%	95%	90%	81%	100%	16%	62%	0.86
16	ELF	FIB4	NFS	Series	Indeterminate	73%	95%	90%	82%	100%	11%	64%	0.87
-	FIB4	-	-	Single	Indeterminate	80%	94%	91%	71%	0%	36%	36%	0.80
4	ELF	FIB4	-	Series	Indeterminate	75%	94%	91%	87%	100%	19%	64%	0.88
3	FIB4 & NFS		-	Parallel	Indeterminate	100%	93%	93%	82%	0%	49%	21%	0.76
11	ELF	NFS	-	Series	Fully Assigned	64%	93%	85%	79%	100%	0%	79%	0.83
17	FIB4	ELF	NFS	Series	Indeterminate	77%	93%	90%	73%	36%	11%	57%	0.84
19	ELF	FIB4	NFS	Series	Fully Assigned	66%	93%	85%	69%	100%	0%	69%	0.83
-	ELF	-	-	Single	Fully Assigned	74%	92%	88%	74%	100%	0%	74%	0.83
6	FIB4	ELF	-	Series	Indeterminate	80%	92%	91%	77%	36%	19%	57%	0.85
10	ELF	FIB4	-	Series	Fully Assigned	63%	91%	84%	74%	100%	0%	74%	0.80
18	FIB4	NFS	ELF	Series	Indeterminate	71%	91%	86%	69%	20%	6%	64%	0.80
20	FIB4	ELF	NFS	Series	Fully Assigned	70%	91%	85%	71%	36%	0%	71%	0.81
7	FIB4	NFS	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
8	NFS	ELF	-	Series	Indeterminate	77%	90%	88%	63%	32%	16%	48%	0.80
9	NFS	FIB4	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
12	FIB4	ELF	-	Series	Fully Assigned	78%	90%	88%	67%	36%	0%	67%	0.80
-	NFS	-	-	Single	Indeterminate	82%	89%	88%	43%	0%	37%	21%	0.74
13	FIB4	NFS	-	Series	Fully Assigned	63%	88%	82%	60%	0%	0%	60%	0.74
21	FIB4	NFS	ELF	Series	Fully Assigned	75%	88%	85%	57%	20%	0%	57%	0.76
14	NFS	ELF	-	Series	Fully Assigned	77%	87%	85%	55%	32%	0%	55%	0.75
15	NFS	FIB4	-	Series	Fully Assigned	55%	85%	79%	50%	0%	0%	50%	0.69

Table 7.7 Diagnostic strategies evaluated presented in order of descending Diagnostic Accuracy

The fully assigned approach using ELF at the 10.51 threshold is highlighted in yellow.

Algorithm	Test 1	Test 2	Test 3	Biomarker Strategy: Single, Serial or Parallel	Terminal Approach: Indeterminate or Fully Assigned	PPV%	NPV%	Diagnostic Accuracy (TP+TN)	% F3-4 detected amongst patients with positive or negative result	% Patients requiring ELF Tests	% of Population with Indeterminate Results	% of F3-4 detected within entire population by strategy	AUC
2	ELF & NFS		-	Parallel	Indeterminate	100%	96%	96%	81%	100%	41%	21%	0.83
1	ELF & FIB4		-	Parallel	Indeterminate	82%	98%	95%	88%	100%	41%	33%	0.84
3	FIB4 & NFS		-	Parallel	Indeterminate	100%	93%	93%	82%	0%	49%	21%	0.76
-	ELF	-	-	Single	Indeterminate	77%	97%	92%	93%	100%	32%	62%	0.86
-	FIB4	-	-	Single	Indeterminate	80%	94%	91%	71%	0%	36%	36%	0.80
4	ELF	FIB4	-	Series	Indeterminate	75%	94%	91%	87%	100%	19%	64%	0.88
6	FIB4	ELF	-	Series	Indeterminate	80%	92%	91%	77%	36%	19%	57%	0.85
5	ELF	NFS	-	Series	Indeterminate	74%	95%	90%	81%	100%	16%	62%	0.86
16	ELF	FIB4	NFS	Series	Indeterminate	73%	95%	90%	82%	100%	11%	64%	0.87
17	FIB4	ELF	NFS	Series	Indeterminate	77%	93%	90%	73%	36%	11%	57%	0.84
-	ELF	-	-	Single	Fully Assigned	74%	92%	88%	74%	100%	0%	74%	0.83
-	NFS	-	-	Single	Indeterminate	82%	89%	88%	43%	0%	37%	21%	0.74
7	FIB4	NFS	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
8	NFS	ELF	-	Series	Indeterminate	77%	90%	88%	63%	32%	16%	48%	0.80
9	NFS	FIB4	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
12	FIB4	ELF	-	Series	Fully Assigned	78%	90%	88%	67%	36%	0%	67%	0.80
18	FIB4	NFS	ELF	Series	Indeterminate	71%	91%	86%	69%	20%	6%	64%	0.80
11	ELF	NFS	-	Series	Fully Assigned	64%	93%	85%	79%	100%	0%	79%	0.83
14	NFS	ELF	-	Series	Fully Assigned	77%	87%	85%	55%	32%	0%	55%	0.75
19	ELF	FIB4	NFS	Series	Fully Assigned	66%	93%	85%	69%	100%	0%	69%	0.83
20	FIB4	ELF	NFS	Series	Fully Assigned	70%	91%	85%	71%	36%	0%	71%	0.81
21	FIB4	NFS	ELF	Series	Fully Assigned	75%	88%	85%	57%	20%	0%	57%	0.76
10	ELF	FIB4	-	Series	Fully Assigned	63%	91%	84%	74%	100%	0%	74%	0.80
13	FIB4	NFS	-	Series	Fully Assigned	63%	88%	82%	60%	0%	0%	60%	0.74
15	NFS	FIB4	-	Series	Fully Assigned	55%	85%	79%	50%	0%	0%	50%	0.69

Table 7.8 Patients with either positive or negative result: diagnostic strategies presented in descending order of % of F3-4 detected

The fully assigned approach using ELF at the 10.51 threshold is highlighted in yellow.

Algorithm	Test 1	Test 2	Test 3	Biomarker Strategy: Single, Serial or Parallel	Terminal Approach: Indeterminate or Fully Assigned	PPV%	NPV%	Diagnostic Accuracy (TP+TN)	% F3-4 detected amongst patients with positive or negative result	% Patients requiring ELF Tests	% of Population with Indeterminate Results	% of F3-4 detected within entire population by strategy	AUC
-	ELF	-	-	Single	Indeterminate	77%	97%	92%	93%	100%	32%	62%	0.86
1	ELF & FIB4		-	Parallel	Indeterminate	82%	98%	95%	88%	100%	41%	33%	0.84
4	ELF	FIB4	-	Series	Indeterminate	75%	94%	91%	87%	100%	19%	64%	0.88
3	FIB4 & NFS		-	Parallel	Indeterminate	100%	93%	93%	82%	0%	49%	21%	0.76
16	ELF	FIB4	NFS	Series	Indeterminate	73%	95%	90%	82%	100%	11%	64%	0.87
2	ELF & NFS		-	Parallel	Indeterminate	100%	96%	96%	81%	100%	41%	21%	0.83
5	ELF	NFS	-	Series	Indeterminate	74%	95%	90%	81%	100%	16%	62%	0.86
11	ELF	NFS	-	Series	Fully Assigned	64%	93%	85%	79%	100%	0%	79%	0.83
6	FIB4	ELF	-	Series	Indeterminate	80%	92%	91%	77%	36%	19%	57%	0.85
-	ELF	-	-	Single	Fully Assigned	74%	92%	88%	74%	100%	0%	74%	0.83
10	ELF	FIB4	-	Series	Fully Assigned	63%	91%	84%	74%	100%	0%	74%	0.80
17	FIB4	ELF	NFS	Series	Indeterminate	77%	93%	90%	73%	36%	11%	57%	0.84
-	FIB4	-	-	Single	Indeterminate	80%	94%	91%	71%	0%	36%	36%	0.80
20	FIB4	ELF	NFS	Series	Fully Assigned	70%	91%	85%	71%	36%	0%	71%	0.81
18	FIB4	NFS	ELF	Series	Indeterminate	71%	91%	86%	69%	20%	6%	64%	0.80
19	ELF	FIB4	NFS	Series	Fully Assigned	66%	93%	85%	69%	100%	0%	69%	0.83
12	FIB4	ELF	-	Series	Fully Assigned	78%	90%	88%	67%	36%	0%	67%	0.80
8	NFS	ELF	-	Series	Indeterminate	77%	90%	88%	63%	32%	16%	48%	0.80
13	FIB4	NFS	-	Series	Fully Assigned	63%	88%	82%	60%	0%	0%	60%	0.74
21	FIB4	NFS	ELF	Series	Fully Assigned	75%	88%	85%	57%	20%	0%	57%	0.76
14	NFS	ELF	-	Series	Fully Assigned	77%	87%	85%	55%	32%	0%	55%	0.75
15	NFS	FIB4	-	Series	Fully Assigned	55%	85%	79%	50%	0%	0%	50%	0.69
-	NFS	-	-	Single	Indeterminate	82%	89%	88%	43%	0%	37%	21%	0.74
7	FIB4	NFS	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
9	NFS	FIB4	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78

Table 7.9 Diagnostic strategies presented in order of descending % F3-4 identified among entire population

The fully assigned approach using ELF at the 10.51 threshold is highlighted in yellow.

Algorithm	Test 1	Test 2	Test 3	Biomarker Strategy: Single, Serial or Parallel	Terminal Approach: Indeterminate or Fully Assigned	PPV%	NPV%	Diagnostic Accuracy (TP+TN)	% F3-4 detected amongst patients with positive or negative result	% Patients requiring ELF Tests	% of Population with Indeterminate Results	% of F3-4 detected within entire population by strategy	AUC
11	ELF	NFS	-	Series	Fully Assigned	64%	93%	85%	79%	100%	0%	79%	0.83
-	ELF	-	-	Single	Fully Assigned	74%	92%	88%	74%	100%	0%	74%	0.83
10	ELF	FIB4	-	Series	Fully Assigned	63%	91%	84%	74%	100%	0%	74%	0.80
20	FIB4	ELF	NFS	Series	Fully Assigned	70%	91%	85%	71%	36%	0%	71%	0.81
19	ELF	FIB4	NFS	Series	Fully Assigned	66%	93%	85%	69%	100%	0%	69%	0.83
12	FIB4	ELF	-	Series	Fully Assigned	78%	90%	88%	67%	36%	0%	67%	0.80
4	ELF	FIB4	-	Series	Indeterminate	75%	94%	91%	87%	100%	19%	64%	0.88
16	ELF	FIB4	NFS	Series	Indeterminate	73%	95%	90%	82%	100%	11%	64%	0.87
18	FIB4	NFS	ELF	Series	Indeterminate	71%	91%	86%	69%	20%	6%	64%	0.80
-	ELF	-	-	Single	Indeterminate	77%	97%	92%	93%	100%	32%	62%	0.86
5	ELF	NFS	-	Series	Indeterminate	74%	95%	90%	81%	100%	16%	62%	0.86
13	FIB4	NFS	-	Series	Fully Assigned	63%	88%	82%	60%	0%	0%	60%	0.74
6	FIB4	ELF	-	Series	Indeterminate	80%	92%	91%	77%	36%	19%	57%	0.85
7	FIB4	NFS	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
9	NFS	FIB4	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
17	FIB4	ELF	NFS	Series	Indeterminate	77%	93%	90%	73%	36%	11%	57%	0.84
21	FIB4	NFS	ELF	Series	Fully Assigned	75%	88%	85%	57%	20%	0%	57%	0.76
14	NFS	ELF	-	Series	Fully Assigned	77%	87%	85%	55%	32%	0%	55%	0.75
15	NFS	FIB4	-	Series	Fully Assigned	55%	85%	79%	50%	0%	0%	50%	0.69
8	NFS	ELF	-	Series	Indeterminate	77%	90%	88%	63%	32%	16%	48%	0.80
-	FIB4	-	-	Single	Indeterminate	80%	94%	91%	71%	0%	36%	36%	0.80
1	ELF & FIB4		-	Parallel	Indeterminate	82%	98%	95%	88%	100%	41%	33%	0.84
-	NFS	-	-	Single	Indeterminate	82%	89%	88%	43%	0%	37%	21%	0.74
2	ELF & NFS		-	Parallel	Indeterminate	100%	96%	96%	81%	100%	41%	21%	0.83
3	FIB4 & NFS		-	Parallel	Indeterminate	100%	93%	93%	82%	0%	49%	21%	0.76

Table 7.10 Diagnostic strategies that mandate ELF testing for all patients presented in order of descending diagnostic accuracy

The fully assigned approach using ELF at the 10.51 threshold is highlighted in yellow

Algorithm	Test 1	Test 2	Test 3	Biomarker Strategy: Single, Serial or Parallel	Terminal Approach: Indeterminate or Fully Assigned	PPV%	NPV%	Diagnostic Accuracy (TP+TN)	% F3-4 detected amongst patients with positive or negative result	% Patients requiring ELF Tests	% of Population with Indeterminate Results	% of F3-4 detected within entire population by strategy	AUC
2	ELF & NFS		-	Parallel	Indeterminate	100%	96%	96%	81%	100%	41%	21%	0.83
1	ELF & FIB4		-	Parallel	Indeterminate	82%	98%	95%	88%	100%	41%	33%	0.84
-	ELF	-	-	Single	Indeterminate	77%	97%	92%	93%	100%	32%	62%	0.86
4	ELF	FIB4	-	Series	Indeterminate	75%	94%	91%	87%	100%	19%	64%	0.88
5	ELF	NFS	-	Series	Indeterminate	74%	95%	90%	81%	100%	16%	62%	0.86
16	ELF	FIB4	NFS	Series	Indeterminate	73%	95%	90%	82%	100%	11%	64%	0.87
-	ELF	-	-	Single	Fully Assigned	74%	92%	88%	74%	100%	0%	74%	0.83
11	ELF	NFS	-	Series	Fully Assigned	64%	93%	85%	79%	100%	0%	79%	0.83
19	ELF	FIB4	NFS	Series	Fully Assigned	66%	93%	85%	69%	100%	0%	69%	0.83
10	ELF	FIB4	-	Series	Fully Assigned	63%	91%	84%	74%	100%	0%	74%	0.80

Table 7.11 Diagnostic strategies that require ELF for between 20-36% of patients presented in order of descending diagnostic accuracy

Algorithm	Test 1	Test 2	Test 3	Biomarker Strategy: Single, Serial or Parallel	Terminal Approach: Indeterminate or Fully Assigned	PPV%	NPV%	Diagnostic Accuracy (TP+TN)	% F3-4 detected amongst patients with positive or negative result	% Patients requiring ELF Tests	% of Population with Indeterminate Results	% of F3-4 detected within entire population by strategy	AUC
6	FIB4	ELF	-	Series	Indeterminate	80%	92%	91%	77%	36%	19%	57%	0.85
17	FIB4	ELF	NFS	Series	Indeterminate	77%	93%	90%	73%	36%	11%	57%	0.84
12	FIB4	ELF	-	Series	Fully Assigned	78%	90%	88%	67%	36%	0%	67%	0.80
8	NFS	ELF	-	Series	Indeterminate	77%	90%	88%	63%	32%	16%	48%	0.80
18	FIB4	NFS	ELF	Series	Indeterminate	71%	91%	86%	69%	20%	6%	64%	0.80
20	FIB4	ELF	NFS	Series	Fully Assigned	70%	91%	85%	71%	36%	0%	71%	0.81
14	NFS	ELF	-	Series	Fully Assigned	77%	87%	85%	55%	32%	0%	55%	0.75
21	FIB4	NFS	ELF	Series	Fully Assigned	75%	88%	85%	57%	20%	0%	57%	0.76

Table 7.12 Diagnostic strategies that do not require ELF testing presented in order of descending diagnostic accuracy

Algorithm	Test 1	Test 2	Test 3	Biomarker Strategy: Single, Serial or Parallel	Terminal Approach: Indeterminate or Fully Assigned	PPV%	NPV%	Diagnostic Accuracy (TP+TN)	% F3-4 detected amongst patients with positive or negative result	% Patients requiring ELF Tests	% of Population with Indeterminate Results	% of F3-4 detected within entire population by strategy	AUC
3	FIB4 & NFS		-	Parallel	Indeterminate	100%	93%	93%	82%	0%	49%	21%	0.76
-	FIB4	-	-	Single	Indeterminate	80%	94%	91%	71%	0%	36%	36%	0.80
-	NFS	-	-	Single	Indeterminate	82%	89%	88%	43%	0%	37%	21%	0.74
7	FIB4	NFS	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
9	NFS	FIB4	-	Series	Indeterminate	76%	90%	88%	38%	0%	20%	57%	0.78
13	FIB4	NFS	-	Series	Fully Assigned	63%	88%	82%	60%	0%	0%	60%	0.74
15	NFS	FIB4	-	Series	Fully Assigned	55%	85%	79%	50%	0%	0%	50%	0.69

Table 7.13 Fully Assigned Strategies presented in order of descending diagnostic accuracy.

The fully assigned approach using ELF at the 10.51 threshold is highlighted in yellow.

Algorithm	Test 1	Test 2	Test 3	Biomarker Strategy: Single, Serial or Parallel	Terminal Approach: Indeterminate or Fully Assigned	PPV%	NPV%	Diagnostic Accuracy (TP+TN)	% F3-4 detected amongst patients with positive or negative result	% Patients requiring ELF Tests	% of Population with Indeterminate Results	% of F3-4 detected within entire population by strategy	AUC
-	ELF	-	-	Single	Fully Assigned	74%	92%	88%	74%	100%	0%	74%	0.83
12	FIB4	ELF	-	Series	Fully Assigned	78%	90%	88%	67%	36%	0%	67%	0.80
11	ELF	NFS	-	Series	Fully Assigned	64%	93%	85%	79%	100%	0%	79%	0.83
14	NFS	ELF	-	Series	Fully Assigned	77%	87%	85%	55%	32%	0%	55%	0.75
19	ELF	FIB4	NFS	Series	Fully Assigned	66%	93%	85%	69%	100%	0%	69%	0.83
20	FIB4	ELF	NFS	Series	Fully Assigned	70%	91%	85%	71%	36%	0%	71%	0.81
21	FIB4	NFS	ELF	Series	Fully Assigned	75%	88%	85%	57%	20%	0%	57%	0.76
10	ELF	FIB4	-	Series	Fully Assigned	63%	91%	84%	74%	100%	0%	74%	0.80
13	FIB4	NFS	-	Series	Fully Assigned	63%	88%	82%	60%	0%	0%	60%	0.74
15	NFS	FIB4	-	Series	Fully Assigned	55%	85%	79%	50%	0%	0%	50%	0.69

DISCUSSION

In this study, as compared to the use of the individual tests alone, direct and indirect biomarkers of liver fibrosis have been successfully combined to enhance the detection of advanced fibrosis in NAFLD. Moreover, these novel combination strategies have been compared to the strategy recommended by NICE (the application ELF at a threshold of 10.51 to distinguish advanced fibrosis from lesser degrees of fibrosis) and those commonly employed in clinical practice (indirect biomarkers applied using a combined cut off).

Of the 3 individual tests that were explored in this study, ELF demonstrated the best performance. ELF at a threshold of 10.51 was also the best performing fully assigned test strategy, detecting 74% of patients with advanced fibrosis in contrast to the indirect tests which detected only 55-60% of patients with advanced fibrosis. The application of a combined cut-off for ELF in a single test strategy improved both disease detection to 93% amongst patients with a positive or negative result in contrast to FIB4 (71%) and NAFLD fibrosis score (43%) but left 32% of indeterminate cases to be resolved by further investigations.

Of note, these findings were predicted by applying Bayes' theorem to the published performance of these biomarkers at their diagnostic thresholds. The application of a parallel strategy increased both the sensitivity and specificity of the diagnostic pathway resulting in negative and positive predictive values of almost 100%. However, the application of a parallel strategy was associated with an increase in the number of cases classified as indeterminate, thus reducing overall AUROC compared with single test evaluation. By contrast, compared with a single test applied with an indeterminate approach, the application of a serial strategies can significantly reduce the number of cases classified as indeterminate whilst preserving diagnostic accuracy, resulting in

an increase in AUROC. In addition, considering ELF containing algorithms, the re-testing of indeterminate cases in serial strategies allowed for up to 80% reduction in the need for ELF testing in this study population.

Given these findings, in clinical practice a positive or negative result from a ‘parallel algorithm’ would certainly provide strong evidence to ‘rule in’ or ‘rule out’ advanced fibrosis respectively whereas an indeterminate result arising from a ‘parallel algorithm’ could then be reassigned by the application of ELF. However, given the cost differential between ELF and the indirect tests, if one was attempting to reduce the number of ELF tests applied, an indeterminate result arising from a parallel strategy could be re-tested using a serial strategy containing ELF itself. Indeed, as ELF in this context is acting as a somewhat of a final arbiter, one could substitute ELF for another non-invasive test with excellent performance such as transient or magnetic resonance elastography.

As such, how these diagnostic strategies are employed in clinical practice is very much dependent on the availability and cost of ELF testing. In a health care system in which ELF testing is freely available, a diagnostic pathway in which a patient is first tested using a parallel strategy by the application of an indirect biomarker and ELF itself. Whereas a positive or negative result would effectively rule in or rule out advanced fibrosis, an indeterminate result could be reassigned by the application of ELF at a threshold of 10.51.

If, however, there is a need to reduce the number of ELF tests that are used, then a diagnostic pathway in which patients are first tested using a parallel strategy composed of 2 indirect tests would seem logical. As before, a positive or negative result will provide strong evidence to rule in or rule out advanced fibrosis. In this context however, an indeterminate result could be tested using a serial strategy (fully

assigned terminal approach) using FIB4 as the first test and ELF as the second test. In this study, this approach produced a 64% reduction in the need for ELF testing despite still detecting 67% of patients with advanced fibrosis and performing with a diagnostic accuracy of 88%. Alternatively, a 3 biomarker serial strategy that terminates with ELF in a fully assigned approach could be applied; this strategy reduced the need for ELF testing in the study cohort by 80% whilst still detecting 57% of patients with advanced fibrosis and demonstrating a diagnostic accuracy of 88%. However, when working within a diagnostic pathway in which ELF tests are unavailable, a serial fully assigned strategy of FIB4 followed by NFS (diagnostic accuracy of 82%) could be employed to retest cases classified as indeterminate by a 2 indirect biomarker parallel strategy. Regardless, prior to the implementation of these proposals a formal health economic analysis should be conducted. This should also take into account factors such as failure rates (elastography), debate over diagnostic thresholds and availability. These findings also need to be considered in the context of the inherent flaws of the reference test used in this study, liver biopsy. Even under optimal conditions, liver biopsy is far from a perfect reference standard. Aside from the hazards associated with the procedure of liver biopsy itself, sampling variability may under or over stage fibrosis in as many of 20% of liver biopsies. [28] Furthermore, liver biopsy specimens may be fragmented or insufficient size (e.g. 15mm or 5 portal tracts) to enable a confident histopathologic fibrosis stage to be determined. It therefore has been argued that the maximum diagnostic accuracy of liver biopsy under optimal circumstances could be no more than 90%.[349] Thus any non-invasive approach which is able to deliver a diagnostic accuracy of approximately 90% while avoiding the associated hazards of invasive testing is certainly more appealing. The findings of this study suggest that an indeterminate approach which offers higher diagnostic accuracy as

compared with a fully assigned approach performs at a level which is comparable to liver biopsy with 90% accuracy and a failure rate of approximately 10%. The indeterminate approach serial strategies with their 10% failure rate also compare favourably with imaging modalities such as transient elastography which itself also suffers from a significant failure rate even when using a probe specifically designed for obese patients (failure rate up to 16%).[261]

Nevertheless, it is important that any algorithm which will be implemented in clinical practice should be easy for clinicians to follow. In the absence of reflex testing within a laboratory or an automated algorithm calculator, clinicians may find these algorithms cumbersome. For this reason, the performance of a variety of biomarker combinations ranging from only two biomarkers with only one additional triage stage, to more complex approaches with three biomarkers and three additional triage stages have been presented. Whereas combinations of direct and indirect biomarkers have been combined retrospectively by logistic regression in validation studies this leads to the generation of yet more diagnostic thresholds which need to be calculated and validated.

The DANA score provides key insights into the cases classified as indeterminate and emphasises their importance. Rather than defining a group with a low prevalence of disease, it is evident that the group classified as indeterminate may contain a significant proportion of patients with advanced fibrosis. For example, half of all the cases with advanced fibrosis are classified by the indirect markers as indeterminate. Furthermore, an evaluation of the spectrum of disease and the degree of spectrum bias (DANA) in the indeterminate group reveals that despite having a higher overall prevalence of advanced fibrosis, there is central clustering of fibrosis stages. The indeterminate group therefore represents a population that is difficult to characterize

containing a significant number of cases with advanced fibrosis. This highlights the importance of including tests that exhibit both good performance and conditional independence when developing diagnostic algorithms with combinations of biomarkers exhibiting greater degrees of conditional independence or scanning modalities such as transient elastography appearing logical.

The results in this chapter also need to be interpreted in the context of the earlier findings relating to the use of PIIINP in the detection of both NASH and advanced fibrosis. Given the level of performance of PIIINP for the detection of NASH, following the application of PIIINP, the presence or absence of advanced fibrosis could be determined by one of the diagnostic approaches outlined above the contents of which are dependent on the availability of ELF testing.

In this study the performance of many of the proposed diagnostic strategies using Bayes theorem and biomarker threshold published performance were estimated and validated. Whereas these results will need to be further validated in larger independent series, the analysis has identified that in addition to sample size, both spectrum bias and test dependence are the principle reasons why the estimated and actual performances of these biomarkers both alone and as algorithms may vary. Regardless, the similarities between the estimated and actual performances are striking suggesting that similar approaches can be used to design diagnostic pathways for NAFLD in both primary and secondary care.

Overall, the findings support the endorsement of ELF at the 10.51 threshold for the detection of advanced fibrosis in NAFLD by NICE. Of the 25 strategies proposed, ELF at this threshold offers the overall best performance offering a diagnostic accuracy of 88%, identifying 74% of the patients with advanced fibrosis and without any indeterminate results. Whilst there is a desire to employ indirect tests due to their lower

cost, given the size of the NAFLD epidemic that is being faced, an economy of scale should dramatically drive down the costs of ELF testing.

In summary, after estimating performance using Bayes theorem, diagnostic strategies have successfully been developed to enhance the diagnostic performance of biomarkers of advanced fibrosis in a cohort of patients with NAFLD in secondary care. Firstly, as compared to single test evaluation including ELF at a threshold of 10.51, the application of a parallel strategy was associated with higher positive and negative predictive values but with a greater proportion of patients classified as indeterminate. Thus, on an individual patient basis, the positive and negative results arising after the application of a parallel algorithm provide compelling evidence to either rule in or ‘rule out advanced fibrosis. Secondly, the indeterminate results arising from a parallel strategy algorithm could be retested by ELF itself or a serial strategy algorithm that can reduce the need for ELF testing whilst maintaining diagnostic accuracy and increasing AUROC. Moreover, diagnostic approaches have been proposed in health care settings in which the ELF test is not available. These strategies can be used together in clinical practice to effectively stratify patients with NALFD into those with and without advanced fibrosis.

CHAPTER 8

NON-INVASIVE ASSESSMENT OF HEPATIC FIBROSIS IN PRIOR NON-RESPONDERS TO HEPATITIS C VIRUS TREATMENT: A COMPARISON OF TEN BIOMARKERS OF LIVER FIBROSIS

ABSTRACT

Background: advancing liver fibrosis is regarded as the most important factor when stratifying patients with CHC for treatment.

Aims: (1) to compare the performance of 10 biomarkers of fibrosis in patients with CHC and treatment failure; (2) to assess the impact on biomarker performance of using 2 different assays of HA.

Methods: for 80 patients, liver histology (Metavir) was compared to biomarker scores using sera obtained within 6 months of liver biopsy (indirect biomarkers -AST:ALT ratio, APRI, Forns index, FIB-4, Fibrometer-3G, direct biomarkers - ELF, Fibrospect-II, Hyularonic acid-HA, and hybrid biomarkers of fibrosis -Fibrometer-2G, Hepascore). Direct and hybrid marker scores were calculated using 2 validated assays for HA (ELISA-Siemens, radiometric-Pharmacia).

Results: using the ELISA assay for HA to calculate the direct and hybrid panels, all 10 of the biomarkers exhibited comparable overall discriminatory performance (ordROC 0.92-0.94, *p-value>0.05*) except AST:ALT ratio and APRI (ordROC 0.86-0.88, *p-value<0.05*). For the detection of moderate (F2-4) and advanced (F3-4) fibrosis, the AUROC of Fibrometer-2G were significantly higher than AST:ALT ratio and APRI but not of any of the other biomarkers. Excellent correlation was observed between the two HA assays (R^2=0.909) with the ELISA assay exhibiting superior diagnostic performance (ordROC 0.92 Vs.0.88, *p-value=0.003*). Importantly, the performance of direct and hybrid biomarkers at their diagnostic thresholds was heavily influenced by the choice of HA assay.

Conclusion: Whilst many biomarkers exhibited good diagnostic performance for the detection of advancing fibrosis these results indicate that diagnostic performance may be significantly affected by the selection of individual component assays.

Acknowledgements: The PROFI-C study was endorsed by the German Hepatitis Network (HepNet). Prof. Otto Dietze co-scored the liver biopsy specimens. The Fibrometer scores were calculated by Prof. Paul Calès.

INTRODUCTION

Amongst patients with CHC and prior treatment failure, the development of advancing fibrosis is regarded as the most important factor when stratifying patients for retreatment.[350] Patients with CHC and prior treatment failure with advanced or rapidly advancing fibrosis are candidates for early retreatment as they are at immediate risk of the complications of CLD including portal hypertension, liver failure and hepatocellular carcinoma. The reference standard for assessing hepatic fibrosis in CHC remains the histological staging of a liver biopsy specimen. The limitations of liver biopsy include sampling error[181], inter- and intra-observer variability[351], and procedural complications.[28] As a result, non-invasive methods have been developed for the cross-sectional staging of liver fibrosis. Hitherto, serum markers of liver fibrosis have been widely studied for their ability to discriminate between different stages of liver fibrosis.[3] Serum markers can be categorised in several ways including into direct serum markers, indirect serum markers and hybrid serum panels (combinations of direct and indirect serum markers). Further distinctions can be made into those markers deemed specific for single disease such as CHC (for example Fibrometer 2G) or those markers which have been validated for use in a variety of liver diseases such as Hepascore or ELF.

Hitherto, non-invasive serum markers have been subjected to numerous derivation and validation studies which have collectively suggested that many of these tests have reliable performance for the detection of severe fibrosis in CHC.[336] Due to factors such as spectrum bias[201], any comparison of the performance of biomarkers derived from patient populations with differing characteristics can be subject to type I and type II error. Furthermore, for most biomarker panels, the algorithms that combine component tests are published and widely available. As the use of biomarkers to

assess fibrosis moves from a research application to become part of clinical practice it is essential to understand whether or not the use of alternative component assays will impact on the discriminatory performance of these tests both overall and at their diagnostic thresholds.

This study had the following aims: (1) to validate and compare the performance of biomarkers of fibrosis in patients with CHC and prior treatment failure (2) to determine whether the diagnostic performance of direct and hybrid panels that contain HA as a constituent component are influenced by the choice of HA assay.

METHODS

Patient Population

Patients in this study were enrolled in the PROFI-C trial which was an investigator-initiated, prospective, randomized trial involving 18 centers in Germany and Austria investigating the effect of high dose silymarin plus pegylated interferon alpha 2b in non-responders or relapsers to standard treatment for CHC. Ethical approval was granted by the local ethics committees of the participating centers with initial ethical approval granted by Clinical Ethics Committee at the University Hospital Erlangen. One hundred and eight patients participated in the PROFI-C study.[352] Participants were male and female patients aged between 18 and 65 years with evidence of CHC (COBAS Amplicor HCV Monitor, Roche Molecular Diagnostics, Mannheim, Germany) after failure of therapy with either interferon or pegylated interferon and ribavirin. Patients were also required to have histologically proven chronic hepatitis on a liver biopsy specimen (at least 8 identifiable portal tracts) within 6 months prior to entry into the study. Written consent was obtained from all patients before admission to the study.

Exclusion criteria included acute hepatitis, therapy with steroids or immunosuppressive drugs in the previous three months, Child-Pugh stage B or C cirrhosis, thrombocytopenia ($<100 \times 10^9$/L), leucopenia ($<3 \times 10^9$/L), other chronic liver diseases, autoimmune diseases, HIV infection, alcohol abuse (defined as the consumption of >40g per day in males and >20g per day in females), active drug abuse, pregnancy or psychiatric diseases including depression.

Histological assessment

Liver biopsies were fixed in formalin and embedded in paraffin. Hematoxylin-Eosin staining was used for grading of inflammation and the Chromotrope-aniline blue staining for staging the amount of liver fibrosis. [353, 354] All specimens were graded and staged according to the 5 stage Metavir Score. [188] Histological assessment was performed by 2 independent pathologists (D.N. and O.D.) who were blinded to the clinical data and randomization status of the patients in the study and one another's scores. Interobserver variability was determined by the Kappa statistic (Kappa=0.624). All liver biopsy specimens that were discordantly staged were re-reviewed by both pathologists with a final score determined after further discussion.

Sample Collection and Calculation of Serum Marker Panels for CHC

Only patients recruited into PROFI-C who underwent liver biopsy prior to therapy and had stored sera were evaluated in this study (n=80). Patient samples were tested for hematological and biochemical parameters. Serum samples were stored at -70ºC prior to transfer to the central laboratory, where serum samples were analyzed for levels of HA, TIMP-1 and PIIINP using the proprietary assays developed for the ELF test by Siemens Healthcare Diagnostics Inc (Tarrytown, New York, USA). The assays are magnetic particle separation immunoassays and were performed on the ADVIA Centaur® immunoassay system (Siemens Medical Solutions Diagnostics Inc, Tarrytown, New York, USA). Hepatitis C virus RNA was quantified, and genotyping performed on all samples. Hepatitis C virus (HCV) RNA was quantified using an in-house HCV RNA real time RT-qPCR[355] using the QIAamp96 Virus nucleic acid purification procedure on the BioRobotMDx (Qiagen, Hilden, Germany) and the ABI Prism 7500 real-time PCR with Qiagen QuantiTect probe RT-PCR reagents. The

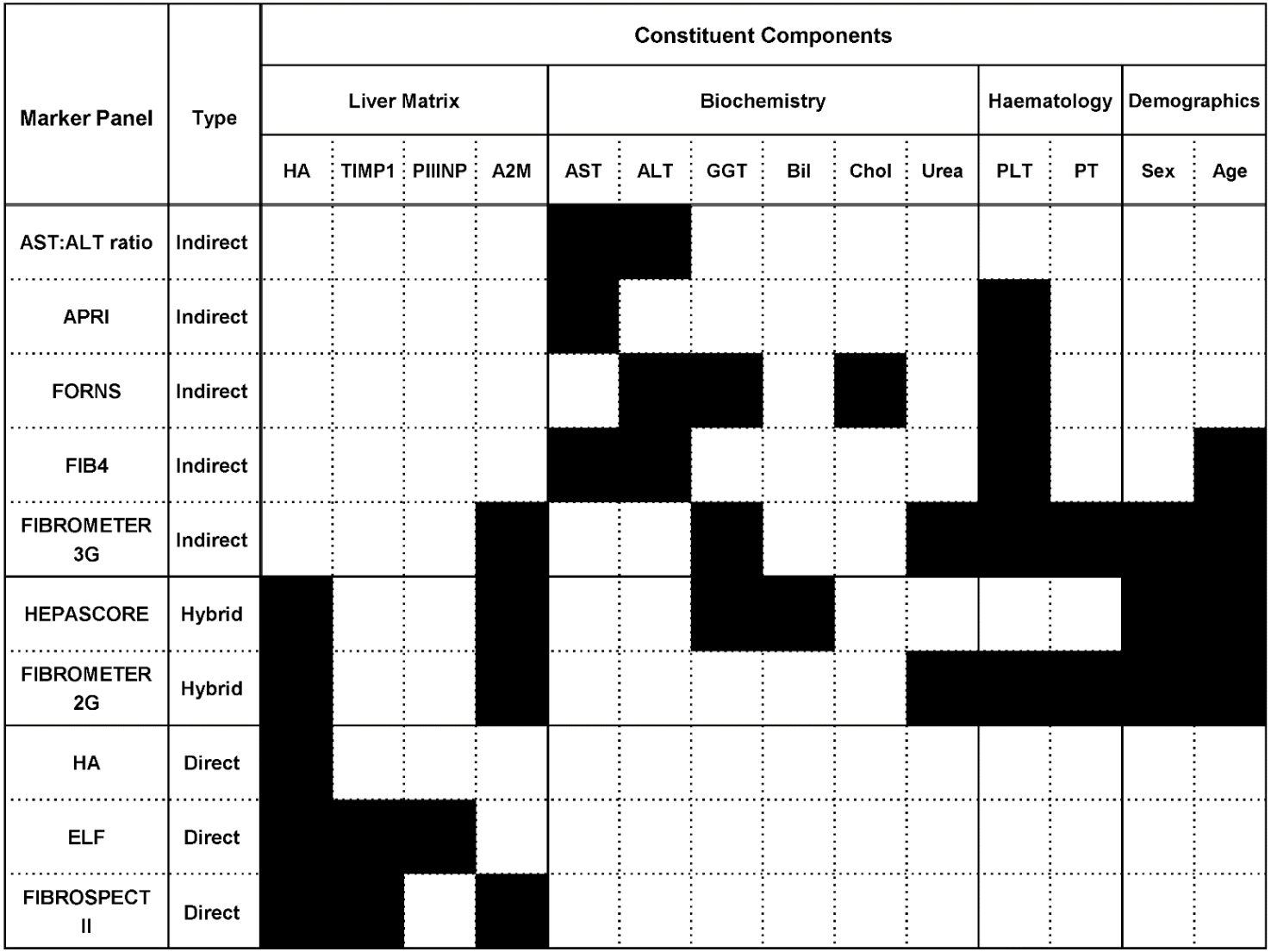
assay uses brome mosaic virus RNA as an internal control, introduced at the extraction stage.

HCV genotyping was performed by amplifying and sequencing a region of the 5'NCR. The sequence was analysed to compare probe binding sites of the LiPa method [356] and by finding the restriction sites.[357] These two virtual methods were compared to give the HCV genotype result.

Markers Evaluated in this study

Ten markers of fibrosis (figure 8.1) were evaluated in this study which can be characterized into indirect (AST to ALT ratio[358], AST to Platelet Ratio Index (APRI)[359], Forns Index[237], FIB-4[240], Fibrometer 3G[360]), hybrid (Hepascore[224], Fibrometer 2G[361]) and direct markers: (ELF [221], Fibrospect II[362], HA[363]). A full description of how these marker scores were calculated is described in the appendix.

Figure 8.1 The constituent components used to calculate the indirect, direct and hybrid markers used in the study



Exploring the Impact of using an alternative assay for HA on the performance of hybrid and direct markers

Serum levels of HA were also measured locally in the PROFI-C study with a radiometric assay (Pharmacia AB, Uppsala, Sweden) which has been validated for the detection of fibrosis in CHC both as a single marker[363] and as the constituent component of biomarker panels.[364] This allowed an exploration of whether the performance of both the hybrid and the direct markers, as well as the HA tests individually would be affected by the choice of assay.

Statistical Analyses

Statistical analyses were performed using SPSS for Windows (version 20, SPSS Inc, Chicago, IL) and R for Windows (version 2.15.1, The R Foundation for Statistical Computing). Patient demographic and clinical laboratory characteristics were descriptively summarized and reported as mean ± standard deviation (SD) and range. All tests were two-sided and statistical significance assessed at the 0.05 threshold. The diagnostic performance of the biomarkers as compared to liver biopsy was assessed using ROC curves. The AUROC and 95% confidence intervals of AUROC were calculated. Good performance for a test within the studied cohort was defined as an AUROC > 0.8.[199] The Obuchowski[202] method of correcting for spectrum effect was applied in a similar fashion to previously published literature

The Obuchowski measure (ordROC) gives an average of the N(N-1)/2 AUROC pairwise comparisons between N categories of gold standard outcome. Thus using the Metavir scale with its N (=5) categories of fibrosis staging (F0-4) there are 10 pair wise comparisons between 2 of the N categories. We defined a penalty function proportional to the difference in Metavir units between fibrosis stages. The penalty

function was 0.25, 0.5, 0.75, and 1 when the difference between Metavir stages was 1, 2, 3 and 4 stages respectively. As the severity of histological liver fibrosis in patients with CHC with prior treatment failure has not been well characterised, the Obuchowski measure presented in this study has not been weighted according to the prevalence of fibrosis stages in a reference population.

AUROC were compared using the method of Delong.[331] Sensitivity, specificity and predictive values were calculated at thresholds derived from ROC curves. The thresholds evaluated included those previously proposed for the respective biomarkers markers for the detection of moderate and advanced fibrosis which were compared with both the Q-point (where sensitivity and specificity are equal) and the Youden cut-off (highest sum of sensitivity and specificity minus 1). Logistic regression was used to determine whether combinations of biomarker panels were more effective than individual panels in discriminating between patients with and without moderate fibrosis, advanced fibrosis and cirrhosis.

RESULTS

Patient characteristics and baseline histology

The baseline characteristics of the 80 patients included in this study are displayed in table 8.1. The majority of the patient population was male with a mean age of 48.5 years. The predominant HCV genotype was genotype 1.

Comparison of all 10 markers (direct and hybrid Panels calculated with Siemens HA assay) (Table 8.2)

Ability to Discriminate Moderate Fibrosis (F0-1 Vs. F2-4)

Whereas all the direct and hybrid panels and HA as an individual assay were able to discriminate between patients with and without moderate fibrosis with an AUROC of >0.8, the only indirect panel that achieved this level of performance was Fibrometer 3G (AUROC 0.86). Of all the markers tested, Fibrometer 2G generated the highest AUROC (0.88) for the detection of moderate fibrosis which was significantly higher *(p=value<0.05)* than all of the indirect panels other than Fibrometer 3G, or any of the other hybrid and direct markers *(p-value>0.05) (table 8.6).*

Ability to Discriminate Advanced Fibrosis (F0-2 Vs. F3-4)

Other than Fibrospect II, all the direct and hybrid panels and HA alone achieved AUROC of >0.8 in their ability to discriminate between patients with and without advanced fibrosis; the only indirect panels that did not achieve AUROC > 0.8 was AST:ALT ratio (AUROC 0.65) and APRI (AUROC 0.71). For the detection of advanced fibrosis, Fibrometer 2G generated the highest AUROC (0.84) which was significantly higher *(p-value<0.05)* than AST:ALT ratio and APRI but none of the other markers tested *(p-value>0.05)* (table 8.2 and table 8.7).

Table 8.1 Baseline Demographics of the study cohort

Variable			n=80
Demographics	Age		48.9 ± 9.8
	Male n,%		40 (57%)
	BMI		24.3 ± 4.2
Virology	Genotype	1 (n,%)	67 (84%)
		2 or 3 (n,%)	7 (9%)
		4 (n,%)	6 (7%)
	Log Viral Load (IU/ml) (median, range)		6.21 (3.84 - 8.20)
Haematology	Hb (g/dl)		15.2 ± 1.4
	WCC (10^9/L)		6.4 ± 1.7
	PLT (g/L)		206 ± 61
	PI (%)		95.9 ± 9.6
Biochemistry	ALT (IU/L)		117 ± 97
	AST (IU/L)		81.6 ± 74.5
	GGT (IU/L)		102.5 ± 92.1
	A2M (g/L)		4.28 ± 1.21
	Bilirubin (µmol/L)		12.7 ± 7.5
	Albumin (g/L)		47.7 ± 7.6
	AFP (ng/ml)		8.2 ± 11.7
Fibrosis Stage	Metavir F0 (n,%)		7 (9%)
	Metavir F1 (n,%)		22 (28%)
	Metavir F2 (n,%)		19 (24%)
	Metavir F3 (n,%)		25 (31%)
	Metavir F4 (n,%)		7 (9%)

Values are presented as mean ± standard deviation unless otherwise specified

Table 8.2 Performance of the 10 biomarkers with respect to discriminating moderate fibrosis (F2-F4), advanced fibrosis (F3-F4) and cirrhosis (F4) and the overall diagnostic accuracy

All the direct and hybrid markers have been calculated using the Siemens assay for HA.

Marker	F0-1 (n=30) Vs. F2-4 (n=50)				F0-2 (n=48) Vs. F3-4 (n=32)				F0-3 (n=73) Vs. F4 (n=7)				Unweighted Obuchowski Measure	
	AUROC	95% CI	P-value	Std. Error	AUROC	95% CI	P-value	Std. Error	AUROC	95% CI	P-value	Std. Error	ordAUROC	Std. Error
AST:ALT ratio	0.61	0.48-0.73	0.116	0.06	0.65	0.52-0.78	0.027	0.06	0.76	0.55-0.97	0.024	0.11	0.86	0.02
APRI	0.71	0.59-0.82	0.002	0.06	0.71	0.60-0.83	0.001	0.06	0.78	0.65-0.91	0.015	0.07	0.88	0.02
FORNS	0.78	0.67-0.88	<0.001	0.05	0.82	0.72-0.91	<0.001	0.05	0.92	0.86-0.98	<0.001	0.03	0.92	0.01
FIB4	0.77	0.67-0.88	<0.001	0.05	0.81	0.71-0.90	<0.001	0.05	0.90	0.82-0.98	<0.001	0.04	0.92	0.01
FIBROMETER 3G	0.86	0.78-0.94	<0.001	0.04	0.81	0.72-0.90	<0.001	0.05	0.86	0.77-0.98	0.002	0.05	0.94	0.01
HA (Siemens)	0.80	0.71-0.90	<0.001	0.05	0.80	0.70-0.90	<0.001	0.05	0.88	0.79-0.98	0.001	0.05	0.92	0.01
HEPASCORE	0.85	0.76-0.93	<0.001	0.04	0.83	0.74-0.92	<0.001	0.05	0.86	0.70-1.00	0.002	0.08	0.93	0.01
FIBROMETER 2G	0.88	0.80-0.95	<0.001	0.04	0.84	0.75-0.93	<0.001	0.04	0.88	0.77-1.00	0.001	0.05	0.94	0.01
ELF	0.84	0.73-0.92	<0.001	0.04	0.82	0.72-0.92	<0.001	0.05	0.89	0.79-1.00	0.001	0.10	0.93	0.02
FIBROSPECT II	0.84	0.76-0.93	<0.001	0.05	0.79	0.70-0.89	<0.001	0.05	0.83	0.62-1.00	0.004	0.11	0.92	0.01

Table 8.3 Performance of the direct and hybrid biomarkers with respect to discriminating moderate fibrosis (F2-F4), advanced fibrosis (F3-F4), cirrhosis (F4) and the overall diagnostic accuracy (unweighted Obuchowski measure) when calculated using 2 HA assays

Marker	HA assay used	F0-1 (n=30) Vs. F2-4 (n=50)				F0-2 (n=30) Vs. F3-4 (n=50)				F0-3 (n=73) Vs. F4 (n=7)				Unweighted Obuchowski Measure	
		AUROC	95% CI	P-value	Std. Error	AUROC	95% CI	P-value	Std. Error	AUROC	95% CI	P-value	Std. Error	ordAUROC	Std. Error
HA	Siemens	0.80	0.71-0.90	<0.001	0.05	0.80	0.70-0.90	0.001	0.05	0.88	0.79-0.98	0.001	0.05	0.92	0.01
	Pharmacia	0.69	0.57-0.80	0.006	0.06	0.72	0.56-0.84	<0.001	0.06	0.85	0.69-1.00	0.002	0.08	0.88	0.01
HEPASCORE	Siemens	0.85	0.76-0.93	<0.001	0.04	0.83	0.74-0.92	<0.001	0.04	0.86	0.70-1.00	0.002	0.08	0.93	0.01
	Pharmacia	0.81	0.72-0.91	<0.001	0.05	0.79	0.69-0.90	<0.001	0.05	0.84	0.65-1.00	0.003	0.10	0.92	0.02
FIBROMETER 2G	Siemens	0.88	0.80-0.95	<0.001	0.04	0.83	0.74-0.91	<0.001	0.04	0.88	0.77-1.00	0.001	0.05	0.94	0.01
	Pharmacia	0.87	0.79-0.95	<0.001	0.04	0.81	0.72-0.91	<0.001	0.04	0.86	0.75-0.97	0.002	0.05	0.94	0.01
ELF	Siemens	0.84	0.73-0.92	<0.001	0.04	0.82	0.72-0.92	<0.001	0.04	0.89	0.79-1.00	0.001	0.10	0.93	0.01
	Pharmacia	0.79	0.69-0.89	<0.001	0.05	0.79	0.69-0.90	<0.001	0.05	0.87	0.73-1.00	0.001	0.05	0.91	0.01
FIBROSPECT II	Siemens	0.84	0.76-0.93	<0.001	0.05	0.79	0.70-0.89	<0.001	0.05	0.83	0.62-1.00	0.004	0.11	0.92	0.02
	Pharmacia	0.81	0.71-0.91	<0.001	0.05	0.66	0.55-0.78	0.013	0.05	0.70	0.54-0.85	0.090	0.08	0.90	0.02

Table 8.4 Performance of the indirect, direct and hybrid biomarkers evaluated in this study using the thresholds described in their original publications in detection of moderate fibrosis (F0-1 Vs F2-4, n=30 Vs n=50, prevalence 63%).

Marker	HA Assay used	Original publication test threshold and performance			Performance of original threshold in current Study								Current Study Thresholds	
		Threshold	Sens	Spec	Sens	Spec	number –ve Vs. +ve	PPV	NPV	LR+	LR-	DOR	Q-point	Youden
AST:ALT	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.69	0.77
APRI	-	0.5	83.2	54	52	80	46 Vs. 34	82%	50%	2.6	0.60	4.3	0.38	0.38
FORNS	-	5	88	71	78	57	28 Vs. 52	76%	60%	1.8	0.39	4.7	5.4	4.5
FIB4	-	1.0	69.4	58.4	90	30	14 Vs. 66	69%	64%	1.3	0.30	4.4	1.3	1.4
FIBROMETER 3G	-	0.440	81.3	74.1	100	30	9 Vs. 71	71%	100%	1.4	0	∞	0.76	0.59
HA	Siemens	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	34	52
	Pharmacia				N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	53	53
HEPASCORE	Siemens	0.5	63	89	82	63	27 Vs. 53	79%	67%	2.2	0.29	7.8	0.62	0.58
	Pharmacia				90	43	17 Vs. 63	73%	72%	1.6	0.23	6.8	0.67	0.67
FIBROMETER 2G	Siemens	0.419	80	76	100	30	8 Vs. 72	71%	100%	1.4	0	∞	0.70	0.69
	Pharmacia				100	27	8 Vs. 72	70%	100%	1.4	0	∞	0.71	0.71
ELF	Siemens	9.13	73	64	70	83	38 Vs. 42	88%	62%	4.1	0.36	10.6	8.99	9.20
	Pharmacia				100	0	0 Vs. 80	63%	37%	1.0	1.0	N/A	9.88	10.02
FIBROSPECT II	Siemens	0.36	77	73	66	82	43 Vs. 37	86%	59%	3.9	0.41	9.5	0.33	0.31
	Pharmacia				62	80	41 Vs. 39	84%	55%	3.1	0.48	6.5	0.33	0.39

Table 8.5 Performance of the indirect biomarkers evaluated in this study using the thresholds described in their original publications in detection of advanced fibrosis (F0-2 Vs F3-4, n=48 Vs n=32, prevalence 40%).

Marker	HA Assay used	Original Publication test threshold and performance			Performance of original threshold in current Study								Current Study Thresholds	
		Threshold	Sens	Spec	Sens	Spec	number –ve Vs. +ve	PPV	NPV	LR+	LR-	DOR	Q-point	Youden
AST:ALT	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.71	0.78
APRI	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.47	0.40
FORNS	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	5.64	5.64
FIB4	-	1.45	74.3	80.1	81.3	64.6	37 Vs. 43	61%	84%	2.3	0.29	7.9	1.68	1.44
FIBROMETER 3G	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.86	0.84
HA	Siemens	60	88	59	59	85	66 Vs. 14	72%	76%	3.9	0.48	8.2	47	57
	Pharmacia				47	91	64 Vs. 16	78%	72%	5.2	0.58	9.0	55	68
HEPASCORE	Siemens	0.5	88	74	90	50	27 Vs. 53	55%	88%	1.8	0.2	9	0.73	0.72
	Pharmacia				91	31	17 Vs. 63	47%	84%	1.3	0.29	4.5	0.74	0.81
FIBROMETER 2G	Siemens	0.628	84	79	97	50	25 Vs. 55	57%	96%	1.94	0.06	32.3	0.82	0.81
	Pharmacia				91	50	27 Vs. 53	55%	89%	1.82	0.18	10.1	0.84	0.73
ELF	Siemens	9.59	85	63	65	82	50 Vs. 30	71%	78%	3.19	0.23	13.8	9.32	9.77
	Pharmacia				94	29	16 Vs. 64	47%	88%	1.32	0.20	6.4	9.94	10.07
FIBROSPECT II	Siemens	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.36	0.42
	Pharmacia				N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.35	0.34

Ability to Discriminate Cirrhosis (F0-3 Vs. F4)

All the markers tested achieved AUROC >0.8 in their ability to discriminate between patients with and without cirrhosis apart from AST:ALT ratio and APRI. Forns index generated the highest AUROC (0.92) (table 8.8).

Overall Performance (Obuchowski Measure) (Table 8.2, Figures 8.2- 8.3)

The highest unweighted Obuchowski measure (ordROC) was attained by Fibrometer 2G (0.94) and Fibrometer 3G (0.94) which were significantly higher *(p-value<0.05)* than those attained by AST:ALT ratio (0.86) and APRI (0.88) but none of the other markers tested (table 8.9). Of the 2 direct markers tested, ELF generated the highest ordROC (0.93) which was the best performing 'non disease-specific' marker of fibrosis together with Hepascore (0.93)

Performance of the markers at their published thresholds for the detection of moderate fibrosis (F2-4) – prevalence 63% (Table 8.4)

Diagnostic thresholds have been described for all of the markers in this context other than AST:ALT ratio.

- **Indirect Markers**

Overall, the sensitivity and specificity of the indirect markers at their published thresholds for detecting moderate fibrosis were comparable to those observed in their original publications other than Fibrometer 3G (more sensitive (100%), less specific (30%)). The highest positive likelihood ratio for the detection of moderate fibrosis (2.6) was generated by the proposed threshold of APRI (PPV of 82% in the study population).

- **Direct and Hybrid Markers**

The sensitivity and specificity of ELF, Hepascore and Fibrospect II at their published thresholds for the detection of moderate fibrosis in this study were comparable to those observed in their original publications. This was not the case for the proposed diagnostic threshold of Fibrometer 2G which was markedly more sensitive (100%) but less specific (30%). The highest PPV (88%) and LR+ (4.1) for the detection of moderate fibrosis were obtained using the published thresholds of ELF.

Performance of the markers at their published thresholds for the detection of advanced fibrosis (F3-4) – prevalence 40% [Table 8.5]

Diagnostic thresholds have not been proposed for a diagnosis of advanced fibrosis for Fibrospect II or any of the indirect markers other than FIB4.

- **Indirect Markers**

The performance attributes of the proposed threshold of FIB4 (1.45) for the detection of advanced fibrosis appeared consistent with its performance in this study population (Q-point 1.44). This FIB-4 threshold generated a positive likelihood ratio of 2.3 (PPV 61%), negative likelihood ratio of 0.29 (NPV 84%) and diagnostic odds ratio of 7.9 for the detection of advanced fibrosis.

- **Direct and Hybrid Markers**

Whilst the performance of direct and hybrid markers at their proposed thresholds for the detection of advanced fibrosis was comparable to that seen in their original publications, the Q-point and Youden index of ELF and HA in this study were the most similar to their proposed thresholds. The lowest negative likelihood ratio for the exclusion of advanced fibrosis was attained by Fibrometer 2G (0.06) which generated a NPV of 96%. The highest positive likelihood ratio for the detection of advanced

fibrosis was generated by HA itself (3.9) which resulted in a PPV of 72%. At their proposed thresholds for the detection of advanced fibrosis, Fibrometer 2G generated the highest diagnostic odds ratio (32.3).

Exploring whether combinations of biomarkers are more effective using logistic regression

- **Moderate fibrosis (F2-4)**

Stepwise logistic regression identified that both Fibrometer 2G *(p-value<0.001, OR 4.98, 95% CI 13.1-19029)* and ELF *(p-value=0.008, OR 3.07, 95% CI 1.16-8.09)* were significantly associated with a diagnosis of at least moderate fibrosis after accounting for the remaining variables as potential confounders. A combination of Fibrometer 2G and ELF had an AUROC of 0.90 (95% CI 0.84-0.97).

- **Advanced Fibrosis (F3-4)**

Stepwise logistic regression identified that both Forns index *(p-value<0.001, OR 2.16, 95% CI 1.32-3.55)* and ELF *(p-value=0.001, OR 2.72, 95% CI 1.39-5.34)* exhibited statistical independence for a diagnosis of advanced fibrosis. A combination of Forns and ELF had an AUROC of 0.87 (95% CI 0.78-0.96)

- **Cirrhosis (F4)**

Stepwise logistic regression identified that the Forns test was the only biomarker significantly associated with a diagnosis of cirrhosis after accounting for the remaining markers as potential confounders *(p-value<0.001, OR 3.019, 95% CI 1.50-6.05).*

Effect of using a different assay for HA to calculate direct and hybrid Panels (table 8.3)

Relationship between the 2 HA assays (figure 8.3)

Both assays exhibited a high degree of correlation (R=0.954, R^2=0.909, p<0.0001). The relationship between the 2 assays is described by the equation:

$HA_{(Pharmacia)} = 33.67 + [0.49 \times HA_{(Siemens)}]$

Effect of the Pharmacia HA on the ability of the direct and hybrid markers to discriminate between different degrees of fibrosis (Tables 8.3, 8.6-8.9))

The AUROC of HA as assessed by the Siemens assay was significantly higher than that attained using the Pharmacia assay for the detection of moderate fibrosis (AUROC 0.79 Vs 0.68, *p-value=0.005*) and advanced fibrosis (AUROC 0.80 Vs.0.72, *p-value=0.042*) but not for the detection of cirrhosis (AUROC 0.89 Vs. 0.85, *p-value>0.05*). Furthermore, the overall discriminatory power of the Siemens assay was significantly higher than that of the Pharmacia assay (ordROC 0.92 Vs. 0.88, *p-value=0.003*). Calculation of the hybrid and direct markers with the Pharmacia assay resulted in numerically lower AUROCs than when calculated with the Siemens assay. This was most marked in the context of Fibrospect II in its ability to discriminate between moderate fibrosis (AUROC 0.66 Vs. 0.79) and advanced fibrosis (AUROC 0.83 Vs. 0.70).

Effect of the Pharmacia HA on the performance of the direct and hybrid markers at their published diagnostic thresholds

Whereas Fibrometer 2G was largely unaffected, the use of the Pharmacia assay resulted in a reduction in the performance of the other direct and hybrid markers at their published diagnostic thresholds. The most marked impact was seen when the

ELF test was calculated with the 2 different assays for HA. When calculated with the ELISA HA assay, the ELF test at a threshold of 9.13 had a positive likelihood ratio of 3.9 (PPV 88%) and negative likelihood ratio of 0.36 (NPV 62%) for the detection of moderate fibrosis. However when calculated with the radiometric assay the ELF test at a threshold of 9.13 had a positive and negative likelihood ratio of both 1. This resulted in an unchanged *a priori* and *a posteriori* probability of moderate fibrosis regardless of the ELF score obtained in this study (figure 8.3).

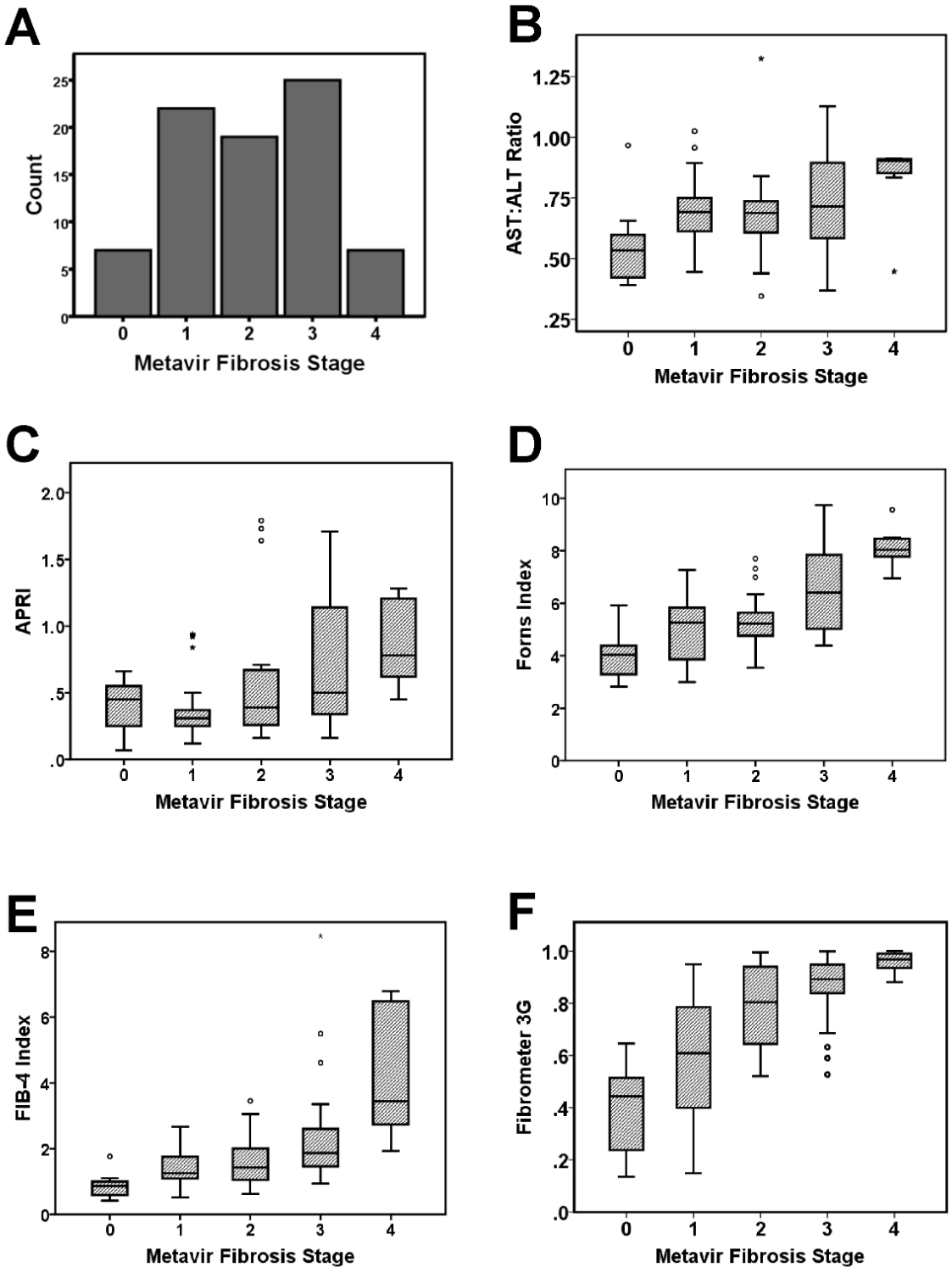


Figure 8.2
Panel A: Distribution of fibrosis stages in the study population. Panels B-F: Boxplots of indirect serum markers with respect to fibrosis stage in study population
(Panel B: AST to ALT ratio, Panel C: APRI, Panel D: Forns Index, Panel E: FIB-4 Index, Panel F: Fibrometer 3G).

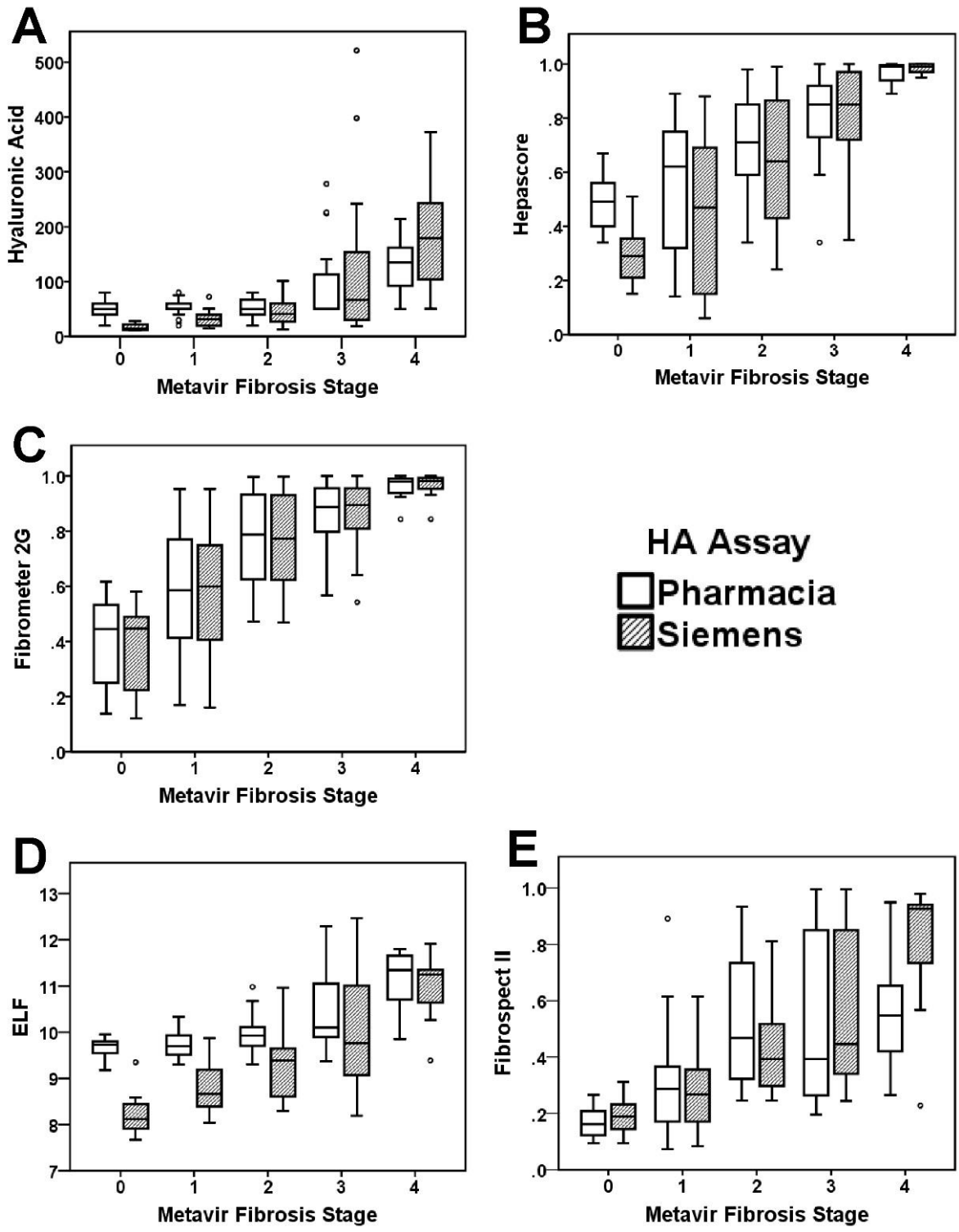


Figure 8.3 Boxplots of direct and hybrid serum markers calculated with two assays for hyaluronic acid with respect to fibrosis stage in the study population

Panel A: Hyaluronic Acid, Panel B: Hepascore, Panel C: Fibrometer 2G, Panel D: ELF, Panel E: Fibrospect II

Table 8.6 Pairwise comparison (Delong) of the AUROC generated by each test for the detection of at least mild fibrosis (F2-4)

The method of Delong is used to compare AUROC for each respective biomarker with the p value for each pairwise comparison displayed in the table if less than 0.05.

AUROC		0.61	0.7	0.77	0.77	0.68	0.79	0.78	0.83	0.81	0.84	0.81	0.84	0.87	0.88
	Test	AST to ALT ratio	APRI	Forns	FIB4	HA (Pharm. HA)	HA (ELISA)	ELF (Pharm. HA)	ELF (ELISA HA)	Hepascore (Pharm. HA)	Hepascore (ELISA HA)	Fibrospect (Pharm. HA)	Fibrospect (ELISA HA)	Fibrometer 2G (Pharm. HA)	Fibrometer 2G (ELISA HA)
0.7	APRI	NS													
0.77	Forns	0.038	NS												
0.77	FIB4	0.018	NS	NS											
0.68	HA (Pharm.)	NS	NS	NS	NS										
0.79	HA (ELISA)	0.012	NS	NS	NS	0.005									
0.78	ELF (Pharm. HA)	0.023	NS	NS	NS	0.021	NS								
0.83	ELF (ELISA HA)	0.003	NS	NS	NS	<0.001	NS	NS							
0.81	Hepascore (Pharm.HA)	0.008	NS	NS	NS	0.006	NS	NS	NS						
0.84	Hepascore (ELISA HA)	0.002	0.025	NS	NS	<0.001	NS	NS	NS	NS					
0.81	Fibrospect (Pharm. HA)	0.01	NS	NS	NS	0.042	NS	NS	NS	NS	NS				
0.84	Fibrospect (ELISA HA)	0.002	0.046	NS	NS	<0.001	NS	NS	NS	NS	NS	NS			
0.87	Fibrometer 2G (Pharm. HA)	<0.001	<0.001	NS	0.018	<0.001	NS	NS	NS	NS	NS	NS	NS		
0.88	Fibrometer 2G (ELISA HA)	<0.001	<0.001	0.045	0.007	<0.001	NS	NS	NS	NS	NS	NS	NS	NS	
0.86	Fibrometer 3G	<0.001	0.002	NS	0.026	0.001	NS	NS	NS	NS	NS	NS	NS	NS	NS

Table 8.7 Pairwise comparison (Delong) of the AUROC generated by each test for the detection of at least moderate fibrosis (F3-4) The method of Delong is used to compare AUROC for each respective biomarker with the p value for each pairwise comparison displayed in the table if less than 0.05.

AUROC		0.65	0.71	0.82	0.81	0.72	0.8	0.79	0.82	0.79	0.83	0.66	0.79	0.81	0.83
	Test	AST to ALT ratio	APRI	Forns	FIB4	HA (Pharm. HA)	HA (ELISA)	ELF (Pharm. HA)	ELF (ELISA HA)	Hepascore (Pharm. HA)	Hepascore (ELISA HA)	Fibrospect (Pharm. HA)	Fibrospect (ELISA HA)	Fibrometer 2G (Pharm. HA)	Fibrometer 2G (ELISA HA)
0.71	APRI	NS													
0.82	Forns	0.035	0.043												
0.81	FIB4	0.027	NS	NS											
0.72	HA (Pharm.)	NS	NS	NS	NS										
0.8	HA (ELISA)	NS	NS	NS	NS	0.042									
0.79	ELF (Pharm. HA)	NS	NS	NS	NS	NS	NS								
0.82	ELF (ELISA HA)	0.031	NS	NS	NS	0.012	NS	NS							
0.79	Hepascore (Pharm.HA)	NS	NS	NS	NS	NS	NS	NS	NS						
0.83	Hepascore (ELISA HA)	0.012	0.035	NS	NS	0.006	NS	NS	NS	NS					
0.66	Fibrospect (Pharm. HA)	NS	NS	0.022	0.023	NS	NS	NS	0.018	0.036	0.004				
0.79	Fibrospect (ELISA HA)	0.049	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS			
0.81	Fibrometer 2G (Pharm. HA)	0.028	0.028	NS	NS	NS	NS	NS	NS	NS	NS	0.006	NS		
0.83	Fibrometer 2G (ELISA HA)	0.018	0.012	NS	NS	0.034	NS	NS	NS	NS	NS	0.003	NS	NS	
0.81	Fibrometer 3G	0.03	0.048	NS	NS	NS	NS	NS	NS	NS	NS	0.006	NS	NS	NS

Table 8.8 Pairwise comparison (Delong) of the AUROC generated by each test for the detection of cirrhosis (F4)

The method of Delong is used to compare AUROC for each respective biomarker with the p value for each pairwise comparison displayed in the table if less than 0.05.

AUROC		0.76	0.78	0.92	0.9	0.85	0.89	0.87	0.89	0.84	0.86	0.67	0.83	0.86	0.88
	Test	AST to ALT ratio	APRI	Forns	FIB4	HA (Pharm. HA)	HA (ELISA)	ELF (Pharm. HA)	ELF (ELISA HA)	Hepascore (Pharm. HA)	Hepascore (ELISA HA)	Fibrospect (Pharm. HA)	Fibrospect (ELISA HA)	Fibrometer 2G (Pharm. HA)	Fibrometer 2G (ELISA HA)
0.78	APRI	NS													
0.92	Forns	NS	NS												
0.9	FIB4	NS	NS	NS											
0.85	HA (Pharm.)	NS	NS	NS	NS										
0.89	HA (ELISA)	NS	NS	NS	NS	NS									
0.87	ELF (Pharm. HA)	NS	NS	NS	NS	NS	NS								
0.89	ELF (ELISA HA)	NS	NS	NS	NS	NS	NS	NS							
0.84	Hepascore (Pharm.HA)	NS	NS	NS	NS	NS	NS	NS	NS						
0.86	Hepascore (ELISA HA)	NS	NS	NS	NS	NS	NS	NS	NS	NS					
0.67	Fibrospect (Pharm. HA)	NS	NS	0.008	0.018	NS	NS	NS	0.044	NS	NS				
0.83	Fibrospect (ELISA HA)	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS			
0.86	Fibrometer 2G (Pharm. HA)	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	0.049	NS		
0.88	Fibrometer 2G (ELISA HA)	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	0.045	NS	NS	
0.86	Fibrometer 3G	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	0.019	NS	NS	NS

Table 8.9 Pairwise comparison (Delong) of the ability of the tests to discriminate between all degrees of fibrosis (ordROC).
The method of Delong is used to compare AUROC for each respective biomarker with the p value for each pairwise comparison displayed in the table if less than 0.05.

ordAUROC		0.86	0.88	0.92	0.92	0.88	0.92	0.91	0.93	0.92	0.93	0.9	0.93	0.94	0.94
	Test	AST to ALT ratio	APRI	Forns	FIB4	HA (Pharm. HA)	HA (ELISA)	ELF (Pharm. HA)	ELF (ELISA HA)	Hepascore (Pharm. HA)	Hepascore (ELISA HA)	Fibrospect (Pharm. HA)	Fibrospect (ELISA HA)	Fibrometer 2G (Pharm. HA)	Fibrometer 2G (ELISA HA)
0.88	APRI	NS													
0.92	Forns	0.016	0.003												
0.92	FIB4	0.02	0.003	NS											
0.88	HA (Pharm.)	NS	NS	0.025	0.019										
0.92	HA (ELISA)	0.026	0.035	NS	NS	0.003									
0.91	ELF (Pharm. HA)	0.042	NS	NS	NS	0.014	NS								
0.93	ELF (ELISA HA)	0.009	0.009	NS	NS	<0.001	NS	NS							
0.92	Hepascore (Pharm.HA)	0.011	NS	NS	NS	0.004	NS	NS	NS						
0.93	Hepascore (ELISA HA)	0.003	0.007	NS	NS	<0.001	NS	NS	NS	NS					
0.9	Fibrospect (Pharm. HA)	NS	NS	NS	NS	NS	NS	NS	NS	NS	0.035				
0.93	Fibrospect (ELISA HA)	0.004	0.033	NS	NS	<0.001	NS	NS	NS	NS	NS	NS			
0.94	Fibrometer 2G (Pharm. HA)	0.003	<0.001	NS	NS	<0.001	NS	NS	NS	NS	NS	0.01	NS		
0.94	Fibrometer 2G (ELISA HA)	0.001	<0.001	NS	NS	<0.001	NS	NS	NS	NS	NS	0.006	NS	NS	
0.94	Fibrometer 3G	0.003	<0.001	NS	NS	<0.001	NS	NS	NS	NS	NS	0.001	NS	NS	NS

DISCUSSION

In this study, the diagnostic performance of 10 serum markers of fibrosis in a population of patients with CHC have been compared by studying their ability to discriminate between moderate fibrosis, advanced fibrosis, cirrhosis, all fibrosis stages (ordROC) and by validating their published diagnostic thresholds. Within this cohort, all the tested direct and hybrid markers had good diagnostic performance for detecting moderate fibrosis. However, the only indirect marker that achieved this level of performance for the detection of moderate fibrosis was Fibrometer 3G. Aside from APRI and AST:ALT ratio, the comparisons between the other 8 markers in their ability to discriminate between advanced fibrosis and cirrhosis were more favorable. Both the Forns index and FIB-4 exhibited excellent performance for discriminating between patients with and without a histologic diagnosis of cirrhosis. This is biologically plausible given that these indirect markers of fibrosis incorporate the platelet count which is affected by cirrhosis and portal hypertension. With regard to overall fibrosis stratification, both of the virus specific patented panels, Fibrometer 2G and 3G, exhibited the highest level of performance. Of the other 8 panels which have been proposed to be used in all aetiologies of liver disease, the best performing tests were ELF and Hepascore.

Using logistic regression, indirect and direct markers demonstrated statistical independence for the detection of mild to advanced fibrosis. Whilst these combinations of indirect and direct markers resulted in an improvement in performance that was statistically significant, this advantage must be balanced against the need to perform more tests with the consequent increase in costs. In this study it was observed that markers with only 3 components (ELF and Fibrospect II) have comparable performance to those with as many as 7 components (Fibrometer 2G and 3G).

Furthermore, some of the more complex tests are dependent on the inclusion of demographic data that requires the collection, transmission and entry of clinical data with attendant costs and potential for error.

Within this population of patients with CHC and prior treatment failure, many of the tests performed differently at their diagnostic thresholds than reported within their reference studies in CHC. This may relate to sample size and variation in spectrum bias. Regardless, within this study, ELF, Fibrospect II, Forns index and FIB-4 exhibited the most consistent performance at their proposed diagnostic thresholds in detecting moderate and advanced fibrosis.

The influence of assay selection for HA was also explored in this study. The use of 2 closely correlated (r^2=0.909) assays for HA to calculate the direct and hybrid markers resulted in changes both in overall diagnostic performance and at the diagnostic thresholds of the biomarkers. Despite the close correlation of these assays, the AUROC of HA as measured by the radiometric assay was significantly lower than that measured by the ELISA assay in the detection of moderate and advanced fibrosis. Furthermore, when the direct and hybrid markers were calculated with HA as measured using the radiometric assay, a modest reduction in the AUROC was observed in most instances. This was not the case, however, for Fibrospect II which suffered a marked reduction ability to discriminate moderate fibrosis (AUROC 0.79 Vs.0.66) and advanced fibrosis (AUROC 0.83 Vs.0.70). With regard to the performance of the biomarkers at their diagnostic thresholds, the use of a different assay for HA had the most marked effect of the performance of the ELF. Despite generating similar AUROC, the use of the Pharmacia assay rendered the ELF test much less useful as a diagnostic tool with both a positive or negative result at its proposed threshold producing an unchanged *a priori* and *a posteriori* probability of

moderate fibrosis in this cohort. These variations in test performance resulting from the choice of assay, have predominantly affected the markers with the fewest number of analytes. Whereas both ELF and Fibrospect II have only 3 constituent components the hybrid assays with up to 7 components are more forgiving of the use of an inferior performing component. This emphasises the need to use the specified individual component assays that have been validated for a particular biomarker both in research studies and in clinical practice. The observation that the diagnostic performance of the second generation Fibrometer was resistant to the choice of HA assay has also been made by previous investigators.[365]

Regardless of treatment status, patents with CHC and a diagnosis of cirrhosis should embark on a surveillance program for the early detection of hepatocellular carcinoma and endoscopic features of portal hypertension. These data have shown that all 10 of the markers studied in this population have good performance (AUROC >0.8) for the detection of cirrhosis with some having excellent performance (AUROC>0.9). This level of performance from even the more basic indirect tests is encouraging particularly as not all centers will have access to some of the expensive proprietary panels that we have investigated.

In summary, this study compared head to head the performance of 10 biomarkers in a population of patients with CHC and prior treatment failure. Many of the biomarkers tested have demonstrated good diagnostic performance for their ability to discriminate between moderate and advanced fibrosis stages. Overall, the best performing markers were the virus specific panels, Fibrometer 2G and 3G. With regard to biomarkers developed for use in all aetiologies of liver disease, the best performing panels were ELF and Hepascore. Importantly, it was identified that the performance of the markers at their diagnostic thresholds can be variable and influenced by the choice of their

component assays. These results emphasise the need to use component assays that have been validated for a particular biomarker.

CHAPTER 9

SERUM MARKERS OF LIVER FIBROSIS: ON TREATMENT CHANGES OF DIRECT MARKERS PREDICT THE OUTCOME OF ANTI-FIBROTIC THERAPY

ABSTRACT

Background & Aims: the utility of non-invasive serum markers to longitudinally monitor liver fibrosis is not established.

Methods: this study included 70 patients with chronic hepatitis C who having previously failed interferon-based antiviral therapy, were randomized to receive pegylated interferon with or without silymarin for 24 months. ELF tests (HA, PIIINP, TIMP-1) were performed on patient sera taken prior to, during and at the end of the study (0, 12, 24 months) and liver histology obtained prior to and at the end of the study.

Results: following the study, absolute changes in Ishak fibrosis stage and ELF ranged from -4 to +4 and -2.41 to +2.68, respectively. Absolute changes in ELF at study mid-point were significantly associated with changes in both ELF and histology at the end of the study. A model combining both baseline ELF and change of ELF at study mid-point was able to predict the end of study ELF *(R^2=0.609, p-value<$1x10^{-11}$)*, a decrease in ELF *(AUROC:0.80-0.85)*, and a rise in ELF *(AUROC:0.81-0.85)*. Furthermore, a model combining both baseline histologic stage and ELF together with the change of ELF at study mid-point was able to predict end of study histology *(R^2=0.601, P-value<$1x10^{-11}$, AUROC:0.88-0.92)*, histologic fibrosis regression *(AUROC:0.81-0.84)* and progression *(AUROC:0.86-0.91).*

Conclusions: these observations suggest that a change in the direct serum marker ELF predicts changes in liver fibrosis over a longer period. These data support the use of ELF as a surrogate marker of liver fibrosis evolution in monitoring anti-fibrotic treatments thus permitting “response-guided” therapy by the early identification of patients who will benefit from prolonged anti-fibrotic treatment.

INTRODUCTION

Progression of CLD to cirrhosis is increasingly recognized as highly important, if not the most important clinical endpoint, since cirrhosis incurs a high risk of portal hypertension, liver failure and hepatocellular carcinoma.[366, 367] The histological staging of a liver biopsy remains the reference standard for assessing hepatic fibrosis. However, reliability of liver biopsy is limited due to sampling error [181], inter- and intra-observer variability[351], and procedural complications.[28] As a result, several non-invasive methods have been developed for the cross-sectional staging of liver fibrosis. These include both direct (ELF[221], Fibrospect[362]) and indirect (APRI[359], Fibrotest[223], Hepascore[224], Fibrometer[368]) serum markers and imaging techniques (Fibroscan[369], ARFI[370], MR elastography).[371] Whereas indirect serum markers of liver fibrosis are combinations of serum parameters which are related to liver function including AST and ALT, biological plausibility links direct serum markers of fibrosis to either fibrolytic or fibrogenic processes involved in liver matrix turnover.[3]

Compared to liver histology, non-invasive methods have demonstrated robust performance in the detection of moderate or advanced hepatic fibrosis in a variety of chronic liver diseases. What has not been demonstrated, however, is their utility in tracking histologic fibrosis longitudinally during the natural history of a patient's liver disease or during anti-fibrotic therapy.

This study was based on the PROFI-C (Progression of Fibrosis Inhibition in Hepatitis C) randomized trial which investigated whether previous non-responders or relapsers to interferon-based therapy showed delayed fibrosis development after 24 months of treatment with a combination of interferon alpha and either silymarin or placebo. Patients in the PROFI-C study had serum samples taken prior to, at the mid-point and

at the end of therapy. The scientific rationale for the study was based upon previous work where both interferon alpha and silymarin were studied as putative anti-fibrotic agents. Earlier studies suggested that the administration of interferon alpha to patients with CHC was associated with significant histological improvement[148, 372] and silymarin demonstrated a marked anti-fibrotic effect in rodent models of hepatic fibrosis.[373-375]

Whilst the PROFI-C study itself did not demonstrate an appreciable difference in histological outcomes between the treatment arms[352], the serum and histological samples taken during the trial have provided an invaluable platform to evaluate longitudinal changes of liver fibrosis as assessed by both non-invasive serum markers and liver histology. This cohort therefore permitted an exploration of whether changes in liver fibrosis (as assessed by both liver histology and serum markers at the end of the study period) were associated with baseline and on-treatment changes in serum markers.

METHODS

Study Design

Patients in this study were enrolled in the PROFI-C trial. Written informed consent was obtained from all patients before admission to the study. Ethical approval was granted by the local ethics committees of the participating centers in accordance with the guidelines of the 1975 Declaration of Helsinki. PROFI-C was an investigator-initiated, prospective, randomized trial involving 18 centers in Germany and Austria, investigating the effect of high dose silymarin plus pegylated interferon alpha 2b (PEG-INFα2b, PegIntron, Essex Pharma GmbH, Munich, Germany) in non-responders or relapsers to standard treatment for CHC. Whereas 108 patients were enrolled into the PROFI-C trial[352], only patients who underwent consecutive liver biopsy (prior to and after the 24 month therapy) and had stored sera taken prior to, at the mid-point, and at the end of therapy (0, 12, 24 months) were evaluated in this study (n=70). Participants were male or female patients aged between 18 and 65 years with chronic hepatitis C infection and had evidence of CHC (positive tests for anti-HCV antibodies and HCV-RNA (COBAS Amplicor HCV Monitor, Roche Molecular Diagnostics, Mannheim, Germany) after failure of first-line therapy with either interferon or pegylated interferon and ribavirin. Patients were also required to have histologically proven chronic hepatitis on a liver biopsy specimen (at least 8 identifiable portal tracts) within 6 months prior to entry into the study.

Exclusion criteria included treatment with silymarin, steroids or immunosuppressive drugs in the preceding three months, acute hepatitis, Child-Pugh stage B or C cirrhosis, thrombocytopaenia (<100 x 10^9/L), leucopenia (<3 x 10^9/L), other chronic liver diseases, history of liver or kidney transplantation, autoimmune diseases, HIV infection, active hepatitis B infection, alcohol abuse (defined as the consumption of

>40g per day in males and 20g per day in females) in the preceding 6 months, active drug abuse, pregnancy and lactation, severe somatic (renal, cardiac, pulmonary, gastrointestinal, oncologic) or psychiatric diseases and depression.

Randomization and data collection were performed at the Department of Medicine, University Hospital Erlangen, and all virologic and serologic analyses were performed according to standardized laboratory routines.

Treatment Schedule

All patients were treated with subcutaneous PEG-INFα2b at either 100 µg per week or at 50 µg on alternate weeks. Treatment was combined with oral silymarin (Bionorica Arzneimittel, Neumarkt, Germany) treatment at 280 mg three times per day (280mg per capsule) or an identically encapsulated placebo filled with glucose and soy bean extract in one of the 4 following treatment regimens:

1) PEG-INFα2b (100 µg/week) + silymarin

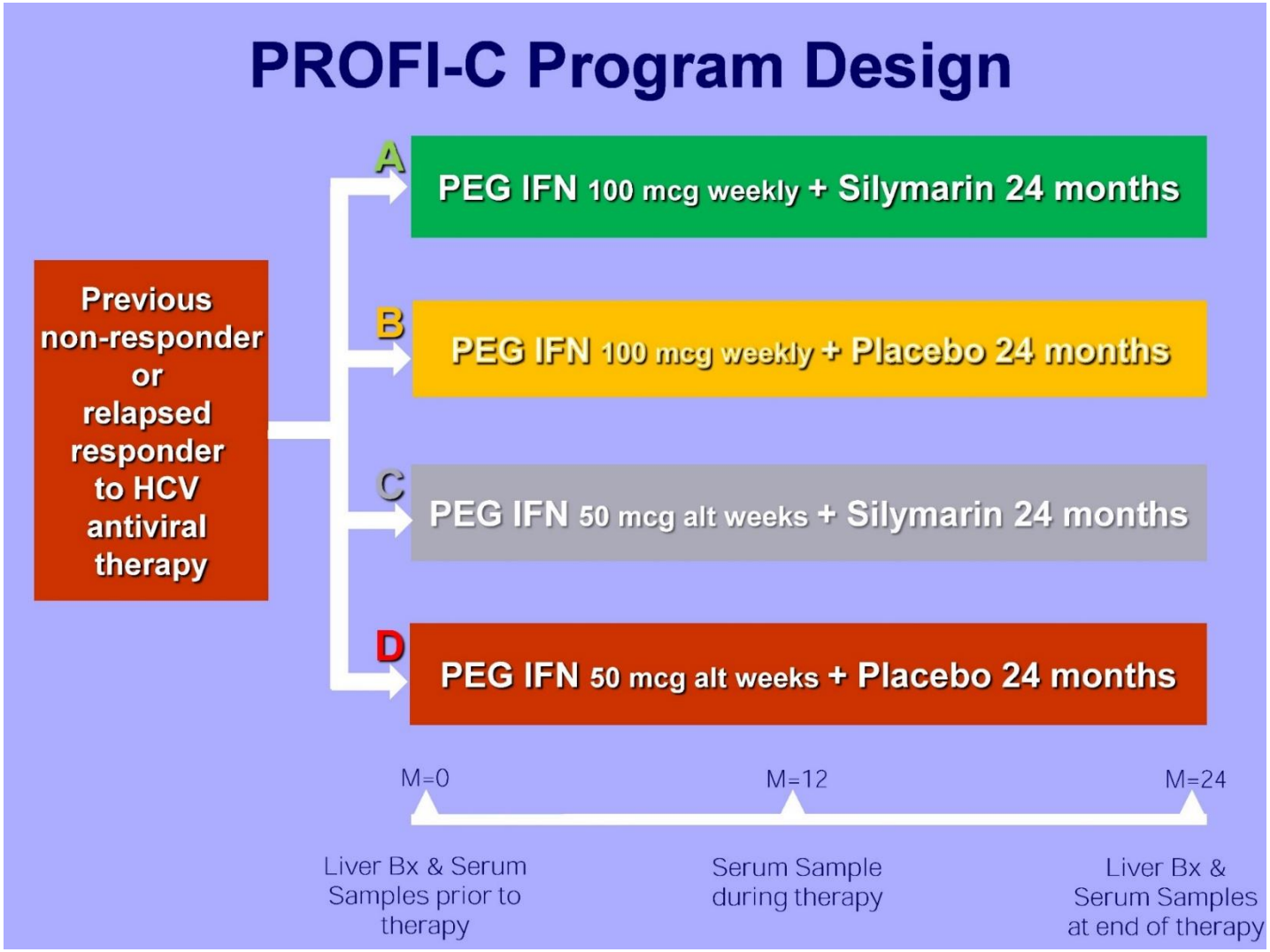
2) PEG-INFα2b (100 µg/week) + placebo

3) PEG-INFα2b (50 µg/every other week) + silymarin

4) PEG-INFα2b (50 µg/every other week) + placebo

The treatment period was 24 months with an additional 3 months of post-treatment surveillance. During the therapy, patients were evaluated 3 monthly to monitor for side effects, compliance and changes in hematological, biochemical and virologic parameters.

Figure 9.1 PROFI-C Randomised Controlled Trial Design



Histological assessment

Liver histology was obtained at the beginning of the study and at the end of the treatment period (month 24 ± 3 months). Biopsies were fixed in formalin and embedded in paraffin. Hematoxylin-Eosin staining was used for grading of inflammation and the Chromotrope-aniline blue staining for staging the amount of liver fibrosis. [353, 354] All specimens were graded and staged by using the Ishak score.[187] Histological assessment was performed by 2 independent pathologists (D.N. and O.D.) who were blinded to the clinical data and randomization status of the patients in the study. Interobserver variability was determined by the Kappa statistic (Kappa=0.624). All liver biopsy specimens that were discordantly staged were re-reviewed by both pathologists with a final score determined after further discussion.

Sample Collection and Serum Marker testing

Sera were stored at -70ºC prior to transfer to the central laboratory, where ELF tests were performed on thawed samples. Serum samples were analyzed for levels of TIMP-1, HA and PIIINP using the proprietary assays developed for the ELF test by Siemens Healthcare Diagnostics Inc (Tarrytown, New York, USA). These assays are magnetic particle separation immunoassays and were performed on the ADVIA Centaur® immunoassay system (Siemens Medical Solutions Diagnostics Inc, Tarrytown, New York, USA). Results were entered in to the manufacturer's published algorithms appropriate for the analyzer used to test the samples and to derive an ELF score.

Virologic Analysis

HCV RNA was quantified using an in-house HCV RNA real time RT-qPCR [355] using the QIAamp96 Virus nucleic acid purification procedure on the BioRobotMDx (Qiagen, Hilden, Germany) and the ABI Prism 7500 real-time PCR with Qiagen QuantiTect probe RT-PCR reagents. The assay uses brome mosaic virus RNA as an internal control, introduced at the extraction stage.

HCV genotyping was performed by amplifying and sequencing a region of the 5'NCR. The sequence was analyzed to compare probe binding sites of the LiPa method [356] and by finding the restriction sites.[357] These two virtual methods were compared to give the HCV genotype result.

Statistical Analyses

Statistical analyses were performed using SPSS for Windows (version 20, SPSS Inc, Chicago, IL) and R for Windows (version 2.15.1, the R Foundation for Statistical Computing). Patient demographic and clinical laboratory characteristics were descriptively summarized and reported as mean ± standard deviation (SD) and range. All tests were two-sided and statistical significance assessed at the 0.05 threshold. The diagnostic performance of ELF as compared to liver biopsy was assessed using ROC curves. The AUROC and 95% confidence intervals of AUROC were calculated. The Obuchowski[376] method of correcting for spectrum effect was applied in a similar fashion to previously published literature and the earlier chapters.[342] The ordROC gives a weighted average of the N(N-1)/2 AUROC pairwise comparisons between N categories of gold standard outcome. Thus, using the Ishak scale with its N (=7) categories of fibrosis staging (F0-6) there are 21 pairwise comparisons between 2 of the N categories. Accordingly, a penalty function proportional to the difference in Ishak

units between fibrosis stages was defined. The penalty function was 0.17, 0.33, 0.5, 0.67, 0.83 and 1 when the difference between Ishak stages was 1, 2, 3, 4, 5 and 6 stages respectively. As the severity of histological liver fibrosis in patients with CHC with prior treatment failure has not been well characterized, the Obuchowski measure presented in this study has not been weighted according to the prevalence of fibrosis stages in a reference population. A 2-sided t-test was used to assess changes of mean Ishak biopsy and serum marker scores as parametric variables arising during the study period. Univariate correlation coefficients (Spearman's Rho) were calculated to assess the association between changes in serum marker scores and liver histology occurring at the end of the study period with baseline serum markers and changes in serum markers occurring at the mid-point of the study period. Linear and logistic regression were used to construct models incorporating baseline and on-treatment variables that were predictive of continuous and categorical variables respectively. The clinical utility of these models for predicting fibrosis progression and regression at the end of the study period was assessed using AUROC analysis. Sensitivity, specificity and predictive values were calculated at thresholds derived from ROC curves.

RESULTS

Patient characteristics and baseline histology

The baseline characteristics of the 70 patients included in this study are displayed in table 9.1. The participants in this study were comprised of patients who had been randomized into each of the 4 arms of the PROFI-C trial. Patients were mostly male and had a mean age of 48.5 years, with predominantly CHC genotype 1. Baseline haematological and biochemical parameters were compatible with compensated chronic liver disease.

The distribution of mean Ishak fibrosis score and mean ELF score prior to therapy are displayed in table 9.2. All 7 Ishak stages are represented with 26% of patients having severe fibrosis/cirrhosis (F5-6). The mean Ishak fibrosis stage prior to therapy was 2.9.

Effect of therapy on HCV RNA

The mean change in HCV RNA at month 12 was -0.78 log (range -7.20 to 2.14). Suppression of HCV RNA was more marked in those patients receiving PEG-IFNα2b 100mcg (*p=0.008*). Seven patients achieved full suppression of HCV RNA during therapy of which 5 had been randomized to the PEG-IFNα2b 100mcg group. No patients achieved SVR. The addition of silymarin did not influence outcome (*p-value>0.05, non-significant (NS)*).

Table 9.1 Baseline Patient Characteristics

Data are presented as mean ± s.d. unless described otherwise

Variable	Total (n=70)
Age (years)	48.5 ± 9.6
Male (n,%)	40, 57%
BMI	24.3 ± 4.2
Genotype 1 (n,%)	47, 67%
Genotype 2 (n,%)	5, 7%
Genotype 3 (n,%)	18, 26%
Log Viral Load IU/ml (range)	6.21(3.84 - 8.20)
Hb (g/dl)	15.2 ± 1.4
WCC (10^9/L)	6.4 ± 1.7
PLT (10^9/L)	210 ± 62
ALT (IU/L)	117 ± 10
Bilirubin (µmol/L)	12.7 ± 7.5
Albumin (g/L)	47.7 ± 7.6
PT (seconds)	12.1 ± 1.2
AFP (ng/ml)	8.2 ± 11.7
PEG 50 + Placebo (n,%)	21 (30%)
PEG 50 + Silymarin (n,%)	18 (26%)
PEG 100 + Placebo (n,%)	10 (14%)
PEG 100 + Silymarin (n,%)	21 (30%)

Table 9.2 The distribution of fibrosis stages prior to and at the end of the study period

Ishak Fibrosis Stage	Pre-Study	Post-Study
	n (%)	n (%)
F0	5 (7)	3 (4)
F1	18 (26)	20 (29)
F2	9 (13)	8 (11)
F3	10 (14)	17 (24)
F4	10 (14)	7 (10)
F5	14 (20)	7 (10)
F6	4 (6)	8 (11)
ELF (range)	7.67-12.47	6.53-13.59

Baseline and post study marker performance in discriminating between fibrosis stages

The ELF test exhibited good performance in the detection of histological fibrosis at baseline and at the end of the study period. The performance of ELF in detecting severe fibrosis or cirrhosis was similar (AUROC 0.87-0.88) with the Obuchowski measure (ordAUROC) comparable both at baseline and at the end of the study period (ordAUROC 0.93-0.92, standard error 0.02).

A heterogeneous effect on both serum markers and histology is observed in the study

Absolute changes in both histology and ELF score observed during the study are presented graphically in figure 9.2. At the end of treatment, absolute changes in Ishak stage and ELF score ranged from -4 to +4 and -2.41 to +2.68 respectively. Twenty one and 24 patients were noted to have an absolute decrease and increase in Ishak fibrosis score at the end of the study period, respectively (table 9.2 and figure 9.2). Similarly, 27 and 43 patients were noted to have an absolute decrease and increase in ELF score at the end of the study period respectively.

The dosage of pegylated interferon and the addition of silymarin did not influence anti-fibrotic outcomes *(p-value=NS).* In addition, the degree of viral suppression on therapy also did not influence outcome *(p-value=NS).*

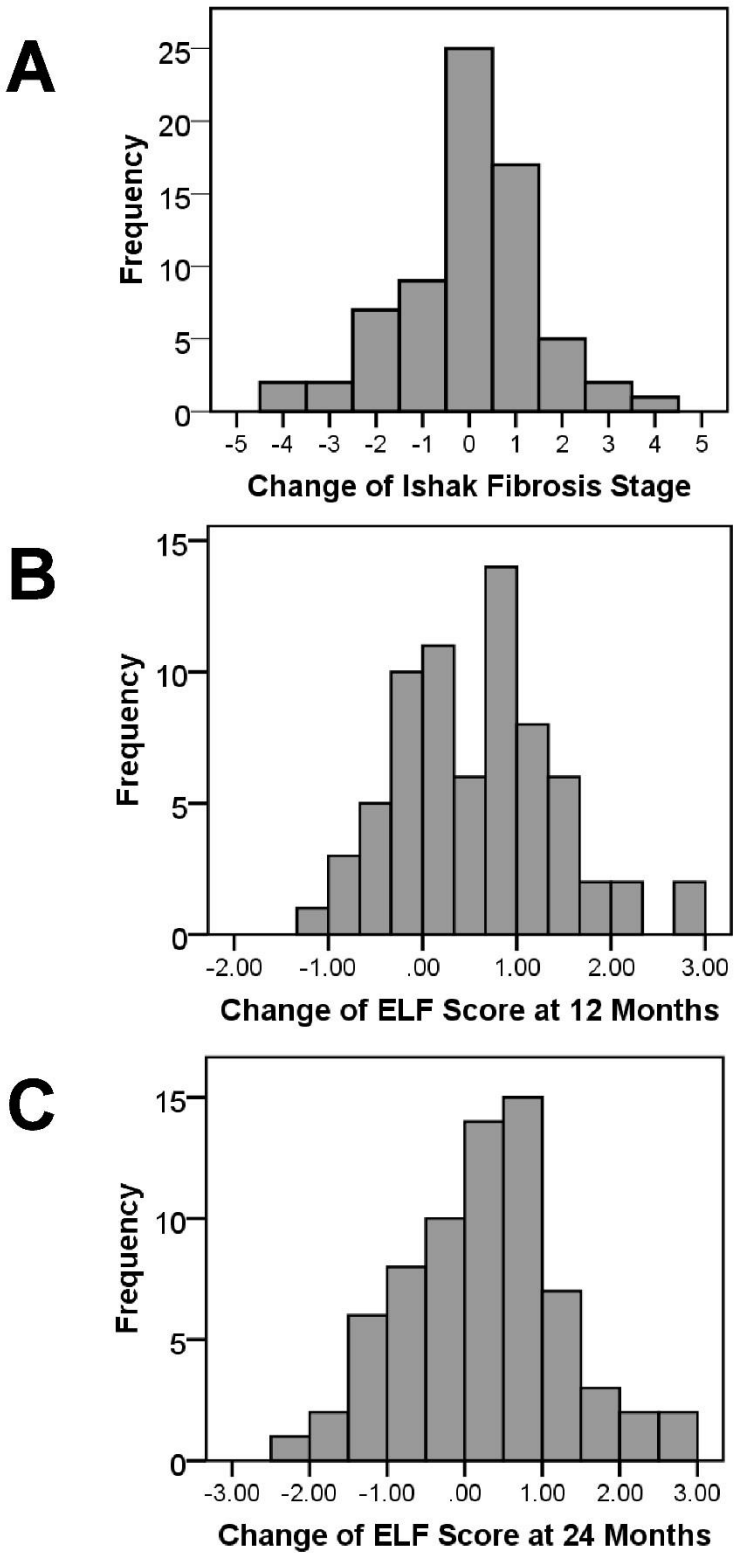


Figure 9.2 Histograms displaying the absolute changes of:

(A) Ishak Fibrosis Stage between 0-24 months, (B) ELF between 0-12 months, and (C) ELF between 0-24 months

Baseline and on treatment levels of ELF scores are associated with the evolution of ELF scores at the end of therapy

After stratifying patients by their end of treatment changes in ELF score (figure 9.3), it was evident that mean baseline ELF scores were significantly higher *(p-value=0.043)* in patients who experienced a decrease in ELF score at the end of study than in those who experienced an increase in ELF score at the end of the study. Stepwise logistic regression identified the baseline ELF score as the only baseline factor associated with an end of study reduction in serum ELF *(OR 1.59, 95% CI 1.01-2.52, p-value=0.049).* With regard to the marker scores during the treatment period, absolute changes in ELF score between months 0 and 12 were found to correlate with absolute changes between months 0 and 24 *(r=0.599, p-value $<10^{-7}$)* suggesting that changes of ELF score that occurred between the start and the mid-point of the study were predictive of ELF scores at the end of the study. Moreover, mean ELF scores were observed to be significantly higher at the mid-point of the trial than they were at the beginning of the study (table 9.3) with the magnitude of this rise reflecting the outcome of the changes in ELF at the end of the study (figure 9.3).

Table 9.3 Mean Ishak fibrosis stage, Ishak inflammatory grade, serum marker scores and ALT observed at months 0, 12, and 24 of all patients and after stratifying patients by those with an end of therapy increase and decrease in inflammatory grade.

	All patients (n=70)			End of therapy Ishak grade reduced (n=34)			End of therapy Ishak grade unchanged or increased (n=36)		
	Month 0	**Month 12**	**Month 24**	**Month 0**	**Month 12**	**Month 24**	**Month 0**	**Month 12**	**Month 24**
Mean ALT	117.1	98.1	89.9*	122.9	86.4*	78.4*	111.6	109.3	100.8
Standard Deviation	100.9	69.0	59.8	115.8	65.2	59.7	85.8	71.6	58.7
Mean Ishak Inflammatory Grade	3.6		3.2	4.4		2.3*	2.8		4.1*
Standard Deviation	1.9		2.0	1.9		1.4	1.6		2.2
Mean Ishak Fibrosis Stage	2.9		2.8	2.9		2.4	2.8		3.3
Standard Deviation	1.8		1.8	1.7		1.6	1.9		1.8
Mean ELF	9.32	9.91*	9.52	9.28	9.61*	9.29	9.36	10.20*	9.75*
Standard Deviation	1.10	1.19	1.26	1.01	1.16	1.30	1.18	1.14	1.21
Mean HA	69.2	124.3*	95.6	65.2	105.2*	77.2	73.1	142.3*	112.9
Standard Deviation	88.2	144.3	141.8	64.7	125.4	120.3	106.6	159.7	159.2
Mean TIMP-1	738.7	743.1	749.6	732.9	715.9	757.3	744.1	768.8	742.2
Standard Deviation	215.5	196.1	226.3	215.2	141.4	203.0	218.6	235.7	249.0
Mean PIIINP	10.69	12.16*	10.50	9.71	10.24	9.04	11.63	13.98*	11.89
Standard Deviation	5.35	6.43	5.53	4.53	4.91	4.46	5.94	7.20	6.12

*significant difference from month 0 *(p-value <0.05)*

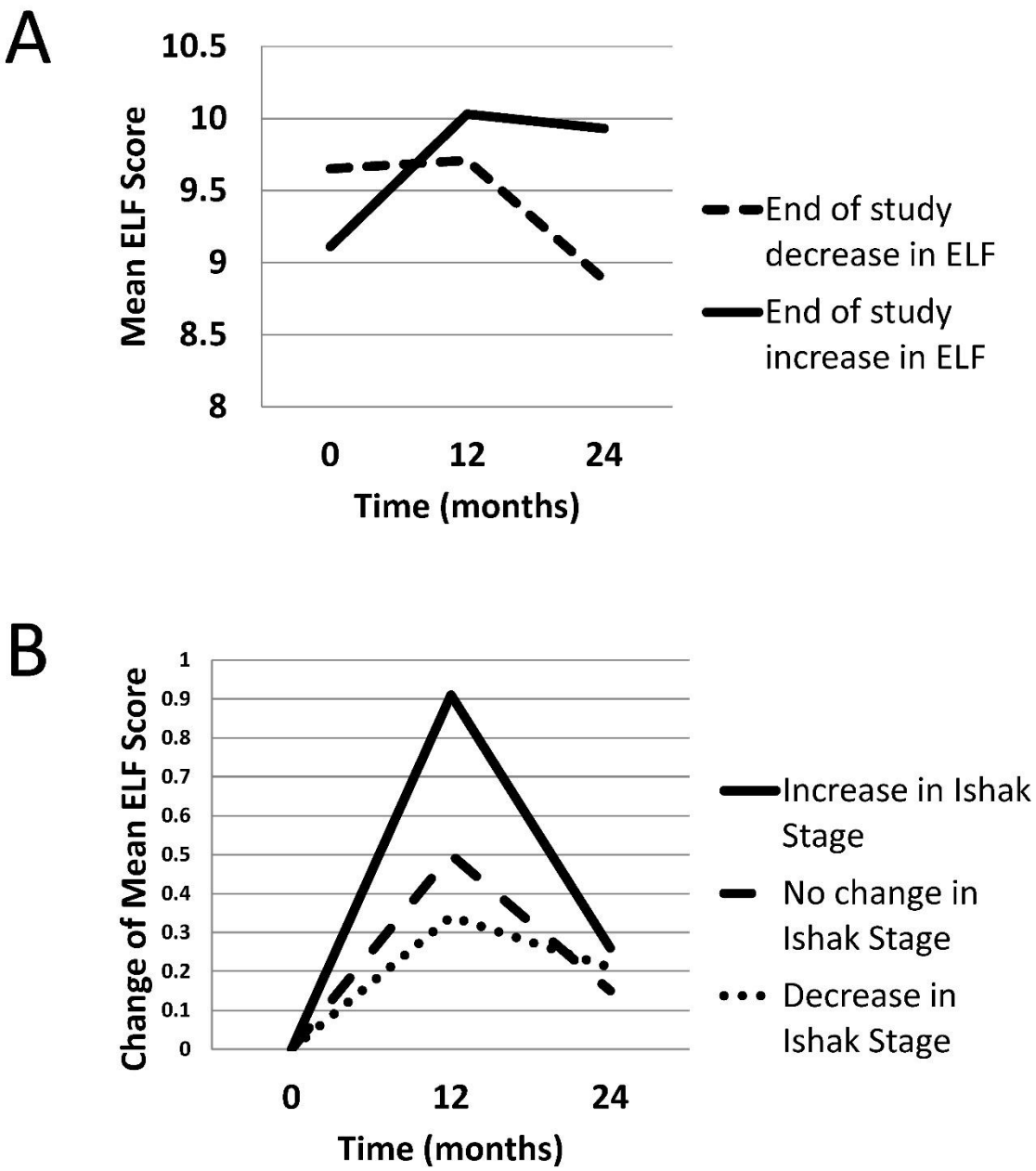


Figure 9.3 Changes in mean ELF during the study period after stratifying patients by the evolution of ELF (A), and by histologic fibrosis evolution (B) at the end of the study

Changes in serum markers during the study period reflect histological outcomes

Patients were divided into 3 categories based upon the histologic outcome determined at the end of the study period. Histological fibrosis regression or progression was defined as ≥1 stage decrease or increase in Ishak fibrosis stage respectively. Within these 3 categories, mean changes of ELF scores were evaluated. These data are presented graphically in figure 9.3. During therapy there was a rise in mean ELF score in all 3 categories of fibrosis evolution with the magnitude of this rise reflecting the change in histology observed at the end of the study. Patients with fibrosis progression had the largest rise in ELF followed by those with unchanged histology; those with fibrosis regression had the smallest change. Analogous to the pattern observed in the evolution of serum markers at the end of therapy, this effect was not sustained at the end of the study period with mean ELF scores at the end being lower than those seen at the mid-point of the study period. This pattern of rise and subsequent fall of serum markers was observed in patients receiving both dose schedules of pegylated interferon and in patients receiving either silymarin or placebo.

Changes in ELF score at the mid-point of the study correlate with changes in Ishak fibrosis score at the end of the study

The relationship between the absolute change in liver histological fibrosis (seen at the end of the study) and the absolute change in serum marker score (at the mid-point and end of the study period) was explored. Individual changes in ELF score observed at the mid-point of the study period significantly correlated *(p=0.007)* with individual changes of histological fibrosis score seen at the end of the study period. However, a significant association was not observed between the change in ELF score occurring prior to, and at the end of the study period with the change in Ishak fibrosis score

occurring at the end of the study. The lack of a significant association between changes in serum markers at the end of the study and histological evolution at the end of the study may be explained by the observation that the end of study ELF scores were seen to regress thus reducing the power of the study to detect a significant association at this time point.

An 'ELF regression model' comprising baseline ELF and changes of ELF at the mid-point of the study can predict the both the ELF score and change of ELF score at the end of the study

Both baseline ELF score *(r=-0.249, p-value=0.038)* and changes in ELF score arising at the mid-point of therapy *(r=0.599, p-value<10^{-7})* were significantly associated with the change of ELF score at the end of the study period. An 'ELF regression model' comprised of both baseline ELF score and the change of ELF score at month 12 were able to predict the change of ELF score at month 24 *(R=0.626, R^2=0.392, p-value<10^{-8})* and the end of study ELF score *(R=0.781, R^2=0.609, p-value<10^{-11})* [figure 9.4]. The change of ELF score at the end of the study is described by the equation: -5.786 - 0.762(ELF_0) + 0.665($\Delta ELF_{0\text{-}12}$).

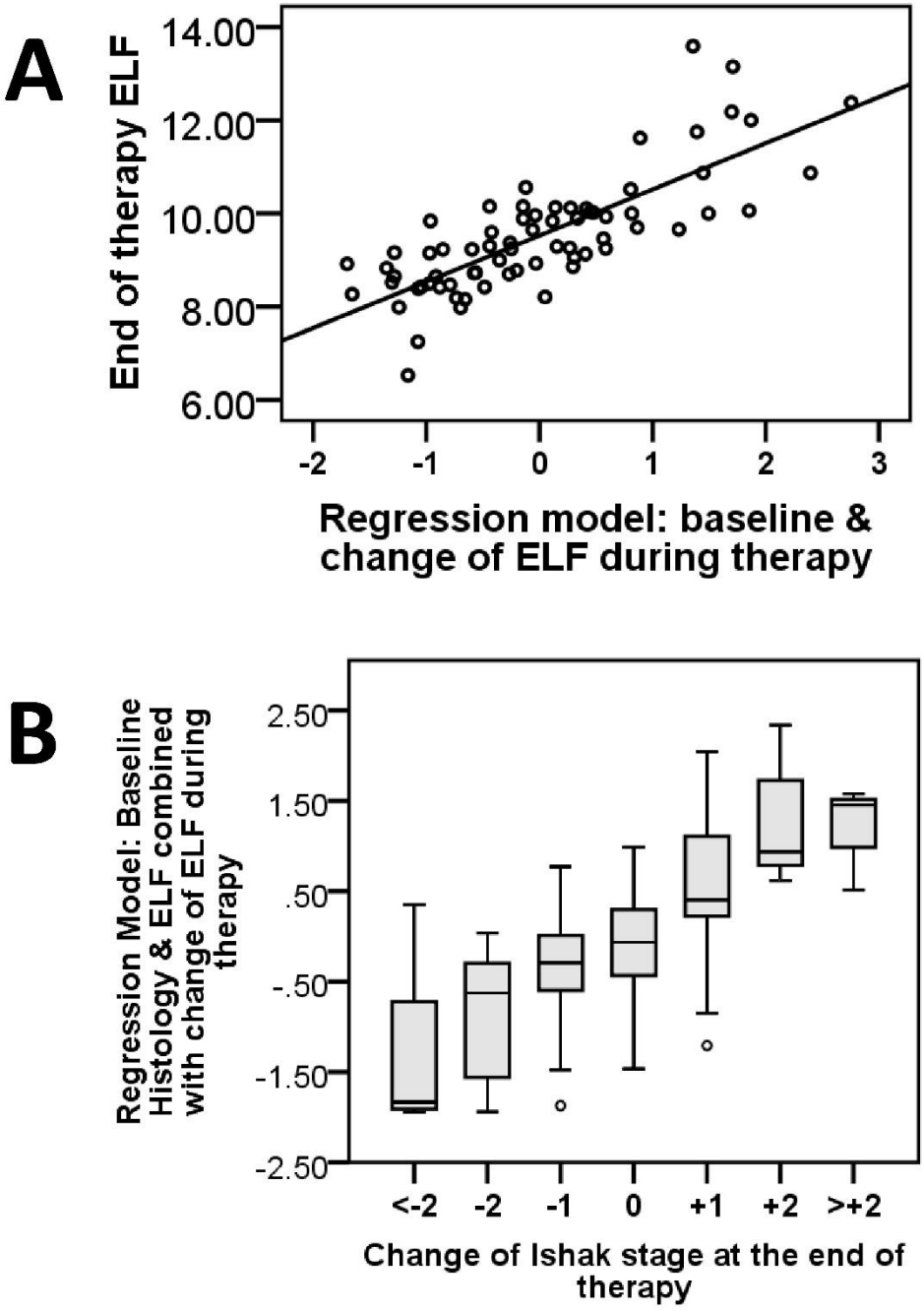


Figure 9.4 models predicting fibrosis evolution at the end of the study period

Panel A: a model combining baseline ELF and change of ELF at the mid-point of the study predicts the end of study ELF.

Panel B: a model combining baseline Ishak stage with baseline ELF and change of ELF at the mid-point of the study predicts the end of study change of Ishak fibrosis stage.

Performance of the 'ELF regression model' in predicting change in ELF

The performance of the 'ELF regression model' in predicting a change of ELF at the end of the study is presented in table 9.4. The performance of the regression model ranged from 0.80 to 0.85 and 0.81 to 0.85 in its ability to predict a fall and rise in ELF score respectively.

Table 9.4 Performance of 'ELF' regression model (ELF0 + ∆ELF0-12) at predicting a change of ELF score (∆ELF0-24) at the end of the study period

Evolution of ELF (n v n)	**AUROC**	**95% Confidence interval**	**Standard Error**	***P-value***
∆ELF$_{0-24}$: Decrease of ELF Score by ≥ 1.5 points (7 v 63)	0.85	*0.70-0.99*	0.074	*0.003*
∆ELF$_{0-24}$: Decrease of ELF Score by ≥ 1 point (9 v 61)	0.80	*0.69-0.81*	0.065	*0.029*
∆ELF$_{0-24}$: Decrease of ELF Score by ≥ 0.5 points (18 v 52)	0.82	*0.72-0.91*	0.054	*<0.001*
∆ELF$_{0-24}$: Increase of ELF Score by ≥ 0.5 points (29 v 41)	0.81	*0.69-0.91*	0.055	*<0.001*
∆ELF$_{0-24}$: Increase of ELF Score by ≥ 1 point (14 v 56)	0.82	*0.71-0.93*	0.057	*<0.001*
∆ELF$_{0-24}$: Increase of ELF Score by ≥ 1.5 points (7 v 67)	0.85	*0.70-0.99*	0.074	*0.003*

Table 9.5 Performance of 'histologic' regression model (Ishak0 + ELF0 + ∆ELF0-12) in predicting both histologic fibrosis evolution (∆Ishak0-24) and the histologic stage (Ishak24) at the end of the study period.

Evolution of Histology (n Vs. n)	**AUROC**	**95% Confidence interval**	**Standard Error**	***P-value***
$Ishak_{0-24}$: Decrease Ishak score ≥ 3 Stages (4 Vs. 66)	0.83	0.56-1.00	0.136	*0.028*
$Ishak_{0-24}$: Decrease Ishak score ≥ 2 Stages (11 Vs. 59)	0.84	0.72-0.97	0.065	*<0.001*
$Ishak_{0-24}$: Decrease Ishak score ≥ 1 Stages (21 Vs. 49)	0.81	0.70-0.91	0.054	*<0.001*
$Ishak_{0-24}$: Increase Ishak Score ≥ 1 Stages (24 Vs. 46)	0.86	0.76-0.95	0.050	*<0.001*
$Ishak_{0-24}$: Increase Ishak Score ≥ 2 Stages (8 Vs. 62)	0.91	0.84-0.98	0.037	*<0.001*
$Ishak_{0-24}$: Increase Ishak Score ≥ 3 Stages (3 Vs. 67)	0.88	0.75-1.00	0.063	*0.029*
$Ishak_{24}$: F0-1 Vs. F2-6 (23 Vs. 47)	0.88	0.79-0.98	0.049	*<0.001*
$Ishak_{24}$: F0-2 Vs. F3-6 (31 Vs. 39)	0.88	0.80-0.96	0.041	*<0.001*
$Ishak_{24}$: F0-3 Vs. F4-6 (48 Vs. 22)	0.91	0.84-0.97	0.035	*<0.001*
$Ishak_{24}$: F0-4 Vs. F5-6 (55 Vs. 15)	0.92	0.84-0.98	0.041	*<0.001*

A 'histologic regression model' incorporating baseline histology and on treatment changes of serum markers can predict the end of therapy histology and histologic change

Stepwise linear regression incorporating baseline serum marker scores was used to develop models able to predict histological outcome as assessed by a change in the semi-continuous Ishak stage arising at the end of the study period. Models incorporating both the baseline and subsequent changes in both ELF score and its constituent components did not result in a significant improvement in the univariate correlation coefficients already identified. However, the incorporation of baseline histology produced models with significantly improved performance. The best performing 'histologic regression model' combined baseline Ishak fibrosis score, baseline ELF score together with the change of ELF score arising at the mid-point of therapy *(R= 0.645, R^2 0.416, P-value = 9.2 x 10^{-8})* (figure 9.4). The 'histologic regression model' score is described by the equation: -5.786 - 0.601($Ishak_0$) + 0.762(ELF_0) + 0.665($\Delta ELF_{0\text{-}12}$). Similarly, the end of therapy Ishak fibrosis stage was best predicted by a model combining baseline Ishak fibrosis score, baseline ELF score together with the change of ELF score arising at the mid-point of therapy *(R=0.775, R^2=0.601, P-value<1 x 10^{-13}).*

Performance of the 'histologic regression model' in predicting histologic change and end of treatment histology

The performance of the 'histologic regression model' in predicting histologic fibrosis evolution and the histologic fibrosis stage at the end of the study period was evaluated (table 9.5). The AUROC for predicting fibrosis regression of greater than 1, 2 and 3 Ishak stages at the end of the study period was 0.81, 0.84 and 0.83 respectively. The AUROC for predicting fibrosis progression of greater than 1, 2 and 3 Ishak stages at

the end of the study period was 0.86, 0.91 and 0.88 respectively. Furthermore, the 'histologic regression model' performed well in its ability to discriminate between the end of study fibrosis stages (AUROC 0.88-0.92, table 9.5).

Examples of how the 'histologic regression model' can be used predict fibrosis change at the end of therapy

A 'histologic regression model' threshold of +0.53 could be used to 'rule in' failure of anti-fibrotic therapy (a rise of more than 2 Ishak fibrosis stage at the end of the study period, sensitivity of 88%, specificity of 83%, diagnostic odds ratio of 34.2 and positive likelihood ratio 5.1) and 'rule out' successful anti-fibrotic therapy (a fall of more than 2 Ishak stages at the end of the study period, sensitivity 100%, specificity 40%, negative likelihood ratio 0.0). As a result, patients with a 'histologic' threshold of greater than 0.53 at the midpoint of therapy could be considered unlikely to achieve significant fibrosis regression and could therefore stop therapy early by meeting a 'futility rule'.

Conversely, a 'histologic regression model' threshold of -0.63 could be used to 'rule out' failure of anti-fibrotic therapy (a rise of more than 2 Ishak fibrosis stage at the end of the study period, sensitivity of 100%, specificity of 30%, negative likelihood ratio 0.0) and 'rule in' successful anti-fibrotic therapy (a fall of more than 2 Ishak stages at the end of the study period, sensitivity 64%, positive likelihood ratio 4.3, diagnostic odds ratio 10.1). A 'histologic threshold' of less than -0.63 at the midpoint of therapy could be applied as a 'continuation rule' with patients scoring less than -0.63 considered likely to achieve fibrosis regression.

DISCUSSION

This was the first study that evaluated the relationship between changes in serum markers occurring during a clinical trial with changes of fibrosis stage at the end of the trial period[377]. As a result, these data provided an invaluable insight into the dynamic interaction and association between changes in fibrosis and the levels of a panel of direct markers which most plausibly may reflect the dynamics of liver fibrogenesis and fibrolysis. [378] Studies had previously compared histological change occurring after putative anti-fibrotic therapy with changes in direct and indirect serum markers at the end of therapy but not during therapy.[379-381] Previous studies had also monitored changes in both direct and indirect serum markers occurring during putative anti-fibrotic therapy but did not have sequential histology as an end point.[382-386]

In this study, patients with lower pre-treatment ELF scores demonstrated greater increases in ELF score at the end of the study period. In addition, it was observed that changes of ELF occurring at the mid-point of the study period were significantly associated with fibrosis evolution at the end of the study.

The 'ELF regression model' which combined these two observations (baseline ELF score and a change of ELF score at the mid-point of the study) performed well at identifying improvement in fibrosis as defined by a fall in ELF score at the end of therapy (AUROC 0.80-0.85) and worsening of fibrosis as defined by a rise in ELF score at the end of therapy (AUROC 0.81-0.85). Furthermore, the 'histologic regression model' which combined these two observations with baseline histologic fibrosis stage performed well at identifying improved fibrosis (AUROC 0.81-0.84) and worsened fibrosis (AUROC 0.86-0.91) as assessed by histological change.

The purely non-invasive ‘ELF regression model’ combining only baseline and changes of ELF score at the mid-point of therapy performed less well at predicting histological fibrosis evolution. For the prediction of histologic fibrosis regression, the performance of this model ranged from AUROC of 0.64 for predicting the regression of at least 1 Ishak stage to AUROC of 0.75 predicting the regression of at least 3 Ishak stages. However, for the prediction of histologic fibrosis progression, the performance of this ‘non-invasive’ model was more comparable to the model incorporating baseline histological stage with AUROC of 0.72 for predicting the progression of at least 1 Ishak stage to AUROC of 0.85 for predicting the progression of at least 3 Ishak stages. Regardless, the non-invasive ‘ELF regression model’ (based on baseline and on treatment changes of ELF) performed well in its ability to discriminate between fibrosis stages at the end of the study (AUROC 0.83-0.91).

Whilst the sample size in this study is modest, the challenge of obtaining paired liver biopsies from patients with serial blood samples over 24 months is difficult, in part due to recognition of the accuracy of non-invasive methods for liver fibrosis staging. The study population represents a subgroup of patients enrolled in a randomized controlled trial that attempted to investigate the anti-fibrotic properties of interferon and or silymarin. The trial found that none of the therapeutic regimes were superior to any of the others studied in delivering a significant benefit in terms of liver fibrosis. Due to the lack of a control arm (in which patients did not receive therapy) the study did not permit exploration of the correlation between any potential anti-fibrotic effect attributable to interferon-based therapy. This however has already been addressed by larger randomised controlled trials which have collectively suggested that interferon is not superior to placebo in preventing histological progression when sustained viral response is not achieved.[151, 387] Nevertheless, in this study, two independent methods

of assessing liver fibrosis, non-invasive serum markers and liver histology, have documented significant changes in liver fibrosis. Up to 70% of the patients in the study experienced either no progression or regression of fibrosis as assessed by either methodology suggesting that the treatment may have had some anti-fibrotic effect. However, these histologic findings are similar to a previous study of 219 untreated patients with CHC who had interval biopsies after a median interval of 2.5 years.[388] In this study 33% and 10% of patients showed progression and regression of liver fibrosis by at least 1 Ishak stage, respectively. Regardless, when considering a single modality to assess liver fibrosis, changes in either histology or serum markers at the end of the study period could be attributed to confounding factors such as sampling error. However, within this study it has been observed that changes in these two independent modalities of assessing liver fibrosis are significantly associated with one another. Thus, the correlation between changes in ELF and changes in histology is likely to be attributable to a true biological association between these two methods of assessing liver fibrosis rather than the result of confounding or a random association. The correlation of change in ELF with improvement of histological fibrosis and the corresponding correlation of increase in ELF with progression of fibrosis provides evidence that the ELF test can be used to assess longitudinal changes of fibrosis. This could be in the context of therapeutic trials of drugs that lead to improvements in liver fibrosis either as a result of successful treatment of underlying pathology (such as clearance of HCV or HBV infection) or due to a direct anti-fibrotic effect, as was anticipated in the PROFI-C trial. Furthermore, the ability to predict longer term changes in fibrosis by the use of on-treatment changes in serum markers would enable clinicians to employ ‘futility rules’ for such therapies by the use of clinically relevant thresholds. Patients meeting ‘futility rules’ based upon relevant diagnostic thresholds

would be able to discontinue therapy early thus avoiding unnecessary exposure to agents.

Nevertheless, due to the introduction of DAA therapies for CHC the concept of maintenance interferon has now been surpassed by treatment regimens that offer SVR for the overwhelming majority of patients with increasingly short treatment durations. Whereas interferon based antiviral therapy is now more relevant for the treatment of hepatitis B infection, in this study it has been demonstrated that the performance of direct serum markers in detecting fibrosis remains consistent after a course of interferon based therapy and on treatment changes can be used to predict fibrosis evolution occurring as a result of therapy. Thus, patients can be stratified into those who have and have not developed severe fibrosis or cirrhosis and screening for hepatoma and endoscopic features of portal hypertension can be instigated. Furthermore, it is anticipated that improvements in hepatic fibrosis will result in improved clinical outcomes in patients with advanced liver disease. Primarily, a reduction in ELF score suggests an improvement in hepatic fibrosis. However, as ELF scores have been shown to predict clinical outcomes more reliably than histology[389], a rise or fall in ELF score resulting from therapy appears to denote an increase or decrease in the risk of subsequent liver related outcomes.

Whether or not these observations are restricted to interferon alpha or silymarin based treatment or are applicable to any anti-fibrotic therapy in general remains to be tested. Interestingly, these observations were consistent regardless of whether or not patients received pegylated interferon with or without silymarin in this study. A rise in ELF score during interferon therapy was noted in this study regardless of the evolution of fibrosis at the end of therapy. This has been previously noted by other investigators and has been attributed to the effect of interferon on extrahepatic serum markers of fibrosis.[384]

However, the lower pegylated interferon dosage of 50 mcg at alternate weeks is indeed considerably lower than what that was normally employed for the treatment of CHC. It is therefore unclear whether the changes in serum markers that were observed in this study are due to interferon itself but instead due to the activation of fibrogenic pathways which result in a rise in direct serum markers.

In summary, this study has demonstrated that baseline and on-treatment changes of the ELF score are significantly associated with fibrosis evolution during a 24 month period of observation. A model combining these parameters was highly predictive of changes in ELF score over a longer period. In addition, when combined in a model with baseline histology, these parameters were highly predictive of histologic fibrosis evolution over a longer period. If confirmed in other cohorts using anti-fibrotic agents, this may qualify ELF as a dynamic marker panel of fibrosis evolution rather than simply a parameter of cross-sectional fibrosis staging; this is highly relevant for the clinical testing of anti-fibrotic agents.[390, 391] It would also permit the refinement of “response-guided therapy” by identifying those patients who will benefit from both continued and prolonged anti-fibrotic treatment.

CHAPTER 10

RANDOMISED CLINICAL TRIAL INVESTIGATING THE SAFETY AND EFFECTIVENESS OF AN ESCALATING DOSE OF PEGINTERFERON ALFA-2A MONOTHERAPY FOR 48 WEEKS COMPARED WITH STANDARD CLINICAL CARE IN PATIENTS WITH HCV CIRRHOSIS

ABSTRACT

Background: A substantial proportion of patients with chronic hepatitis C develop progressive liver fibrosis which culminates in cirrhosis and its complications. Despite evidence that interferon-α exerts an anti-fibrotic effect that is independent of viral eradication this effect was not demonstrated in clinical trials in which low-dose maintenance interferon was administered to patients with compensated CHC cirrhosis who remained viraemic following a lead-in phase of interferon.

Aim: to study the effect on clinical outcomes and indirect biomarkers of fibrosis of escalating dose pegylated interferon alfa-2a as compared with standard clinical care and amongst patients with CHC and Child A or B cirrhosis.

Methods: in a prospective study, 40 patients with CHC and Child A or B cirrhosis were randomised to receive either standard clinical care (without antiviral therapy) or 48 weeks treatment with pegylated interferon alfa-2a starting at 90mcg escalating to 180mcg weekly if tolerated. Patients thereafter were followed for a mean duration of 41 months. The primary outcome variables were liver related death (LRD); all-cause mortality and sustained virological response. Secondary outcomes were “liver-related events” (LRE) and health-related quality of life (HRQL). Indirect markers of fibrosis (FIB4 and APRI) were calculated for both groups prior to and during the study period.

Results: both groups were well matched with treatment well tolerated. The incidence of all-cause mortality *(p=0.024)* and non-oncological liver morbidity *(p=0.04)*, were significantly higher in the control arm after a mean of 47 months follow up.

Improvements in biomarker fibrosis scores were identified in the patients receiving pegylated interferon.

Conclusion: 48 week escalating dose pegylated interferon2a was associated with a significant reduction in all cause mortality, non-oncological liver related morbidity and indirect fibrosis biomarker scores in this trial. These findings suggest that not only does pegylated interferon alpha 2 a have an antifibrotic effect independent of viral clearance but that changes in indirect serum marker scores could be used to monitor and predict this effect.

ID: ISRCTN88008850

INTRODUCTION

Following on from the molecular cloning of HCV in 1989[392], pegylated interferon (pegylated interferon alfa 2a or pegylated interferon alfa 2b) formed the spine of antiviral therapy for CHC until the development of DAA therapy.[393, 394] In contrast to current DAA therapy which is highly efficacious, abbreviated, and tolerable, interferon bases therapies resulted in SVR in approximately 50% of patients.[395, 396] Retreatment strategies with interferon and ribavirin including dose optimisation and an extended duration of treatment were in the main, unsuccessful in achieving viral clearance.[387]

Regardless, benefits were observed with interferon-based therapy even when therapy failed to achieve viral clearance. Both reduction in portal pressure[146] and significant histological improvement with interferon based antiviral therapy were observed in the absence of achieving SVR.[147] These observations were supported by studies in animal models where type 1 interferons have been shown to have a direct antifibrotic effect.[148, 149] These basic science observations were translated into the three large clinical trials that addressed the use of long term interferon in patients with advanced fibrosis.[150-152]

These trials showed that low maintenance doses of pegylated interferon did not improve outcomes in patients with compensated cirrhosis who had not responded to a lead-in phase of interferon. However, the value of full dose pegylated interferon in patients with more advanced stages of cirrhosis was not investigated. These factors may have resulted in an underestimation of the impact of pegylated interferon in this setting.

In this study the use of PEG-INFα2a for patients with CHC Child A and B Cirrhosis, escalating to full dose for 48 weeks without a lead-in treatment phase was investigated

as a putative antifibrotic agent with response to therapy assessed by liver related events and indirect of serum marker scores.

METHODS

PACIFIC (PEGasys in Cirrhosis Study) was a randomized controlled trial of 48 weeks of escalating dose PEG-INFα2a monotherapy compared to standard clinical care (no further antiviral therapy), in patients with CHC and Child A or B cirrhosis. The patients had all either failed previous antiviral treatment with non-pegylated or pegylated interferon (with or without ribavirin therapy) or were treatment naive who were felt unable to tolerate the conventional dosing schedule of pegylated interferon and ribavirin due to advanced liver disease (thrombocytopaenia, anaemia or ascites).

The study was conducted at 5 clinical centres in the United Kingdom (Southampton, London, Plymouth, Nottingham and Glasgow). Randomisation was conducted by telephone from a central data coordinating centre in Southampton. Blood testing was performed locally and virological testing was supplemented at a single laboratory (iQUR Limited Southampton). Blood test results were used to calculate APRI and FIB4 scores as surrogate marker scores of liver fibrosis by published formulae (full description is found in the appendix)[231, 397]

40 patients were recruited between October 2004 and March 2008. Further recruitment was halted by the Data and Safety Monitoring Committee (DSMC) on publication of HALT-C and EPIC trial results.

Patient population

Patients eligible for the study were male or female aged over 18 years old with both serologic evidence of hepatitis C infection by an anti-HCV antibody test and quantifiable HCV RNA (>50 IU /mL Roche TaqMan 48 kit) and were classified as having either Child-Pugh A or B cirrhosis. Patients were considered to be eligible for pegylated interferon monotherapy if they had previously failed antiviral therapy, were

unwilling to take ribavirin or were judged to be unable to tolerate ribavirin in conjunction with pegylated interferon. Evidence of cirrhosis was based upon liver biopsy as judged by a local pathologist (METAVIR F4); or on cross-sectional imaging with features of portal hypertension. Concurrent hepatocellular carcinoma was excluded by imaging and alpha-fetoprotein (AFP) monitoring for 3 months prior to entry. Patients were excluded if they had other forms of liver disease, human immunodeficiency virus (HIV) infection, hepatitis B infection, pre-existing severe depression or other conditions preventing the use of interferon including psychiatric disease, cardiac disease, renal disease, seizure disorders, pregnancy or severe retinopathy. Laboratory values that excluded patients from enrolment included neutrophil count <1000 cells/mm3, or platelet count <60,000 cells/mm3.

STUDY DESIGN AND SUBSEQUENT FOLLOW UP

Patients were randomised to standard clinical care (no further antiviral therapy) or 48 weeks treatment with PEG-INFα2a at a dose of 90mcg per week escalating to 135 mcg per week for 4 weeks and thereafter to 180mcg per week if tolerated. Randomisation was stratified based on HCV genotype Patients were subsequently assessed at 12-weekly intervals for the next 48 weeks and then resumed standard clinical care. Standard clinical care included regular upper GI endoscopy to evaluate endoscopic features of portal hypertension and six monthly clinical assessments for symptoms and signs of liver disease, with routine blood tests, ultrasound scanning and AFP measurement as surveillance for hepatoma (figure 10.1).

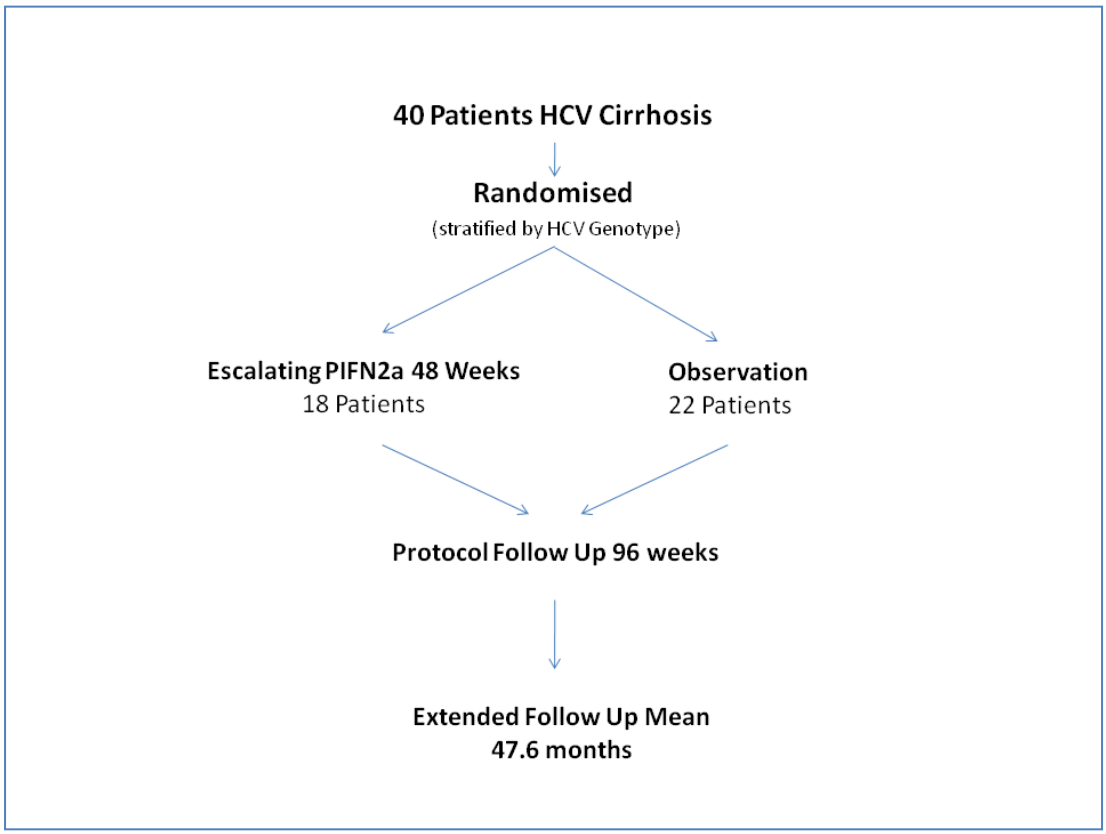


Figure 10.1 Enrolment, Randomization, and Follow-up of Study Participants.

Whilst on therapy, patients were evaluated at monthly intervals to monitor for side effects and changes in serum liver chemistries, complete blood count, serum HCV RNA level, APRI and FIB4. Quality of life assessments were conducted by measuring the validated SF-36 and fatigue severity scale score at baseline and every 12 weeks thereafter for 96 weeks.

Following on from the end of the 96 week study period all patients resumed standard clinical care. This comprised regular clinical assessments at a maximum interval of 6 months, including gastroscopy, ultrasound and AFP. At study closure, patient medical records and clinic reports were reviewed at the 5 sites.

Outcomes

The primary outcome variables were SVR, liver related death and all-cause mortality. Secondary outcomes were “liver related events” including variceal haemorrhage, new development of ascites and spontaneous bacterial peritonitis (SBP), hepatocellular cancer, health related quality of life (HRLQ) and changes in indirect biomarkers of fibrosis.

Definitions of Response

SVR was defined as the absence of detectable HCV RNA (<50 IU /mL Roche TaqMan 48 kit) in serum 24 weeks after PEG-INFα2a. End-of-treatment (EOT) virologic response was defined as absence of detectable HCV RNA in serum at treatment week 48. SVR was defined as absence of detectable serum HCV RNA at week 72, at least 24 weeks after treatment was discontinued.

Dose Modification

All patients randomized to treatment were commenced on 90mcg per week of PEG-INFα2a for 4 weeks. Their dosage was then escalated to 135mcg per week for 4 weeks, and thereafter to 180mcg per week thereafter according to tolerability. Poor tolerability was assessed by the presence of laboratory abnormalities including anaemia and leucopaenia or by the occurrence of adverse events deemed related to interferon by the principle investigator. Pegylated interferon alfa 2a dosage adjustments in response to biochemical abnormalities were uniform across centres decreasing at 45mcg intervals until the medication was withheld.

Planned recruitment and statistical analysis

The primary outcome measures for the study were sustained viral response to treatment as judged by negative HCV PCR 24 weeks after cessation of therapy; death; liver transplantation; and hepatocellular cancer. Secondary outcome measures were end of treatment response judged by negative HCV PCR at the end of treatment; liver related events (LRE) including: the detection of bleeding varices; development of ascites; spontaneous bacterial peritonitis; hepatic encephalopathy; and quality of life as measured using the SF36 and Fatigue Severity Scale.

For the primary outcome variable of SVR, the planned recruitment for this pilot study was for 45 patients in each treatment arm. As a superiority study, this would enable a power of 90% to detect a clinically significant difference (p=0.05) in SVR rates between the 2 groups of 20%.

For the both the primary outcome variables of liver related death; and all cause mortality and the secondary outcomes of LRE with full recruitment the study would be powered to have an 80% chance of detecting a significant (p=0.05) difference in the

frequency of stated events assuming a 25% and 5% rate of events in the control and treatment arms respectively.

Data were analysed using SPSS v.18.0 (SPSS Inc., Chicago, IL). Analyses were conducted on an intention to treat (ITT) basis. Differences in outcomes between the IFN treated and the control group were assessed using a 2-sided Fisher's Exact test. Differences in continuous data between the PEG-INFα2a treated and the control group were assessed by an independent samples t-test and independent samples Mann-Whitney test for parametric and non-parametric data respectively. Stepwise binary logistic regression was used to predict which factors were associated with clinical outcomes.

RESULTS

All patients randomised had Child-Pugh A or B cirrhosis and were HCV RNA positive. Except for 1 patient in the control arm, all patients had been free from interferon therapy for at least 12 months prior to entry. Both groups were well matched in terms of demographics, severity of liver disease and health related quality of life with no significant differences between the treatment and control groups (tables 10.1 and 10.2). Eighteen patients were randomised to treatment with PEG-INFα2a and 22 randomised to the control arm. 17 of the 18 patients randomised to PEG-INFα2a received treatment as 1 patient withdrew from follow up prior to receiving interferon.

Escalating dose PEG-INFα2a was generally well tolerated in patients with cirrhosis. Two patients (11%) were unable to tolerate 48 weeks of treatment. One patient had PEG-INFα2a withdrawn at 2 weeks due to pancytopaenia (Hb 9.8 g/dl, neutrophils 800 cells/mm3, platelets 47,000 cells /mm3). A further patient had pegylated interferon alfa 2a withdrawn at 4 weeks due to thrombocytopaenia (platelets 27,000 cells/mm^3). The patient with pancytopaenia was treatment naive. Four patients (22%) required a dose reduction of PEG-INFα2a due to thrombocytopaenia (average weekly dose 159 mcg). Seven patients (39%) achieved viral suppression during the treatment period and 3 patients (16%) in the treatment group achieved a SVR (1 previous non-responder, 2 previous relapsed-responders to pegylated interferon and Ribavirin).

Table 10.1 Baseline Demographic, Biochemical and Liver indices of the Patients

Variable	Treatment Group (n=18)	Control Group (n=22)	*P Value*
Mean Age (yr)	54.9 ± 8.5	52.1 ± 8.5	*NS*
Female Sex (%)	5 (27.8%)	5 (22.7%)	*NS*
Mean Weight (kg)	80.4 ± 12.5	81.6 ± 15.1	*NS*
Race or Ethnic group			
• **White**	16 (94%)	19 (86%)	
• **Black**	0	1 (5%)	
• **South Asian**	1 (6%)	2 (9%)	
Mean Alcohol Consumption (units/wk)	1.7 ± 3.7	3.7 ± 8.6	*NS*
Genotype			*NS*
• **1**	9 (50.0%)	12 (54.9%)	
• **2,3**	8 (44.4%)	9 (40.9%)	
• **4,5,6**	1 (5.6%)	1 (4.5%)	
Prior treatment status			
• **Non responder**	9 (50%)	11 (50%)	
• **Virological Breakthrough**	0	2 (9%)	
• **Relapsed Responder**	7 (39%)	9 (41%)	
• **Treatment naïve**	2 (11%)	0	
Diagnosis of cirrhosis			
• **Histological**	12 (66%)	14 (64%)	
• **Imaging**	6 (33%)	8 (36%)	
Serum Albumin	39.7 ± 3.5	39.5 ± 5.9	*NS*
Serum ALT (IU/ml)	84.7 ± 47.2	107 ± 77.8	*NS*
Serum Bilirubin	16.2 ± 8.1	16.7 ± 7.4	*NS*
Serum Creatinine	86.5	80.5	*NS*
Platelets	137.1 ± 64.8	138.2 ± 51.9	*NS*
Child-Pugh Mean Score	5.4 ± 0.6	5.3 ± 0.7	*NS*
• **A (%)**	16 (88%)	19 (86%)	
• **B (%)**	2 (12%)	3 (14%)	
MELD	8.2	8.0	*NS*

Parametric data are displayed as mean ± standard deviation with significance testing using 2-sided independent samples t-test. Non-parametric data are displayed as median values with significance testing using a 2 sided Mann-Whitney independent samples test.

Table 10.2 Baseline quality of life indices for the patients

The 8 components of the SF-36 and the 2 components of the fatigue severity scale are displayed as mean ± standard deviation with significance testing using 2-sided independent samples t-test

Variable	Treatment	Control	*P value*
SF-36			
Physical Functioning	57.0 ± 33.3	64.7 ± 28.9	*NS*
Physical role limitation	55.0 ± 50.2	51.5 ± 44.6	*NS*
Emotional Role Limitation	60.0 ± 50.7	51.0 ± 48.8	*NS*
Energy / Fatigue	45.0 ± 25.8	43.5 ± 27.1	*NS*
Emotional Well being	62.9 ± 20.5	61.4 ± 25.4	*NS*
Social Functioning	70.0 ± 33.0	58.1 ± 32.5	*NS*
Pain	67.5 ± 29.3	58.1 ± 26.7	*NS*
General Health	38.3 ± 20.2	37.7 ± 23.9	*NS*
Fatigue Severity Scale			
Mean Fatigue severity score	5.2 ± 1.9	5.4 ± 1.9	*NS*
Fatigue Index	4.9 ± 2.8	5.5 ± 2.6	*NS*

Events during trial duration

Within the 96 week trial duration, 1 LRE and no deaths were observed in the treatment group compared to 4 LRE *(p=0.211),* 3 liver related deaths *(p=0.096)* and 4 all cause deaths *(p=0.051)* amongst the controls (table 8.3). The LRE in the treatment group was a hepatocellular carcinoma (oncological event). Of the LRE in the controls all were non-oncological events and each was attributable to portal hypertension (1 episode of decompensated cirrhosis and 2 episodes of variceal bleeding).

Table 10.3 Clinical and Virological Outcomes at week 96

Death (all cause or liver related), liver events (all or non-oncologic) and virologic outcomes (SVR, relapsed response and non-response) are compared between the treatment and control groups:differences in outcomes between the IFN treated and the control group are assessed using a 2-sided Fisher's Exact test.

Variable	Treatment (n=18)	Control (n=22)	*P Value*
All Cause Death	0	4 (18.2%) *b,b,d,m*	*0.114*
Liver Related Death	0	3 (13.6%) *b,b,d*	*0.238*
Liver Event	1 (5.6%) *c*	4 (18.2%) *b,b,b,d*	*0.355*
Non-oncological Liver Event	0	4 (18.2%) *b,b,b,d*	*0.114*
SVR	3 (17%)	0	*0.165*
Relapsed Response	4	0	
Non-response	11	22	

Key:
a= SBP, b=Variceal Bleed, c= HCC, d=Decompensated Cirrhosis, m=Myocardial Infarction

Subsequent Follow up

Thereafter, patients resumed normal clinical care. All patients continued to be followed in the hepatology outpatient clinic in the absence of a clinical event. This was comprised of 6 monthly clinical assessment including ultrasound and AFP.

The duration of subsequent follow up was similar *(p=0.311)* for both the treatment (mean 47.3 months, median 47 months) and control groups (mean 47.6 months, median 46 months) (table 10.4).

During this period there were 2 LRE and no deaths were observed in the treatment group compared to 6 LRE *(p=0.258),* 5 liver-related deaths *(p=0.053)* and 6 all-cause deaths *(p=0.024)* amongst the controls. Only 1 of the patients who died had Child-Pugh B cirrhosis at the start of the study. Both LRE in the treatment arm were hepatocellular cancers (table 10.5). SVR did not appear to confer benefit in the treatment arm as both patients with hepatocellular cancers had viral suppression during treatment of which one achieved SVR. In the control arm, all LRE were non-oncological.

Table 10.4 Clinical and Virological Outcomes at the end of extended follow

Death (all cause or liver related), liver events (all or non-oncologic) and SVR are compared between the treatment and control groups; differences in outcomes between the IFN treated and the control group are assessed using a 2-sided Fisher's Exact test.

Variable	Treatment (n=18)	Control (n=22)	*P value*
Follow Up (months)	41.3 ± 6.9	41.1 ± 9.1	*0.311*
All Cause Death	0	6 (27.3%) *a,b,b,d,d,m*	*0.024**
Liver Related Death	0	5 (22.7%) a,b,b,d,d	*0.053*
Liver Event	2 (11.1%) c,c	6 (27.3%) a,b,b,b,d,d	*0.258*
Non-oncological Liver Event	0	6 (27.3%) a,b,b,b,d,d	*0.024**
SVR	3 (17%)	0	*0.165*

Key
a= SBP, b=Variceal Bleed, c= HCC, d=Decompensated Cirrhosis, m=Myocardial Infarction.

Indirect marker scores of liver fibrosis

The indirect markers of fibrosis, APRI and FIB4, were calculated prior to, during and after the study period (table 10.5). In the PEG-INFα2a treated group a non-significant fall in mean APRI and FIB4 scores was observed when comparing the scores prior to and 48 weeks after PEG-INFα2a therapy. By contrast in the control arm, when comparing equivalent time points, a non-significant increase in APRI and FIB4 scores was observed. Furthermore, whilst the serum marker scores in the respective PEG-INFα2a and control arms were not significantly different at the start of therapy after 96 weeks a significant difference in the respective scores was seen *($p<0.05$).*

Table 10.5 Indirect Serum Fibrosis Marker Scores in the Control and Interferon treated groups at the start and end of the 96 week follow up period

Marker scores are presented as mean ± standard deviation after being categorised into week 0 and week 96 for both the treatment and control groups. Significance testing between these categories are assessed using 2-sided independent samples t-test.

Time Point	APRI Scores during PACIFIC study			FIB-4 Scores during PACIFIC Study		
	Control Group	PEG-IFN Group	*P value* (Control v PEG-IFN)	Control Group	PEG-IFN Group	*P value* (Control v PEG-IFN)
Start of Study (Week 0)	3.42 ± 2.58	3.36 ± 2.17	NS	4.31 ± 3.66	3.91 ± 2.58	NS
End of Study (Week 96)	5.49 ± 4.38	2.24 ± 1.49	0.014	6.88 ± 6.37	3.42 ± 2.01	0.007
***P-Value* (Week 0 v Week 96)**	NS/3	NS/3		NS	NS	

Safety and quality of life indices

Quality of life indices were obtained on 17 patients in the control arm and 14 patients in the treatment arm (table 10.2). They were not significantly differences at the start of the study. During therapy, however, levels of pain were significantly higher in the treatment arm compared to controls (p=0.001). Adverse events were noted with similar frequency in both the treatment and control arms (82% and 66%%, p=NS). One serious adverse event (SAE) was noted in a patient in the control arm of the study. This was the occurrence of a fatal myocardial infarction.

Binary Logistic Regression

Using multivariate analysis, both not receiving PEG-INFα2a and thrombocytopaenia were the only factors predictive of all cause mortality *(p=0.023 & p=0.07),* and a non-oncological liver event (p=0.023 & p=0.04) at the end of study follow up (table 10.5).

Table 10.6 Multivariate analysis of factors predicting 'all cause mortality' and 'non-oncological liver events' at the end of extended follow up.

Differences in outcomes between the IFN treated and the control group are assessed using a 2-sided Fisher's Exact test.

Variable	All Cause Mortality OR (95%CI)	P-value	Non-oncological liver event OR (95%CI)	*P value*
Treatment with pegylated interferon ∝2a	119.215 (1.939-7328)	0.023	43.542 (1.694-1119)	0.023
Platelets	0.953 (0.921-0.987)	0.07	0.962 (0.936-0.988)	0.04

Factors not significantly associated with 'all cause mortality' and 'non-oncological liver events' after multiple binary logistic regression:

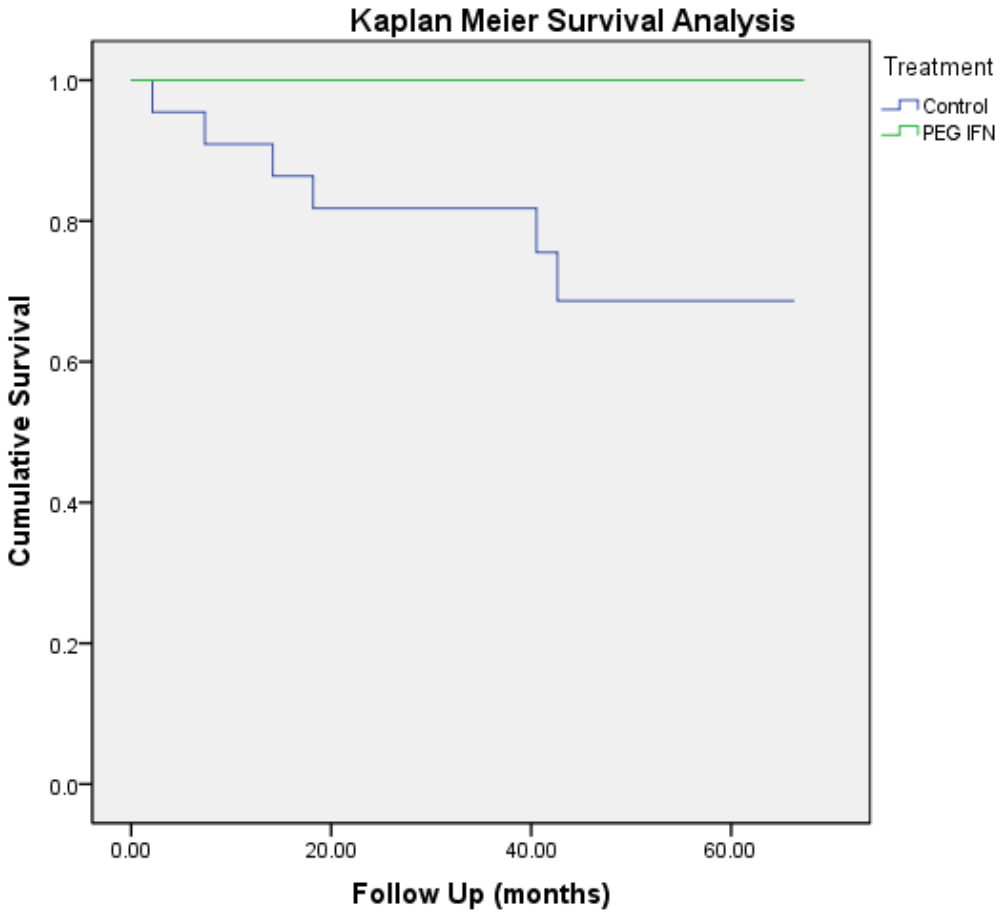
- Age, Gender, Weight, Alcohol consumption, ALT, AST, Bilirubin, ALP Albumin, Creatinine, Serum Sodium, INR, Child Pugh Score, MELD score, HCV RNA

Survival Analysis

A Kaplan-Meier survival curve is displayed in figure 10.2. A significant difference in survival at the end of the study duration between the 2 study groups was identified using the Mantel–Cox test (p=0.017).

Figure 10.2 Kaplan–Meier Survival Analysis

This figure shows the time to death according to group assignment (treatment or control).



DISCUSSION

Prior to the introduction of DAA therapy, patients with cirrhosis due to CHC who had failed antiviral treatment were managed with a 'watch and wait' approach. Decompensation in this group of individuals without successful antiviral treatment stands at a rate of 10% per year[398] with transplantation for HCV infection associated with graft infection and rapid progression to cirrhosis [399]. When this study was conducted, equipoise remained about the antifibrotic effects of interferon for patients with CHC cirrhosis. In this study, a 48 week escalating dose PEG-INFα2a in patients with CHC cirrhosis was associated with a significant reduction in all-cause mortality and non-oncological liver related morbidity after a mean of 47 months extended follow up. Whereas one hypothesises that the benefit from pegylated interferon in this study was derived by its antifibrotic effect and associated reduction in portal pressure, the differences between these findings and those of previous studies may relate to the modest study sample size. However, as these findings are supported by changes in indirect fibrosis biomarker scores during the study period, these differences may instead relate to study methodology

PACIFIC was designed as a pragmatic trial to address the needs of patients with established and advanced cirrhosis. Whilst subgroups of patients in the former trials had advanced fibrosis or cirrhosis, the overwhelming majority had non-cirrhotic fibrosis. Furthermore, those cirrhotic patients who were enrolled had Child-Pugh A status and presumably did not have significant portal hypertension and so represent a milder spectrum of disease than in the present study. It is therefore likely that the benefit from pegylated interferon results from its antifibrotic effect that results in a reduction in portal pressure as seen in animal models. This benefit appears to occur even in the absence of virologic response in this study. Approximately 50% of the

PACIFIC patients randomised to the pegylated interferon arm demonstrated virologic non-response during the treatment period. Moreover, the plausibility of an anti-fibrotic effect in this context is supported by data from the HALT-C trial from a single centre.[400] At this centre, in a sub-study, patients with advanced fibrosis or cirrhosis treated with low dose pegylated interferon were seen to have both histological improvement and reduction in portal pressure. These benefits were observed independently of viral suppression.

A reduction in hepatocellular carcinoma was not seen in the interferon treated group in PACIFIC. 2 patients developed hepatocellular carcinoma during the subsequent follow-up but both had received PEG-INFα2a. These 2 cases of hepatocellular carcinoma occurred in 2 of the 3 patients who achieved SVR in the treatment arm. Both patients were therefore referred for transplantation in the absence of ongoing HCV infection, a situation that greatly increased their prognosis post-transplant.

The changes in indirect scores suggest that these beneficial effects could have been predicted. Whereas mean indirect serum marker scores were noted to reduce in the group receiving escalating dose PEG-INFα2a, this effect was not identified in all patients. As identified in chapter 9 response guided treatment would allow patients with an ongoing antifibrotic effect to continue treatment whereas patients meeting futility rules would stop treatment. It is therefore unfortunate that patient samples were not also tested for ELF as this would have proved a useful cohort on which to validate the findings that were derived in chapter 9.

In summary, the results of PACIFIC are compatible with PEG-INFα2a exerting an anti-fibrotic effect in cirrhotic patients with CHC that predominantly manifests in a reduction in the complications of portal hypertension with this hypothesis supported by changes in direct serum markers during the study. The study design and patient population of

PACIFIC are sufficiently different from earlier studies that the differences and similarities in outcomes can be explained.

CHAPTER 11

SUMMARY OF FINDINGS
&
DISCUSSION

SUMMARY OF FINDINGS

Chapter 4: DERIVATION AND VALIDATION OF A BIOMARKER FOR THE DETECTION AND ASSESSMENT OF NON-ALCOHOLIC STEATOHEPATITIS IN PATIENTS WITH NAFLD

- 17 variables are assessed in derivation and validation cohorts of patients with NAFLD without advanced fibrosis to identify a biomarker of NASH. PIIINP is identified by logistic regression as the only variable associated with a histological diagnosis of NASH in both cohorts.
- PIIINP discriminates between simple steatosis and NASH amongst patients without advanced fibrosis in both a derivation and validation cohort of patients with NAFLD with good diagnostic performance as assessed by AUROC.
- PIIINP levels also correlate with NAS score and its constituent components.
- Amongst patients with NAFLD with all degrees of fibrosis, PIIINP levels define a hierarchy of liver disease in NAFLD starting with simple steatosis, NASH without advanced fibrosis, and finally advanced fibrosis.
- To detect NASH in clinical practice, three diagnostic thresholds of PIIINP (5.2, 7.2 and 11 ng/ml) are proposed.

Chapter 5: FURTHER VALIDATION OF PIIINP OF FOR THE DETECTION AND ASSESSMENT OF NON-ALCOHOLIC STEATOHEPATITIS IN PATIENTS WITH NON-ALCOHOLIC FATTY LIVER DISEASE

- PIIINP is further validated in a cohort of patients with NAFLD both with and without advanced fibrosis.
- PIIINP again demonstrates good discriminatory performance for the detection of NASH and/or advanced fibrosis.
- Good diagnostic performance is demonstrated for PIIINP at the diagnostic thresholds proposed in chapter 4 (5.2, 7.2 and 11 ng/ml).
- By contrast, serum ALT is unable to discriminate between simple steatosis and NASH or advanced fibrosis.

Chapter 6: COMPARISON OF SERUM BIOMARKERS IN DISCRIMINATING BETWEEN FIBROSIS STAGES IN NAFLD: DIRECT SERUM MARKERS ARE MORE ACCURATE THAN INDIRECT SERUM MARKERS

- The diagnostic performance of 9 biomarkers of fibrosis is validated in a cohort of patients with NAFLD.
- Biomarkers are categorised into direct, hybrid and indirect biomarkers. Diagnostic targets include the fibrosis stages of mild, moderate, advanced, cirrhosis and overall discriminatory performance (Obuchowski).
- Detection of mild fibrosis: none of the 9 biomarkers exhibit good performance for the detection of mild fibrosis in NAFLD as defined by histological stage.
- Detection of moderate fibrosis: direct biomarkers (ELF, HA, PIIINP) all exhibit good performance, but indirect biomarkers demonstrate poorer diagnostic performance.
- Detection of advanced fibrosis: all the direct and 2 of the indirect biomarkers (FIB4 and NAFLD Fibrosis Score) demonstrate good performance for the detection of advanced fibrosis.
- Detection of cirrhosis: whereas direct biomarkers can demonstrate excellent performance, indirect biomarkers demonstrate good diagnostic performance.
- Best overall discriminatory performance (Obuchowski): ELF exhibits the best diagnostic performance of the biomarkers tested both overall and at the 4 stages of fibrosis assessed; direct biomarkers exhibit better diagnostic performance than indirect biomarkers.

Chapter 7: SERIAL AND PARALLEL COMBINATIONS OF SERUM BIOMARKERS RESULT IN ENHANCED DIAGNOSTIC PERFORMANCE FOR THE DETECTION OF ADVANCED FIBROSIS (KLEINER F3-4)

- In this cohort, at the 10.51 threshold endorsed by NICE, ELF exhibits good diagnostic performance for detection of advanced fibrosis (diagnostic accuracy 88%) detecting 74% of patients with advanced fibrosis.
- The 3 biomarkers (ELF, NAFLD Fibrosis score and FIB4) with good performance for the detection of advanced fibrosis are selected for use as the constituent components of diagnostic algorithms with the aim of enhancing the detection of advanced fibrosis in NAFLD.
- Parallel and serial strategy algorithms are designed using both an indeterminate and fully assigned approach after defining combined cut-off and single cut-off thresholds: Bayes' theorem is used to develop and estimate biomarker algorithm performance.
- In total, the performance of 21 diagnostic algorithms are evaluated and compared to the performance of ELF at 10.51.
- Overall, the best performing fully assigned strategy evaluated is ELF at 10.51.
- As compared to ELF alone, the serial combination of FIB 4 and ELF has similar diagnostic accuracy while significantly reducing the number of ELF tests required;
- In this cohort, diagnostic accuracy is enhanced by parallel algorithms combining any 2 of the 3 biomarkers (ELF, NAFLD Fibrosis Score and FIB4) but with more cases classified as indeterminate.
- Using these data, diagnostic strategies are proposed in health care systems with full, partial and no access to ELF testing.

Chapter 8: NON-INVASIVE ASSESSMENT OF HEPATIC FIBROSIS IN PRIOR NON-RESPONDERS TO HEPATITIS C VIRUS TREATMENT: A COMPARISON OF TEN BIOMARKERS OF LIVER FIBROSIS

- The diagnostic performance of 10 biomarkers of fibrosis is validated in a cohort of patients with CHC.
- The best performing markers are the virus specific panels, Fibrometer 2G and 3G; ELF and Hepascore were the best performing panels developed for use in all aetiologies of CLD.
- Detection of Cirrhosis: only 2 indirect biomarkers demonstrate excellent performance for the detection of cirrhosis whereas 6 biomarkers (direct or hybrid) demonstrate good performance.
- Detection of Advanced Fibrosis: 7 of the 10 biomarkers tested exhibit good performance for the detection of advanced fibrosis.
- Detection of moderate fibrosis: all the direct and hybrid biomarkers (but none of the indirect biomarkers) exhibit good performance for the detection of moderate fibrosis.
- The overall discriminatory performance of direct and hybrid biomarkers and their performance at diagnostic thresholds was heavily influenced by the choice of assay used to measure HA.

Chapter 9: SERUM MARKERS OF LIVER FIBROSIS: ON TREATMENT CHANGES OF DIRECT MARKERS PREDICT THE OUTCOME OF ANTI-FIBROTIC THERAPY

- ELF tests taken before, during and after putative antifibrotic therapy (treatment with pegylated interferon with or without silymarin) are used to longitudinally monitor fibrosis evolution of a period of 24 months as assessed by liver histology and ELF itself.
- This analysis identifies that both baseline and on-treatment changes of ELF scores are significantly associated with fibrosis evolution during a 24 month period of observation.
- Predicting changes in liver histology: a model combining baseline and on-treatment changes of the ELF with baseline histology can predict histologic fibrosis evolution at the end of treatment.
- Predicting changes in ELF: a model combining baseline and on-treatment changes of the ELF can predict of changes in ELF score at the end of treatment;
- These models could permit the refinement of “response-guided” antifibrotic therapy” by identifying those patients who will benefit from both continued and prolonged anti-fibrotic treatment.

Chapter 10: RANDOMISED CLINICAL TRIAL: A PILOT STUDY INVESTIGATING THE SAFETY AND EFFECTIVENESS OF AN ESCALATING DOSE OF PEGINTERFERON ALFA-2A MONOTHERAPY FOR 48 WEEKS COMPARED WITH STANDARD CLINICAL CARE IN PATIENTS WITH HCV CIRRHOSIS.

- Clinical outcomes and the evolution of indirect biomarker scores are assessed in a randomised controlled trial of patients with CHC comparing putative antifibrotic therapy (48 week escalating dose PEG-INFα2a) with observation only.
- 48 week escalating dose PEG-INFα2a (as compared to observation) is associated with a significant reduction in all-cause mortality, non-oncological liver related morbidity and improvement indirect fibrosis biomarker scores.
- Given these clinical outcomes and associated changes to the indirect serum marker scores it is hypothesised that 48 week escalating dose pegylated interferon alfa 2a in patients with CHC cirrhosis results in an improvement in liver fibrosis and portal hypertension.
- If confirmed in other cohorts, changes in indirect serum marker scores could be used to re-stratify prognosis following antifibrotic therapy.

DISCUSSION

Standard liver function tests are unable to detect and stage liver injury in the population at risk of chronic liver disease. At present, more than a third of the UK population is at risk of chronic liver disease. The reference standard, liver biopsy, is not practicable for use as the primary method for detecting CLD in the population at risk. As a result, chronic liver disease is often detected at an advanced stage when interventions are ineffective. It is therefore essential that strategies are developed to effectively detect liver disease in the population at risk. Imaging modalities that detect CLD are time consuming and labour-intensive rendering them less suitable than serum tests for detecting liver disease particularly in primary care. Given the ubiquity of chronic liver disease, biomarkers as measured through blood tests, are more suited to embracing disease detection in population at risk. This thesis has explored ways in which candidate biomarkers of liver fibrosis and injury can be used to detect alternative diagnostic targets, enhance disease detection and to monitor putative antifibrotic therapies.

NAFLD: Biomarker of NASH

NAFLD, the most common cause of CLD worldwide, results in hepatic complications only in the minority of patients with NAFLD who develop NASH and progressive liver fibrosis. As pharmacologic treatment for NAFLD is, at present investigational, the current focus is on the detection of NASH and/or fibrosis amongst patients with NAFLD. Prior to this thesis being conducted, CK-18 was the only biomarker that had been identified for the detection of NASH but with limited diagnostic performance in validation studies that precluded its use in clinical practice. By contrast, several biomarkers had been identified for the detection of fibrosis in NAFLD but with equipoise remaining over how best to use these biomarkers in clinical practice. In Chapter 4, a candidate biomarker approach was used to identify a test for NASH using 17 independent variables. Logistic regression identified that amongst patients without advanced fibrosis, serum levels of PIIINP discriminated between NASH and advanced fibrosis. Furthermore, amongst patients with all degrees of fibrosis, PIIINP levels represented a hierarchy of liver disease ranging from simple steatosis, NASH (without advanced fibrosis), and finally advanced fibrosis.[401] These findings were then further validated in chapter 5 in which PIIINP once again discriminated between simple steatosis and/or advanced fibrosis and performed well at the proposed diagnostic thresholds. By contrast, serum levels of ALT could not be used to discriminate between patients with and without severe forms of NAFLD.

These findings have subsequently been explored and validated by other independent groups. PIIINP levels, as measured by an alternative ELISA assay targeting a neo-epitope of PIIINP (PRO-C3)[402], was identified in a cohort of 433 patients (derivation n=320, validation n=113)[403] to correlate with both the NAS score ($p<0.0001$) and the presence of advanced fibrosis ($p<0.0001$). PRO-C3 levels were then combined by

logistic regression into a panel consisting of age, BMI, type 2 Diabetes and platelet count to enhance diagnostic performance for the detection of NASH and advanced fibrosis. Interestingly, this assay was also assessed in a recent UK study in which 44 variables were studied in a cohort of 374 patients with NAFLD for their association with NASH. Of these 44 variables, 7 variables including PIIINP itself and PRO-C3 were associated with NASH; a predictive model using these 7 variables identified NASH with good performance (AUROC 0.87). Thereafter, regression analysis identified that only 5 of these variables (PIIINP, AST, Acetyl-HMGB1, CK18 (M30) and PNPLA3 (rs738409) were significantly associated with a diagnosis of NASH after accounting for the remaining variables as confounders. These recent independent validation studies confirm that PIIINP exhibits robust performance for the detection of NASH. Furthermore, these data identify that the diagnostic performance of PIIINP for the detection of NASH can be enhanced by its combination in a biomarker panel with other biomarkers such as CK18. Further studies are now required in a primary care setting to confirm the diagnostic performance of PIIINP in this context. Furthermore, given the inherent limitations of liver biopsy as a modality by which evolution of NASH can be monitored during therapy for NAFLD, the evolution of PIIINP levels would seem an alternative surrogate endpoint.

NAFLD: Biomarker of Fibrosis

In chapter 6, the ability of 9 biomarkers to detect different degrees of fibrosis in NAFLD was assessed[404]. Overall, ELF was the best biomarker tested. Of the indirect biomarkers tested, NAFLD Fibrosis Score and FIB4 were the best biomarkers exhibiting good performance for the detection of both advanced fibrosis and cirrhosis related due to NAFLD. Further analysis identified that none of the biomarkers tested demonstrated good performance for the detection of mild fibrosis. Although ELF itself demonstrated good performance for the detection of moderate fibrosis, this level of performance was inferior to that demonstrated by PIIINP for the detection of NASH. The ability of ELF to detect advanced fibrosis and cirrhosis however, ranged from very good to excellent. Given these findings and those made in chapters 4 and 5, a diagnostic approach that combines PIIINP and ELF to detect NASH and stage fibrosis (advanced fibrosis and cirrhosis) respectively appears logical.

A recent study in which a cohort of 646 patients with NAFLD was followed up over a mean of 20 years identified that fibrosis itself was a stronger predictor of outcome in NAFLD than the presence or absence of NASH *per se.*[405] Patients in this study were biopsied only once at the study outset. Interestingly patients with mild fibrosis had a similar outcome regardless of whether NASH was present (albeit most patients with advanced fibrosis had NASH). These observations are in keeping with the “multiple hit” hypothesis that describes the progression from simple steatosis into NASH after an unspecified and variable interval. The results of this study confirm that the absence of histologic NASH on a biopsy taken at a single instant in time does not indicate a benign prognosis as NASH can develop at a later stage. Moreover, it is likely that NASH and mild fibrosis may regress in response to interventions such as weight loss and/or exercise. Regardless, these studies highlight that patients with NAFLD but

without advanced fibrosis or NASH at their index assessment need to be assessed subsequently for transformation into a more aggressive phenotype of NAFLD.

In chapter 7, diagnostic algorithms were developed to improve the detection of advanced fibrosis due to NAFLD.[406] Given the findings reported in chapter 6, three biomarkers (ELF, NALFD Fibrosis Score and FIB4) were selected as the component biomarkers of the diagnostic algorithms. Both parallel and serial algorithms using both indeterminate and fully assigned approaches were developed and compared to the performance of ELF itself at the 10.51 threshold (as endorsed by NICE). In total, of the 21 diagnostic approaches compared, ELF at the 10.51 threshold was identified as the best performing, fully assigned approach (non-indeterminate) tested. This analysis also confirmed that diagnostic accuracy could be enhanced by employing parallel algorithms but with a greater number of indeterminate results. In addition, these results also confirmed that serial algorithms could reduce the number of ELF tests required whilst maintaining diagnostic accuracy.[406] These results (published prior to the publication of NICE guidance on the assessment and management of NAFLD[29]) led to an ongoing pilot study in Camden and Islington in North London in which the serial application of FIB4 and ELF have been used to improve the detection of advanced fibrosis from primary care to secondary care.[407, 408] Thereafter, this approach has been recommended in the recent BSG guideline on the management of abnormal liver blood tests.[409] This serial approach limits the number of ELF tests required thereby reducing the costs as compared to an approach in which all patients are tested using ELF. However, this algorithm precludes the use of an ELF containing parallel algorithm (with its high positive and negative predictive values). In a health care system in which ELF testing is freely available, it would seem logical to implement a diagnostic pathway in which a patient is first tested using a parallel strategy (indeterminate approach) by the

application of an indirect biomarker (FIB4) and ELF itself. Whereas a positive or negative result would effectively rule in or rule out advanced fibrosis, an indeterminate result could be reassigned by the application of ELF at a threshold of 10.51. If, however, there is a need to reduce the number of ELF tests that are used, then a diagnostic pathway in which patients are first tested using a parallel strategy composed of 2 indirect tests would be preferred. As before, a positive or negative result will provide strong evidence to rule in or rule out advanced fibrosis: cases classified as indeterminate could then be retested by an ELF containing serial strategy (fully assigned approach) algorithm or ELF alone. However, when working within a diagnostic pathway in which ELF tests are unavailable, one would suggest employing the serial fully assigned strategy of FIB4 followed by NFS (diagnostic accuracy of 82%) to retest cases classified as indeterminate following a parallel strategy composed of the 2 indirect biomarkers. Given these findings a formal health economic analysis should be conducted. Such an analysis must also consider the concurrent availability of imaging modalities such as Fibroscan and ARFI which could be incorporated within these diagnostic strategies. A study examining the serial diagnostic algorithm for advanced fibrosis in NALFLD composed of indirect tests and Fibroscan has recently been published[410] and further studies are therefore required comparing this approach with these proposed algorithms.

Chronic Hepatitis C: Biomarker of Fibrosis

Whereas treatment for NAFLD continues to remain in its infancy, this body of work was conducted during a period which has seen an unimaginable transformation in the treatment of CHC. In such a short space of time, highly toxic treatment regimens that often failed to attain viral clearance have evolved to the highly abbreviated, efficacious and tolerable treatment regimens of today. Whilst there is hope that CHC may eventually be eradicated, globally it is still paramount that treatment to be prioritised for patients with advanced or rapidly advancing fibrosis due to CHC. Moreover, it is essential that patients with CHC are evaluated for the presence of cirrhosis. Regardless of this change in emphasis, treatment for CHC remains a useful platform on which serum markers can be assessed for their ability to monitor fibrosis evolution.

In chapter 8, the performance of 10 biomarkers of fibrosis in CHC were compared in their ability to detect differing degrees of fibrosis ranging from mild to cirrhosis itself. Although the need to accurately stage fibrosis in CHC is now less relevant following the introduction of DAA therapies this is not the case in many parts of the world where access to these therapies is more limited. The detection of cirrhosis related to CHC, however, remains highly relevant due to the ongoing risk of HCC and portal hypertension. The results presented in chapter 8 confirm that many biomarkers (including some indirect panels) exhibit good performance for the detection of liver mild and moderate liver fibrosis. In addition, the diagnostic performance ELF and Hepascore, biomarkers designed to be used in all aetiologies of liver disease was not dissimilar to that demonstrated by the viral specific biomarkers. This is an important finding, as in clinical practice it is not practicable to use multiple different specific biomarkers to evaluate fibrosis in different aetiologies of liver disease. There is a

significant value associated with the ability to use a single test in most or all causes of chronic liver disease.

Perhaps more importantly, these results highlight that biomarker performance can be heavily influenced by the choice of assay used to measure the constituent component of a biomarker panel. This finding is highly relevant given that the open publication of biomarker panel formulae permits the calculation of biomarker scores by clinicians using alternative assays for constituent components. These data confirm that the use of assays that have not been validated for a specific biomarker panel may generate biomarkers scores that are uninterpretable using pre-determined diagnostic thresholds.

Chronic Hepatitis C: biomarkers and the evolution of fibrosis

In Chapter 9, it was assessed whether both liver fibrosis regression and progression arising due to putative antifibrotic therapy could be predicted by changes in ELF score. The evolution of liver fibrosis in this study was assessed by both histology and the ELF test itself. Having failed to attain SVR following interferon based standard of care for CHC, the patients in this study were vulnerable to histologic progression of liver disease. These results identified that baseline and on treatment changes in ELF score could predict fibrosis evolution as assessed by these 2 modalities (histology and ELF). Whilst this study provided the first demonstration that ELF could be used in this context[343], these observations have subsequently been replicated in NAFLD.[411] In this study, baseline and on-treatment factors including ELF were assessed in patients with either advanced fibrosis or cirrhosis due to NAFLD during treatment with Simtuzimab as a putative antifibrotic or placebo (two phase 2b studies). Whilst these studies were terminated prematurely due to a lack of efficacy of Simtuzimab, subsequent analyses confirmed that the evolution of liver histology was associated with both baseline and on-treatment changes in ELF. These data and the data presented in chapter 9 further qualify ELF as a biomarker of fibrosis evolution that can be used when assessing the efficacy of putative antifibrotic therapy. This is a highly important finding given that the risks of liver biopsy render repeated histologic assessment challenging for both patients and clinicians.

In Chapter 10, indirect serum markers were measured in patients with CHC cirrhosis who were randomised to receive either no treatment or escalating dose pegylated interferon alfa 2 a monotherapy as a putative antifibrotic. A reduction in both morbidity and mortality was identified amongst patients receiving pegylated interferon alpha 2a which was associated with an improvement in indirect serum marker scores. Indirect

serum marker scores have been identified to predict longer term outcomes[412] and these results of this study suggest that an improvement in indirect biomarker scores may denote an improved prognosis. Analogous to the findings in chapter 9, changes in indirect serum marker scores could also be used as a surrogate endpoint in clinical trials of investigational antifibrotics. Conceptually, this is most plausible amongst patients with cirrhosis where a reduction of indirect serum marker scores is likely to denote an improvement in portal hypertension.

Other research conducted alongside the work presented in this thesis

ELUCIDATE (ELF to uncover cirrhosis and an indication for diagnosis and action for treatable events)

Trial component of NIHR Programme Grant for Applied Research RP-PG-0707-10101 Evaluating the benefits for patients and the NHS of new and existing biologic fluid biomarkers in liver and renal disease

Translational research projects are conducted with the ultimate intention of improving patient care and clinical outcomes. The studies presented in this thesis have confirmed that serum markers, such as ELF, are able to accurately identify disease states such as advanced fibrosis and cirrhosis. The limitations of liver biopsy preclude its use as a modality to regularly and repeatedly assess liver disease severity. Moreover, ELF levels have been used to denote an increased risk of liver related outcomes.[389] These observations directly led to the formulation of the ELUCIDATE randomised controlled trial that was conducted alongside the work presented in this thesis. In this multicentre randomised controlled trial, patients with suspected or diagnosed CLD in secondary care were randomised 1:1 to receive standard clinical monitoring for the development of cirrhosis with or without regular ELF testing (figure 11.1). As compared to standard clinical care alone, it was hypothesised that patients randomised to receiving ELF testing (alongside standard clinical care) would be diagnosed with cirrhosis in a more timely fashion. This more timely diagnosis of cirrhosis would then lead to improved clinical outcomes by the earlier detection of HCC and varices using screening ultrasound and gastroscopy respectively. In addition, the knowledge that CLD has progressed to a state of advanced fibrosis and cirrhosis may encourage clinicians to more aggressively treat the underlying aetiology of CLD itself; this may result in stabilisation or regression of fibrosis (figure 11.2).

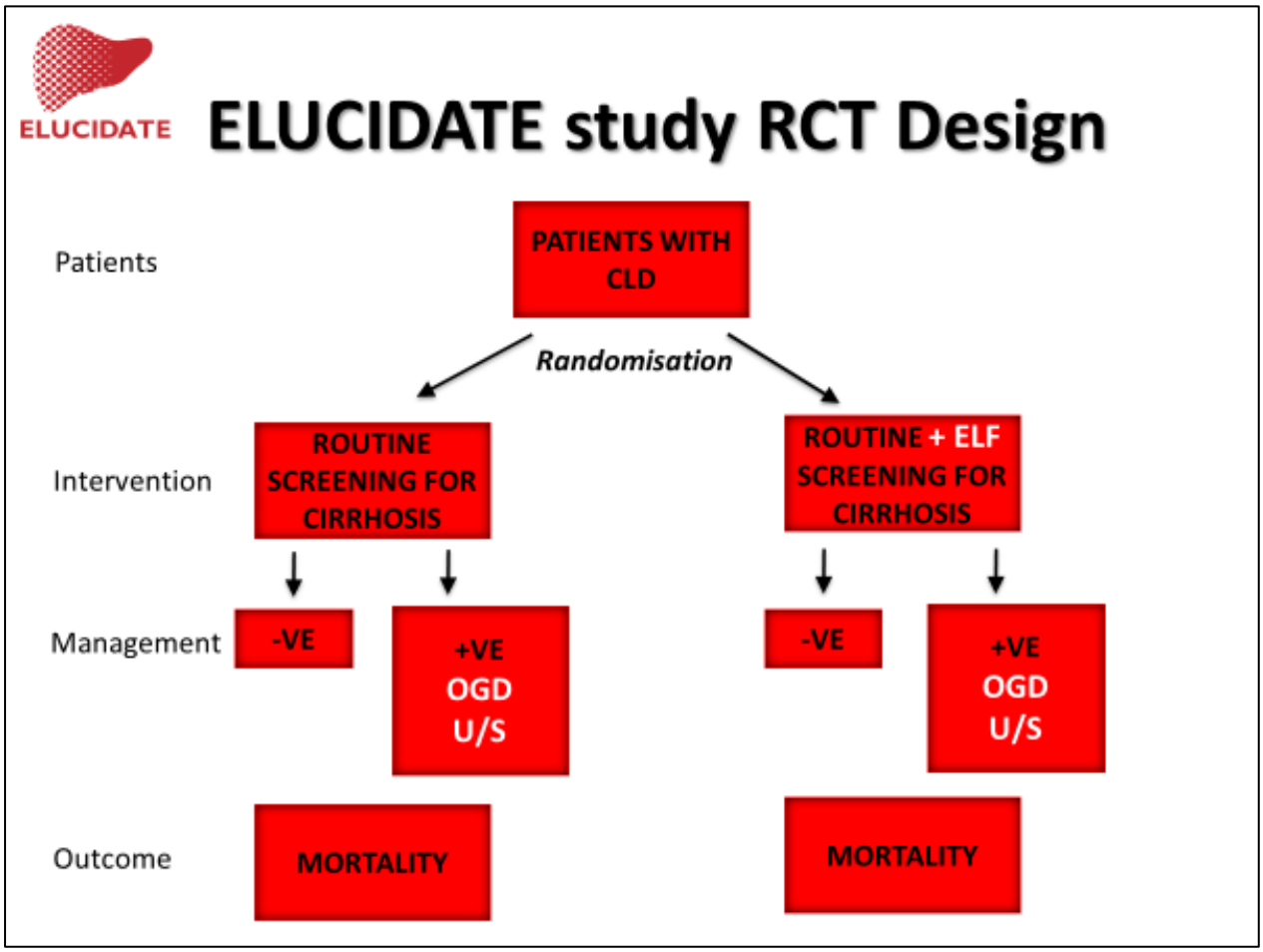


Figure 11.1 Study design of the ELUCIDATE randomised controlled trial

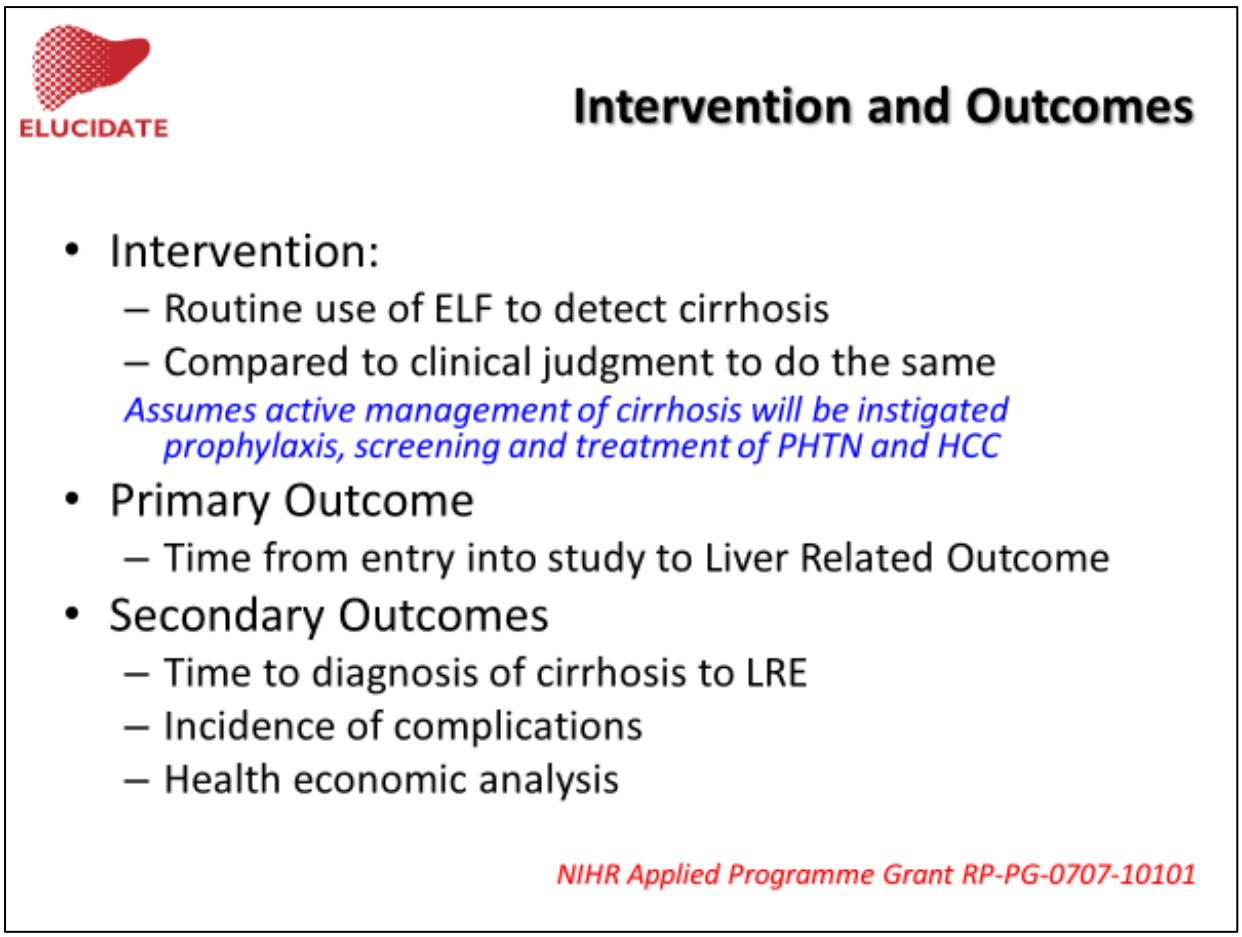


Figure 11.2 Interventions, Primary and Secondary Outcomes of the ELUCIDATE Randomised Controlled Trial

Whereas an ELF threshold of 8.4 was required for entry into the ELUCIDATE study, a threshold of 9.5 was used to denote a diagnosis of cirrhosis amongst patients who were randomised to the ELF arm of study. The 9.5 threshold was selected to detect cirrhosis with 80% sensitivity and 90% specificity[221] whilst also conferring an increase in liver related morbidity.[389] In both ELUCIDATE study arms, patients with cirrhosis (as detected by any means) were required to embark on surveillance for HCC and to undergo gastroscopy to assess for portal hypertension.

In total, 1040 patients were recruited into the ELUCIDATE study across 28 clinical sites in the UK. Whereas the initial process-of-care outcomes have already been

reported[413], the main clinical outcome results are awaited. The impact and consequences of biomarker determined management of cases randomised to the ELF arm of the trial will determine whether or not this approach should be implemented in secondary and tertiary care. Clinical effectiveness, patient and clinician acceptance and health economics should all be determined.

ELF test reproducibility

In the ELUCIDATE study, patient serum samples were sent to a reference laboratory by post in a secure pod. This also may be after a period of freezing at -80 degrees Celsius and subsequent thawing. It was therefore essential that ELF tests results were not influenced by these handling factors. This was investigated in our group's recently published study that confirmed that ELF test results are reproducible using standard clinical handling conditions. [414]

Further research work to be conducted

The work presented in this thesis has contributed to the field of study of biomarkers for the management of chronic liver disease. Firstly, PIIINP has been identified as a biomarker of NASH with its performance validated in 2 cohorts. Secondly, ELF and other biomarkers have demonstrated good to excellent performance for the detection of fibrosis in both NAFLD and CHC. In NAFLD itself, diagnostic algorithms have been proposed and validated thereby successfully enhancing diagnostic performance. Thereafter, changes in ELF score have been used to predict fibrosis evolution during putative antifibrotic therapy. Finally, an improvement of indirect serum marker scores has been found to denote improved prognosis following putative antifibrotic therapy.

Going forward several research questions remain unanswered including:

- Is PIIINP able to detect NASH in large independent studies in a primary care setting?
- How can PIIINP be optimally combined with other variables to maximise the detection of NASH?
- Can PIIINP be used to determine prognosis in NAFLD?
- Can PIIINP be used as a biomarker of NASH to monitor treatment for non-alcoholic steatohepatitis?
- What is the most cost-effective biomarker strategy for the detection of NASH and fibrosis in NAFLD?
- Can the described on-treatment changes in biomarkers be used to predict fibrosis evolution in other cohorts?

- Can changes in biomarker scores be used as a surrogate end point for the development of antifibrotic agents?
- Can the systematic use of biomarkers for the detection of cirrhosis in secondary care result in an improvement in liver related morbidity and mortality?
- Further evaluation should be conducted to determine the factors that impact on the effectiveness of ELF including feeding effects, the influence of age on test performance and the impact of comorbidities on test performance, especially fibrotic conditions such as joint, respiratory and renal disease.

Although this list is not exhaustive, answering these questions will help identify and successfully treat inflammation and fibrosis in the ever-increasing population at risk of CLD.

Final Concluding Remarks

Whereas liver biopsy remains the reference standard by which liver disease is assessed, this technique not practicable for use as the primary method for detecting CLD in the population at risk as it is invasive, hazardous and costly. Approximately one third of the population in the West are currently at risk of developing CLD with standard liver function tests unable to identify CLD with sufficient diagnostic accuracy. These shortcomings also limit the use of liver biopsy as a modality to assess the response to treatment for hepatic inflammation and fibrosis.

This thesis has identified novel biomarker strategies for the detection of progressive liver disease (NASH and advanced fibrosis) due to the most common cause of liver disease worldwide, NAFLD. Moreover, this thesis has validated and compared the performance of biomarkers for the detection of fibrosis due to NAFLD and CHC. This thesis has also confirmed that biomarkers can be used to monitor the evolution of fibrosis due to putative antifibrotic therapy. These results provide a platform from which both diagnostic and therapeutic strategies for CLD can be developed and implemented.

APPENDIX

FORMULAE: SERUM NIT FOR FIBROSIS

1. **ALT/AST ratio**[12]

   $ALT_{(IU/L)}$ / $AST_{(IU/L)}$

2. **APRI**[340]

   $AST_{(IU/L)}$ /ULN × 100/platelet count (10^9/L)

3. **BAAT Score**[338]

   BMI (≥28=1, 28<0),

   age at liver biopsy (≥50 years=1, 50<0),

   ALT (≥2ULN=1, <2N=0),

   serum triglycerides(≥1.7 mmol/l= 1,<1.7= 0)

4. **BARD**[255]

   composed of 3 variables:

   an AST/ALT ratio >0.8 = 2 points; a BMI >28 =1 point; presence of diabetes =1 point.

5. **Cirrhosis Discriminant Score**[339]

Platelet count >340=0, 280–339=1, 220–279=2, 160–219=3, 100–159=4, 40–99=5, <40=6

AST/ALT ratio 1.7=0 1.2–1.7=1 0.6–1.19=2 0.6=3

INR <1.1=0, 1.1–1.4=1, >1.4=2

Score is the sum of three (0–11)

6. **ELF score** [221]

$= 2.278 + 0.851 \ln(HA_{ng/mL}) + 0.751 \ln(PIIINP_{ng/mL}) + 0.394 \ln(TIMP1_{ng/mL})$

7. **FIB4**[341]

$(Age_{years} \times AST_{IU/L}) / (Platelets_{10^9/L} \times (\sqrt{} (ALT_{IU/L}))$

8. **Fibroindex** [238]

$= 1.738 - 0.064 \times (platelets_{10^3/ml})$ [$+ 0.005 \times (AST_{IU/L})$

$+ 0.463 \times (gamma\ globulin_{g/dl})$

9. **Fibrometer**[361]

−0.007 × platelets G/l

−0.049 × prothrombin time %

+ 0.012 × AST IU/L

+ 0.005 × α2 macroglobulin mg/dl

+ 0.021 × hyaluronate mg/l

−0.270 × urea mmol/l

+ 0.027 × age years

+ 3.718

10. **Fibrometer 2G (Second generation Fibrometer)**[361]**:**

Fibrometer 2G is a serum marker of fibrosis developed for use in viral liver disease including CHC. The Fibrometer 2G (virus) scores used in this study were kindly calculated by Prof. Paul Calès using the patented formulae (Echosens) incorporating the values of assays for α2-macroglobulin, AST, Age, Platelets, Prothrombin Time, Sex, Urea, HA.

11. **Fibrometer 3G (Third generation Fibrometer (Fibrometer 3G)**[360]**:**

Fibrometer 3G is a serum marker of fibrosis developed for use in viral liver disease including CHC. The Fibrometer 3G scores used in this study were kindly calculated by Prof. Paul Calès using the patented formulae (Echosens) incorporating the values of assays for α2-macroglobulin, AST, Age, Platelets, Protrombin Time, Sex, GGT, Urea.

12. Fibrospect II[362]

Fibrospect II scores were calculated by entering the values of HA, TIMP-1 and α2-macroglobulin into the patented formula (United States Patent 7,670,764 B2).

$$\text{Index} = \frac{\text{Exp}^{[-4.3633+(0.0108*\text{HA ng/ml})+(0.0015*\text{TIMP}-1\text{ ng/ml})+(0.5357*\text{A2M mg/ml})]}}{1+\text{Exp}^{[-4.3633+(0.0108*\text{HA ng/ml})+(0.0015*\text{TIMP}-1\text{ ng/ml})+(0.5357*\text{A2M mg/ml})]}}$$

13. Fibrotest[223]

Patented formula of

α2macroglobulin, γGT, apolipoprotein A1, haptoglobin, bilirubin, age and gender

14. Forns Index[237]

$= 7.811 - 3.131 \times \ln(\text{platelet count}_{G/}) + 0.781 \times \ln \gamma\text{GT}_{IU/L}$

$+ 3.467 \times \ln \text{age}_{years} - 0.014 \times \text{cholesterol}_{g/l}$

15. Goteborg University Cirrhosis Index (GUCI)[415]

$(AST_{(IU/L}$ / Upper limit of normal) × INR × 100(Platelet Count)

16. HALT-C model[235]

= -3.66 - 0.00995 x $platelets_{10^3/ml}$ + 0.008 x serum $TIMP\text{-}1_{ng/mL}$ + 1.42 x log($hyaluronate_{ng/mL}$)

17. Hepascore[224]

= y / (1 + y)

with y =exp

(−4.185818 − (0.0249 × age) + (0.7464 × 1 if male, 0 if female gender) + (1.0039 × α2 macroglobulin) + (0.0302 × $hyaluronate_{ng/mL}$) + (0.0691 × bilirubin) − (0.0012 × γGT))

18. Lok Index[239]

= -5.56 – (0.0089 x $platelet_{10^3/ml}$) +1.26 x AST/ALT ratio+ 5.27 x INR

19. MP3 [229]

= 0.5903 × Log PIIINP $_{ng/ml}$ −0.1749 × Log MMP1 $_{ng/ml}$

20. NAFLD Fibrosis Score[252]

− 1.675 + 0.037 × age (years) + 0.094 × BMI (kg/m2)

+ 1.13 × IFG/diabetes (yes = 1, no = 0) + 0.99 × AST/ALT ratio

- 0.013 × platelet (×109/l) − 0.66 × albumin (g/dl)

21. SHASTA[416]

-3.84

+ 1.70 (1 if HA 41–85 ng/ml, 0 otherwise)

+ 3.28 (1 if HA>85 ng/ml, 0 otherwise)

+ 1.58 (albumin<3.5 g/dl, 0 otherwise)

+1.78 (1 if AST >60 IU/l, 0 otherwise)

22. Virahep-C model [416]

= -5.17 + 0.20 x race + 0.07 x age (yr) + 1.19 ln($AST_{IU/L}$)

- 1.76 ln(platelet count $_{10^3/ml}$) + 1.38 ln(alkaline $phosphatase_{IU/L}$)

ADVIA Centaur® systems' Enhanced Liver Fibrosis (ELF™)

test procedure

Blood Sample Preparation Instructions:

- 5ml of Blood is collected in SST® Serum Separation Tubes (Gold Top).
- Thereafter, blood is allowed to clot adequately prior to centrifugation.
- Serum should be separated from whole blood within six hours of bloodletting.
- Serum is separated by centrifugation at 1500 x g for 10 minutes at room temperature.
- Whole blood may be stored within the range 2 to 25°C prior to separation but cannot be frozen.
- The minimum recommended serum volume is 250 µL.
- After centrifugation, serum is transferred to a sterile polypropylene tube with screw cap capable of withstanding freezing.
- The sample tube is labelled with at least two unique sample identifiers.

Blood Sample Transport Instructions:

- Should a delay in transport occur, samples may be stored at 2-8°C for up to 48 hours or at -70°C for longer term storage.

- Serum may be transported within the range from -70°C to ambient temperature.
- It is recommended that frozen samples are sent on dry ice (UN1845).

To calculate the ELF score:

Quantitative measurements of hyaluronic acid (HA), amino-terminal propeptide of type III procollagen (PIIINP), and tissue inhibitor of metalloproteinase 1 (TIMP-1) in human serum on the ADVIA Centaur XP, ADVIA Centaur XPT, and ADVIA Centaur CP systems only.

ADVIA Centaur XP/XPT:

$$\text{ELF score} = 2.278 + 0.851 \ln(C_{HA}) + 0.751 \ln(C_{PIIINP}) + 0.394 \ln(C_{TIMP1})$$

ADVIA Centaur CP:

$$\text{ELF score} = 2.494 + 0.846 \ln(C_{HA}) + 0.735 \ln(C_{PIIINP}) + 0.391 \ln(C_{TIMP1})$$

Concentrations (C) of each assay are in ng/mL

REFERENCES

1. Mortality statistics: deaths registered in England and Wales (series DR). Office for National Statistics, Newport, UK (2011). Office for National Statistics, 2010.
2. Williams R, Aspinall R, Bellis M, et al. Addressing liver disease in the UK: a blueprint for attaining excellence in health care and reducing premature mortality from lifestyle issues of excess consumption of alcohol, obesity, and viral hepatitis. Lancet 2014;384:1953-97.
3. Schuppan D, Afdhal NH. Liver cirrhosis. Lancet 2008;371:838-51.
4. 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;34:274-85.
5. Younossi ZM, Koenig AB, Abdelatif D, et al. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology 2016;64:73-84.
6. Goldberg D, Ditah IC, Saeian K, et al. Changes in the Prevalence of Hepatitis C Virus Infection, Non-alcoholic Steatohepatitis, and Alcoholic Liver Disease Among Patients with Cirrhosis or Liver Failure on the Waitlist for Liver Transplantation. Gastroenterology 2017.
7. Browning JD, Szczepaniak LS, Dobbins R, et al. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. Hepatology 2004;40:1387-95.
8. Bedogni G, Miglioli L, Masutti F, et al. Prevalence of and risk factors for nonalcoholic fatty liver disease: the Dionysos nutrition and liver study. Hepatology 2005;42:44-52.
9. Adams LA, Lymp JF, St Sauver J, et al. The natural history of nonalcoholic fatty liver disease: a population-based cohort study. Gastroenterology 2005;129:113-21.
10. Armstrong MJ, Houlihan DD, Bentham L, et al. Presence and severity of non-alcoholic fatty liver disease in a large prospective primary care cohort. J Hepatol 2012;56:234-40.
11. Feldman A, Eder SK, Felder TK, et al. Clinical and Metabolic Characterization of Lean Caucasian Subjects With Non-alcoholic Fatty Liver. Am J Gastroenterol 2017;112:102-110.

12. WHO weekly epidemiological record. 2002.
13. Hepatitis C in the UK Health Protection Agency Annual Report 2011.
14. De Angelis D, Sweeting M, Ades A, et al. An evidence synthesis approach to estimating Hepatitis C prevalence in England and Wales. Stat Methods Med Res 2009;18:361-79.
15. Di Bisceglie AM. Natural history of hepatitis C: its impact on clinical management. Hepatology 2000;31:1014-8.
16. Thomas DL, Seeff LB. Natural history of hepatitis C. Clin Liver Dis 2005;9:383-98, vi.
17. Fleming KM, Aithal GP, Card TR, et al. The rate of decompensation and clinical progression of disease in people with cirrhosis: a cohort study. Aliment Pharmacol Ther 2010;32:1343-50.
18. Ratib S, Fleming KM, Crooks CJ, et al. 1 and 5 year survival estimates for people with cirrhosis of the liver in England, 1998-2009: a large population study. J Hepatol 2014;60:282-9.
19. Bialecki ES, Di Bisceglie AM. Clinical presentation and natural course of hepatocellular carcinoma. Eur J Gastroenterol Hepatol 2005;17:485-9.
20. National Health Service Blood and TransplantTaking organ transplantation to 2020.
21. National Institute for Health and Care Excellence: Cirrhosis in over 16s: assessment and management. guideline [NG50].
22. Oseini AM, Sanyal AJ. Therapies in non-alcoholic steatohepatitis (NASH). Liver Int 2017;37 Suppl 1:97-103.
23. Choo Q, Kuo G, Weiner A, et al. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 1989;244:359-362.
24. EASL Recommendations on Treatment of Hepatitis C 2016. J Hepatol 2017;66:153-194.
25. Harris RJ, Thomas B, Griffiths J, et al. Increased uptake and new therapies are needed to avert rising hepatitis C-related end stage liver disease in England: modelling the predicted impact of treatment under different scenarios. J Hepatol 2014;61:530-7.
26. Puoti C, Guarisco R, Spilabotti L, et al. Should we treat HCV carriers with normal ALT levels? The '5Ws' dilemma. J Viral Hepat 2012;19:229-35.

27. Park JJ, Park JY, Kim DY, et al. Prediction of significant fibrosis in chronic hepatitis C patients with normal ALT. Hepatogastroenterology 2011;58:1321-7.
28. Rockey DC, Caldwell SH, Goodman ZD, et al. Liver biopsy. Hepatology 2009;49:1017-44.
29. National Institute for Health and Care Excellence. Non-alcoholic fatty liver disease (NAFLD): assessment and management (guideline NG49), 2016.
30. Bacon BR, O'Grady JG. Comprehensive Clinical Hepatology: Elsevier Mosby, 2006.
31. Rosenberg W, Baldrick T, Tanwar S. Liver Disease. In: Rifai N, Horvath AR, Wittwer C, eds. Tietz textbook of clinical chemistry and molecular diagnostics. Sixth edition. ed.
32. Ebaid A. Analytical solutions for the mathematical model describing the formation of liver zones via Adomian's method. Comput Math Methods Med 2013;2013:547954.
33. Bedossa P, Paradis V. Liver extracellular matrix in health and disease. J Pathol 2003;200:504-15.
34. Schuppan D. Structure of the extracellular matrix in normal and fibrotic liver: collagens and glycoproteins. Semin Liver Dis 1990;10:1-10.
35. OpenStax. Anatomy & Physiology. Volume 2017. v 8.56 ed, 2013.
36. Chouhan MD, Lythgoe MF, Mookerjee RP, et al. Vascular assessment of liver disease-towards a new frontier in MRI. Br J Radiol 2016:20150675.
37. Protzer U, Maini MK, Knolle PA. Living in the liver: hepatic infections. Nat Rev Immunol 2012;12:201-13.
38. Delneste Y, Beauvillain C, Jeannin P. [Innate immunity: structure and function of TLRs]. Med Sci (Paris) 2007;23:67-73.
39. Kawai T, Akira S. Innate immune recognition of viral infection. Nat Immunol 2006;7:131-7.
40. Alexopoulou L, Holt AC, Medzhitov R, et al. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. Nature 2001;413:732-8.
41. Kono H, Rock KL. How dying cells alert the immune system to danger. Nat Rev Immunol 2008;8:279-89.
42. McDonald B, Pittman K, Menezes GB, et al. Intravascular danger signals guide neutrophils to sites of sterile inflammation. Science 2010;330:362-6.

43. Xu R, Huang H, Zhang Z, et al. The role of neutrophils in the development of liver diseases. Cell Mol Immunol 2014;11:224-31.
44. Nahrendorf M, Pittet MJ, Swirski FK. Monocytes: protagonists of infarct inflammation and repair after myocardial infarction. Circulation 2010;121:2437-45.
45. Geissmann F, Manz MG, Jung S, et al. Development of monocytes, macrophages, and dendritic cells. Science 2010;327:656-61.
46. Dal-Secco D, Wang J, Zeng Z, et al. A dynamic spectrum of monocytes arising from the in situ reprogramming of CCR2+ monocytes at a site of sterile injury. J Exp Med 2015;212:447-56.
47. Crispe IN. The liver as a lymphoid organ. Annu Rev Immunol 2009;27:147-63.
48. Italiani P, Boraschi D. From Monocytes to M1/M2 Macrophages: Phenotypical vs. Functional Differentiation. Front Immunol 2014;5:514.
49. Rossaint J, Zarbock A. Platelets in leucocyte recruitment and function. Cardiovasc Res 2015;107:386-95.
50. Seino K, Taniguchi M. Functionally distinct NKT cell subsets and subtypes. J Exp Med 2005;202:1623-6.
51. Radaeva S, Sun R, Jaruga B, et al. Natural killer cells ameliorate liver fibrosis by killing activated stellate cells in NKG2D-dependent and tumor necrosis factor-related apoptosis-inducing ligand-dependent manners. Gastroenterology 2006;130:435-52.
52. Romagnani S. Type 1 T helper and type 2 T helper cells: functions, regulation and role in protection and disease. Int J Clin Lab Res 1991;21:152-8.
53. Scheibner KA, Lutz MA, Boodoo S, et al. Hyaluronan fragments act as an endogenous danger signal by engaging TLR2. J Immunol 2006;177:1272-81.
54. Fernandes-Alnemri T, Yu JW, Datta P, et al. AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. Nature 2009;458:509-13.
55. Chan FK, Luz NF, Moriwaki K. Programmed necrosis in the cross talk of cell death and inflammation. Annu Rev Immunol 2015;33:79-106.
56. Szabo G, Csak T. Inflammasomes in liver diseases. J Hepatol 2012;57:642-54.
57. Friedman SL. Mechanisms of hepatic fibrogenesis. Gastroenterology 2008;134:1655-69.
58. Friedman SL. Liver fibrosis: from mechanisms to treatment. Gastroenterol Clin Biol 2007;31:812-4.

59. Moreira RK. Hepatic stellate cells and liver fibrosis. Arch Pathol Lab Med 2007;131:1728-34.
60. Giampieri MP, Jezequel AM, Orlandi F. The lipocytes in normal human liver. A quantitative study. Digestion 1981;22:165-9.
61. Rockey DC. Endothelial Dysfunction in Advanced Liver Disease. The American Journal of the Medical Sciences;349:6-16.
62. Canbay A, Higuchi H, Bronk SF, et al. Fas enhances fibrogenesis in the bile duct ligated mouse: a link between apoptosis and fibrosis. Gastroenterology 2002;123:1323-30.
63. Hernandez-Gea V, Friedman SL. Pathogenesis of liver fibrosis. Annu Rev Pathol 2011;6:425-56.
64. Borkham-Kamphorst E, van Roeyen CR, Ostendorf T, et al. Pro-fibrogenic potential of PDGF-D in liver fibrosis. J Hepatol 2007;46:1064-74.
65. Nobili V, Parola M, Alisi A, et al. Oxidative stress parameters in paediatric non-alcoholic fatty liver disease. Int J Mol Med 2010;26:471-6.
66. Thabut D, Routray C, Lomberk G, et al. Complementary vascular and matrix regulatory pathways underlie the beneficial mechanism of action of sorafenib in liver fibrosis. Hepatology 2011;54:573-85.
67. Friedman SL. Hepatic stellate cells: protean, multifunctional, and enigmatic cells of the liver. Physiol Rev 2008;88:125-72.
68. Knittel T, Mehde M, Kobold D, et al. Expression patterns of matrix metalloproteinases and their inhibitors in parenchymal and non-parenchymal cells of rat liver: regulation by TNF-alpha and TGF-beta1. J Hepatol 1999;30:48-60.
69. Bourd-Boittin K, Bonnier D, Leyme A, et al. Protease profiling of liver fibrosis reveals the ADAM metallopeptidase with thrombospondin type 1 motif, 1 as a central activator of transforming growth factor beta. Hepatology 2011;54:2173-84.
70. Day CP, James OF. Hepatic steatosis: innocent bystander or guilty party? Hepatology 1998;27:1463-6.
71. Friedman SL. Liver fibrosis in 2012: Convergent pathways that cause hepatic fibrosis in NASH. Nat Rev Gastroenterol Hepatol 2013;10:71-2.
72. Tilg H, Moschen AR. Evolution of inflammation in nonalcoholic fatty liver disease: the multiple parallel hits hypothesis. Hepatology 2010;52:1836-46.

73. Speliotes EK, Yerges-Armstrong LM, Wu J, et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. PLoS Genet 2011;7:e1001324.
74. Xu R, Tao A, Zhang S, et al. Association between patatin-like phospholipase domain containing 3 gene (PNPLA3) polymorphisms and nonalcoholic fatty liver disease: a HuGE review and meta-analysis. Sci Rep 2015;5:9284.
75. Santoro N, Kursawe R, D'Adamo E, et al. A common variant in the patatin-like phospholipase 3 gene (PNPLA3) is associated with fatty liver disease in obese children and adolescents. Hepatology 2010;52:1281-90.
76. Xu C, Wang G, Hao Y, et al. Correlation analysis between gene expression profile of rat liver tissues and high-fat emulsion-induced nonalcoholic fatty liver. Dig Dis Sci 2011;56:2299-308.
77. Brunt EM, Kleiner DE, Wilson LA, et al. Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. Hepatology 2011;53:810-20.
78. Li ZZ, Berk M, McIntyre TM, et al. Hepatic lipid partitioning and liver damage in nonalcoholic fatty liver disease: role of stearoyl-CoA desaturase. J Biol Chem 2009;284:5637-44.
79. Liu J, Han L, Zhu L, et al. Free fatty acids, not triglycerides, are associated with non-alcoholic liver injury progression in high fat diet induced obese rats. Lipids Health Dis 2016;15:27.
80. Zhang XQ, Xu CF, Yu CH, et al. Role of endoplasmic reticulum stress in the pathogenesis of nonalcoholic fatty liver disease. World J Gastroenterol 2014;20:1768-76.
81. Lake AD, Novak P, Hardwick RN, et al. The adaptive endoplasmic reticulum stress response to lipotoxicity in progressive human nonalcoholic fatty liver disease. Toxicol Sci 2014;137:26-35.
82. Meakin PJ, Chowdhry S, Sharma RS, et al. Susceptibility of Nrf2-null mice to steatohepatitis and cirrhosis upon consumption of a high-fat diet is associated with oxidative stress, perturbation of the unfolded protein response, and disturbance in the expression of metabolic enzymes but not with insulin resistance. Mol Cell Biol 2014;34:3305-20.

83. Vonghia L, Michielsen P, Francque S. Immunological mechanisms in the pathophysiology of non-alcoholic steatohepatitis. Int J Mol Sci 2013;14:19867-90.
84. Mouzaki M, Comelli EM, Arendt BM, et al. Intestinal microbiota in patients with nonalcoholic fatty liver disease. Hepatology 2013;58:120-7.
85. Henao-Mejia J, Elinav E, Jin C, et al. Inflammasome-mediated dysbiosis regulates progression of NAFLD and obesity. Nature 2012;482:179-85.
86. Zhu L, Baker SS, Gill C, et al. Characterization of gut microbiomes in nonalcoholic steatohepatitis (NASH) patients: a connection between endogenous alcohol and NASH. Hepatology 2013;57:601-9.
87. Wan J, Benkdane M, Teixeira-Clerc F, et al. M2 Kupffer cells promote M1 Kupffer cell apoptosis: a protective mechanism against alcoholic and nonalcoholic fatty liver disease. Hepatology 2014;59:130-42.
88. Wenfeng Z, Yakun W, Di M, et al. Kupffer cells: increasingly significant role in nonalcoholic fatty liver disease. Ann Hepatol 2014;13:489-95.
89. Eguchi A, Wree A, Feldstein AE. Biomarkers of liver cell death. J Hepatol 2014;60:1063-74.
90. Huang W, Metlakunta A, Dedousis N, et al. Depletion of liver Kupffer cells prevents the development of diet-induced hepatic steatosis and insulin resistance. Diabetes 2010;59:347-57.
91. Sharifnia T, Antoun J, Verriere TG, et al. Hepatic TLR4 signaling in obese NAFLD. Am J Physiol Gastrointest Liver Physiol 2015;309:G270-8.
92. Walenbergh SM, Koek GH, Bieghs V, et al. Non-alcoholic steatohepatitis: the role of oxidized low-density lipoproteins. J Hepatol 2013;58:801-10.
93. Bieghs V, van Gorp PJ, Walenbergh SM, et al. Specific immunization strategies against oxidized low-density lipoprotein: a novel way to reduce nonalcoholic steatohepatitis in mice. Hepatology 2012;56:894-903.
94. Rensen SS, Slaats Y, Nijhuis J, et al. Increased hepatic myeloperoxidase activity in obese subjects with nonalcoholic steatohepatitis. Am J Pathol 2009;175:1473-82.
95. Alkhouri N, Morris-Stiff G, Campbell C, et al. Neutrophil to lymphocyte ratio: a new marker for predicting steatohepatitis and fibrosis in patients with nonalcoholic fatty liver disease. Liver Int 2012;32:297-302.

96. Ibusuki R, Uto H, Arima S, et al. Transgenic expression of human neutrophil peptide-1 enhances hepatic fibrosis in mice fed a choline-deficient, L-amino acid-defined diet. Liver Int 2013;33:1549-56.
97. Talukdar S, Oh DY, Bandyopadhyay G, et al. Neutrophils mediate insulin resistance in mice fed a high-fat diet through secreted elastase. Nat Med 2012;18:1407-12.
98. Brunt EM. Pathology of nonalcoholic fatty liver disease. Nat Rev Gastroenterol Hepatol 2010;7:195-203.
99. Gadd VL, Skoien R, Powell EE, et al. The portal inflammatory infiltrate and ductular reaction in human nonalcoholic fatty liver disease. Hepatology 2014;59:1393-405.
100. Sumarac-Dumanovic M, Stevanovic D, Ljubic A, et al. Increased activity of interleukin-23/interleukin-17 proinflammatory axis in obese women. Int J Obes (Lond) 2009;33:151-6.
101. Tang Y, Bian Z, Zhao L, et al. Interleukin-17 exacerbates hepatic steatosis and inflammation in non-alcoholic fatty liver disease. Clin Exp Immunol 2011;166:281-90.
102. Harley IT, Stankiewicz TE, Giles DA, et al. IL-17 signaling accelerates the progression of nonalcoholic fatty liver disease in mice. Hepatology 2014;59:1830-9.
103. Winer DA, Winer S, Chng MH, et al. B Lymphocytes in obesity-related adipose tissue inflammation and insulin resistance. Cell Mol Life Sci 2014;71:1033-43.
104. Sutti S, Jindal A, Locatelli I, et al. Adaptive immune responses triggered by oxidative stress contribute to hepatic inflammation in NASH. Hepatology 2014;59:886-97.
105. Winer DA, Winer S, Shen L, et al. B cells promote insulin resistance through modulation of T cells and production of pathogenic IgG antibodies. Nat Med 2011;17:610-7.
106. Csak T, Ganz M, Pespisa J, et al. Fatty acid and endotoxin activate inflammasomes in mouse hepatocytes that release danger signals to stimulate immune cells. Hepatology 2011;54:133-44.
107. Wree A, McGeough MD, Pena CA, et al. NLRP3 inflammasome activation is required for fibrosis development in NAFLD. J Mol Med (Berl) 2014;92:1069-82.

108. Kamari Y, Shaish A, Vax E, et al. Lack of interleukin-1alpha or interleukin-1beta inhibits transformation of steatosis to steatohepatitis and liver fibrosis in hypercholesterolemic mice. J Hepatol 2011;55:1086-94.
109. Dixon LJ, Berk M, Thapaliya S, et al. Caspase-1-mediated regulation of fibrogenesis in diet-induced steatohepatitis. Lab Invest 2012;92:713-23.
110. Wanninger J, Neumeier M, Hellerbrand C, et al. Lipid accumulation impairs adiponectin-mediated induction of activin A by increasing TGFbeta in primary human hepatocytes. Biochim Biophys Acta 2011;1811:626-33.
111. Bansal R, van Baarlen J, Storm G, et al. The interplay of the Notch signaling in hepatic stellate cells and macrophages determines the fate of liver fibrogenesis. Sci Rep 2015;5:18272.
112. Schnabl B, Bradham CA, Bennett BL, et al. TAK1/JNK and p38 have opposite effects on rat hepatic stellate cells. Hepatology 2001;34:953-63.
113. Guy CD, Suzuki A, Zdanowicz M, et al. Hedgehog pathway activation parallels histologic severity of injury and fibrosis in human nonalcoholic fatty liver disease. Hepatology 2012;55:1711-21.
114. Lauer GM, Walker BD. Hepatitis C virus infection. N Engl J Med 2001;345:41-52.
115. Blight KJ, Kolykhalov AA, Reed KE, et al. Molecular virology of hepatitis C virus: an update with respect to potential antiviral targets. Antivir Ther 1998;3:71-81.
116. Tomei L, Failla C, Santolini E, et al. NS3 is a serine protease required for processing of hepatitis C virus polyprotein. J Virol 1993;67:4017-26.
117. Lohmann V, Korner F, Koch J, et al. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. Science 1999;285:110-3.
118. Bartenschlager R. Hepatitis C virus replicons: potential role for drug development. Nat Rev Drug Discov 2002;1:911-6.
119. Tomei L, Altamura S, Paonessa G, et al. HCV antiviral resistance: the impact of in vitro studies on the development of antiviral agents targeting the viral NS5B polymerase. Antivir Chem Chemother 2005;16:225-45.
120. Tanwar S, Trembling PM, Dusheiko GM. TMC435 for the treatment of chronic hepatitis C. Expert Opin Investig Drugs 2012;21:1193-209.
121. Wang N, Liang Y, Devaraj S, et al. Toll-like receptor 3 mediates establishment of an antiviral state against hepatitis C virus in hepatoma cells. J Virol 2009;83:9824-34.

122. Ank N, Iversen MB, Bartholdy C, et al. An important role for type III interferon (IFN-lambda/IL-28) in TLR-induced antiviral activity. J Immunol 2008;180:2474-85.
123. Coenen M, Nischalke HD, Kramer B, et al. Hepatitis C virus core protein induces fibrogenic actions of hepatic stellate cells via toll-like receptor 2. Lab Invest 2011;91:1375-82.
124. Schulze-Krebs A, Preimel D, Popov Y, et al. Hepatitis C virus-replicating hepatocytes induce fibrogenic activation of hepatic stellate cells. Gastroenterology 2005;129:246-58.
125. Aweya JJ, Tan YJ. Modulation of programmed cell death pathways by the hepatitis C virus. Front Biosci (Landmark Ed) 2011;16:608-18.
126. Mazzocca A, Sciammetta SC, Carloni V, et al. Binding of hepatitis C virus envelope protein E2 to CD81 up-regulates matrix metalloproteinase-2 in human hepatic stellate cells. J Biol Chem 2005;280:11329-39.
127. Bataller R, Paik YH, Lindquist JN, et al. Hepatitis C virus core and nonstructural proteins induce fibrogenic effects in hepatic stellate cells. Gastroenterology 2004;126:529-40.
128. Lin W, Tsai WL, Shao RX, et al. Hepatitis C virus regulates transforming growth factor beta1 production through the generation of reactive oxygen species in a nuclear factor kappaB-dependent manner. Gastroenterology 2010;138:2509-18, 2518 e1.
129. Tanwar S, Trembling PM, Dusheiko GM. Hepatitis C Therapy: Lessons of the Last Two Decades. Curr Hepat Rep 2012;11:119-127.
130. Liu Z, Wei X, Chen T, et al. Characterization of Fibrosis Changes in Chronic Hepatitis C Patients after Virological Cure: A Systematic Review with Meta-analysis. J Gastroenterol Hepatol 2016.
131. Akhtar E, Manne V, Saab S. Cirrhosis regression in hepatitis C patients with sustained virological response after antiviral therapy: a meta-analysis. Liver Int 2015;35:30-6.
132. Issa R, Zhou X, Constandinou CM, et al. Spontaneous recovery from micronodular cirrhosis: evidence for incomplete resolution associated with matrix cross-linking. Gastroenterology 2004;126:1795-808.

133. Wanless IR, Nakashima E, Sherman M. Regression of human cirrhosis. Morphologic features and the genesis of incomplete septal cirrhosis. Arch Pathol Lab Med 2000;124:1599-607.
134. Fernandez M, Semela D, Bruix J, et al. Angiogenesis in liver disease. J Hepatol 2009;50:604-20.
135. Trautwein C, Friedman SL, Schuppan D, et al. Hepatic fibrosis: Concept to treatment. J Hepatol 2015;62:S15-24.
136. Iredale JP, Benyon RC, Pickering J, et al. Mechanisms of spontaneous resolution of rat liver fibrosis. Hepatic stellate cell apoptosis and reduced hepatic expression of metalloproteinase inhibitors. J Clin Invest 1998;102:538-49.
137. Fallowfield JA, Mizuno M, Kendall TJ, et al. Scar-associated macrophages are a major source of hepatic matrix metalloproteinase-13 and facilitate the resolution of murine hepatic fibrosis. J Immunol 2007;178:5288-95.
138. Ramachandran P, Pellicoro A, Vernon MA, et al. Differential Ly-6C expression identifies the recruited macrophage phenotype, which orchestrates the regression of murine liver fibrosis. Proc Natl Acad Sci U S A 2012;109:E3186-95.
139. Iredale JP. Hepatic stellate cell behavior during resolution of liver injury. Semin Liver Dis 2001;21:427-36.
140. Gressner AM. The cell biology of liver fibrogenesis - an imbalance of proliferation, growth arrest and apoptosis of myofibroblasts. Cell Tissue Res 1998;292:447-52.
141. Oakley F, Trim N, Constandinou CM, et al. Hepatocytes express nerve growth factor during liver injury: evidence for paracrine regulation of hepatic stellate cell apoptosis. Am J Pathol 2003;163:1849-58.
142. Fischer R, Cariers A, Reinehr R, et al. Caspase 9-dependent killing of hepatic stellate cells by activated Kupffer cells. Gastroenterology 2002;123:845-61.
143. Schnabl B, Purbeck CA, Choi YH, et al. Replicative senescence of activated human hepatic stellate cells is accompanied by a pronounced inflammatory but less fibrogenic phenotype. Hepatology 2003;37:653-64.
144. Troeger JS, Mederacke I, Gwak GY, et al. Deactivation of hepatic stellate cells during liver fibrosis resolution in mice. Gastroenterology 2012;143:1073-83 e22.

145. Kisseleva T, Cong M, Paik Y, et al. Myofibroblasts revert to an inactive phenotype during regression of liver fibrosis. Proc Natl Acad Sci U S A 2012;109:9448-53.
146. Rincon D, Ripoll C, Lo Iacono O, et al. Antiviral therapy decreases hepatic venous pressure gradient in patients with chronic hepatitis C and advanced fibrosis. Am J Gastroenterol 2006;101:2269-74.
147. Pockros PJ, Hamzeh FM, Martin P, et al. Histologic outcomes in hepatitis C-infected patients with varying degrees of virologic response to interferon-based treatments. Hepatology 2010;52:1193-200.
148. Fort J, Pilette C, Veal N, et al. Effects of long-term administration of interferon alpha in two models of liver fibrosis in rats. J Hepatol 1998;29:263-70.
149. Inagaki Y, Nemoto T, Kushida M, et al. Interferon alfa down-regulates collagen gene transcription and suppresses experimental hepatic fibrosis in mice. Hepatology 2003;38:890-9.
150. Poynard T, McHutchison J, Manns M, et al. Impact of pegylated interferon alfa-2b and ribavirin on liver fibrosis in patients with chronic hepatitis C. Gastroenterology 2002;122:1303-13.
151. Di Bisceglie AM, Shiffman ML, Everson GT, et al. Prolonged therapy of advanced chronic hepatitis C with low-dose peginterferon. N Engl J Med 2008;359:2429-41.
152. Afdhal NH, Levine R, Brown Jr R, et al. Colchicine versus PEG-interferon alfa 2b long term therapy: results of the 4 year COPILOT trial." Journal of Hepatology 2008;48, Supplement 2:S4.
153. Klein S, Van Beuge MM, Granzow M, et al. HSC-specific inhibition of Rho-kinase reduces portal pressure in cirrhotic rats without major systemic effects. J Hepatol 2012;57:1220-7.
154. Kaps L, Nuhn L, Aslam M, et al. In Vivo Gene-Silencing in Fibrotic Liver by siRNA-Loaded Cationic Nanohydrogel Particles. Adv Healthc Mater 2015;4:2809-15.
155. Canbay A, Guicciardi ME, Higuchi H, et al. Cathepsin B inactivation attenuates hepatic injury and fibrosis during cholestasis. J Clin Invest 2003;112:152-9.
156. Canbay A, Feldstein A, Baskin-Bey E, et al. The caspase inhibitor IDN-6556 attenuates hepatic injury and fibrosis in the bile duct ligated mouse. J Pharmacol Exp Ther 2004;308:1191-6.

157. Patsenker E, Popov Y, Stickel F, et al. Inhibition of integrin alphavbeta6 on cholangiocytes blocks transforming growth factor-beta activation and retards biliary fibrosis progression. Gastroenterology 2008;135:660-70.
158. Omenetti A, Porrello A, Jung Y, et al. Hedgehog signaling regulates epithelial-mesenchymal transition during biliary fibrosis in rodents and humans. J Clin Invest 2008;118:3331-42.
159. Philips GM, Chan IS, Swiderska M, et al. Hedgehog signaling antagonist promotes regression of both liver fibrosis and hepatocellular carcinoma in a murine model of primary liver cancer. PLoS One 2011;6:e23943.
160. Thabut D, Shah V. Intrahepatic angiogenesis and sinusoidal remodeling in chronic liver disease: new targets for the treatment of portal hypertension? J Hepatol 2010;53:976-80.
161. Tugues S, Fernandez-Varo G, Munoz-Luque J, et al. Antiangiogenic treatment with sunitinib ameliorates inflammatory infiltrate, fibrosis, and portal pressure in cirrhotic rats. Hepatology 2007;46:1919-26.
162. Karlmark KR, Zimmermann HW, Roderburg C, et al. The fractalkine receptor CX(3)CR1 protects against liver fibrosis by controlling differentiation and survival of infiltrating hepatic monocytes. Hepatology 2010;52:1769-82.
163. Sahin H, Borkham-Kamphorst E, Kuppe C, et al. Chemokine Cxcl9 attenuates liver fibrosis-associated angiogenesis in mice. Hepatology 2012;55:1610-9.
164. Barry-Hamilton V, Spangler R, Marshall D, et al. Allosteric inhibition of lysyl oxidase-like-2 impedes the development of a pathologic microenvironment. Nat Med 2010;16:1009-17.
165. Harrison SA, Fecht W, Brunt EM, et al. Orlistat for overweight subjects with nonalcoholic steatohepatitis: A randomized, prospective trial. Hepatology 2009;49:80-6.
166. Promrat K, Kleiner DE, Niemeier HM, et al. Randomized controlled trial testing the effects of weight loss on nonalcoholic steatohepatitis. Hepatology 2010;51:121-9.
167. Vilar-Gomez E, Martinez-Perez Y, Calzadilla-Bertot L, et al. Weight Loss Through Lifestyle Modification Significantly Reduces Features of Nonalcoholic Steatohepatitis. Gastroenterology 2015;149:367-78 e5; quiz e14-5.
168. Cariou B, Hanf R, Lambert-Porcheron S, et al. Dual peroxisome proliferator-activated receptor alpha/delta agonist GFT505 improves hepatic and peripheral

insulin sensitivity in abdominally obese subjects. Diabetes Care 2013;36:2923-30.

169. Belfort R, Harrison SA, Brown K, et al. A placebo-controlled trial of pioglitazone in subjects with nonalcoholic steatohepatitis. N Engl J Med 2006;355:2297-307.
170. Promrat K, Lutchman G, Uwaifo GI, et al. A pilot study of pioglitazone treatment for nonalcoholic steatohepatitis. Hepatology 2004;39:188-96.
171. Sanyal AJ, Chalasani N, Kowdley KV, et al. Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. N Engl J Med 2010;362:1675-85.
172. Porez G, Prawitt J, Gross B, et al. Bile acid receptors as targets for the treatment of dyslipidemia and cardiovascular disease. J Lipid Res 2012;53:1723-37.
173. Neuschwander-Tetri BA, Loomba R, Sanyal AJ, et al. Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. Lancet 2015;385:956-65.
174. Safadi R, Konikoff FM, Mahamid M, et al. The fatty acid-bile acid conjugate Aramchol reduces liver fat content in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2014;12:2085-91 e1.
175. Dongiovanni P, Petta S, Mannisto V, et al. Statin use and non-alcoholic steatohepatitis in at risk individuals. J Hepatol 2015;63:705-12.
176. Carbone LJ, Angus PW, Yeomans ND. Incretin-based therapies for the treatment of non-alcoholic fatty liver disease: A systematic review and meta-analysis. J Gastroenterol Hepatol 2016;31:23-31.
177. Sanyal AJ, Mofrad PS, Contos MJ, et al. A pilot study of vitamin E versus vitamin E and pioglitazone for the treatment of nonalcoholic steatohepatitis. Clin Gastroenterol Hepatol 2004;2:1107-15.
178. Lavine JE, Schwimmer JB, Van Natta ML, et al. Effect of vitamin E or metformin for treatment of nonalcoholic fatty liver disease in children and adolescents: the TONIC randomized controlled trial. JAMA 2011;305:1659-68.
179. Zein CO, Lopez R, Fu X, et al. Pentoxifylline decreases oxidized lipid products in nonalcoholic steatohepatitis: new evidence on the potential therapeutic mechanism. Hepatology 2012;56:1291-9.
180. Yano M, Kumada H, Kage M, et al. The long-term pathological evolution of chronic hepatitis C. Hepatology 1996;23:1334-40.

181. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. Hepatology 2003;38:1449-57.
182. Theodossi A, Skene AM, Portmann B, et al. Observer variation in assessment of liver biopsies including analysis by kappa statistics. Gastroenterology 1980;79:232-41.
183. Colloredo G, Guido M, Sonzogni A, et al. Impact of liver biopsy size on histological evaluation of chronic viral hepatitis: the smaller the sample, the milder the disease. J Hepatol 2003;39:239-44.
184. Brunetti E, Silini E, Pistorio A, et al. Coarse vs. fine needle aspiration biopsy for the assessment of diffuse liver disease from hepatitis C virus-related chronic hepatitis. J Hepatol 2004;40:501-6.
185. Sporea I, Popescu A, Sirli R. Why, who and how should perform liver biopsy in chronic liver diseases. World J Gastroenterol 2008;14:3396-402.
186. Knodell RG, Ishak KG, Black WC, et al. Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. Hepatology 1981;1:431-5.
187. Ishak K, Baptista A, Bianchi L, et al. Histological grading and staging of chronic hepatitis. J Hepatol 1995;22:696-9.
188. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. Hepatology 1996;24:289-93.
189. Scheuer PJ. Classification of chronic viral hepatitis: a need for reassessment. J Hepatol 1991;13:372-4.
190. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313-21.
191. Goodman ZD. Grading and staging systems for inflammation and fibrosis in chronic liver diseases. Journal of Hepatology 2007;47:598-607.
192. Brunt EM. Nonalcoholic steatohepatitis: definition and pathology. Semin Liver Dis 2001;21:3-16.
193. Sanyal A, Harrison SA, Ratziu V, et al. Changes in Fibrosis, but not the Nafld Activity Score (NAS), are Associated with Disease Progression in Patients with Nonalcoholic Steatohepatitis (NASH) and Advanced Fibrosis. Gastroenterology;152:S1055.

194. Stigliano R, Marelli L, Yu D, et al. Seeding following percutaneous diagnostic and therapeutic approaches for hepatocellular carcinoma. What is the risk and the outcome? Seeding risk for percutaneous approach of HCC. Cancer Treat Rev 2007;33:437-47.
195. Calvaruso V, Burroughs AK, Standish R, et al. Computer-assisted image analysis of liver collagen: relationship to Ishak scoring and hepatic venous pressure gradient. Hepatology 2009;49:1236-44.
196. Tsochatzis E, Bruno S, Isgro G, et al. Collagen proportionate area is superior to other histological methods for sub-classifying cirrhosis and determining prognosis. J Hepatol 2014;60:948-54.
197. Jangouk P, Turco L, Oliveira A, et al. Validating, deconstructing and refining Baveno criteria for ruling out high-risk varices in patients with compensated cirrhosis. Liver Int 2017.
198. Weinstein S, Obuchowski NA, Lieber ML. Clinical evaluation of diagnostic tests. AJR Am J Roentgenol 2005;184:14-9.
199. Weinstein MC, Fineberg HV. Clinical Decision Analysis (Saunders). 1980.
200. Mehta SH, Lau B, Afdhal NH, et al. Exceeding the limits of liver histology markers. J Hepatol 2009;50:36-41.
201. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. N Engl J Med 1978;299:926-30.
202. Obuchowski NA. An ROC-type measure of diagnostic accuracy when the gold standard is continuous-scale. Stat Med 2006;25:481-93.
203. Poynard T, Halfon P, Castera L, et al. Standardization of ROC curve areas for diagnostic evaluation of liver fibrosis markers based on prevalences of fibrosis stages. Clin Chem 2007;53:1615-22.
204. Fagan TJ. Letter: Nomogram for Bayes theorem. N Engl J Med 1975;293:257.
205. Piton A, Poynard T, Imbert-Bismut F, et al. Factors associated with serum alanine transaminase activity in healthy subjects: consequences for the definition of normal values, for selection of blood donors, and for patients with chronic hepatitis C. MULTIVIRC Group. Hepatology 1998;27:1213-9.
206. Guechot J, Laudat A, Loria A, et al. Diagnostic accuracy of hyaluronan and type III procollagen amino-terminal peptide serum assays as markers of liver fibrosis in chronic viral hepatitis C evaluated by ROC curve analysis. Clin Chem 1996;42:558-63.

207. Boeker KH, Haberkorn CI, Michels D, et al. Diagnostic potential of circulating TIMP-1 and MMP-2 as markers of liver fibrosis in patients with chronic hepatitis C. Clin Chim Acta 2002;316:71-81.
208. Lieber CS, Weiss DG, Paronetto F. Value of fibrosis markers for staging liver fibrosis in patients with precirrhotic alcoholic liver disease. Alcohol Clin Exp Res 2008;32:1031-9.
209. Nelson DR, Gonzalez-Peralta RP, Qian K, et al. Transforming growth factor-beta 1 in chronic hepatitis C. J Viral Hepat 1997;4:29-35.
210. Weissmann B, Meyer K, Sampson P, et al. Isolation of oligosaccharides enzymatically produced from hyaluronic acid. J Biol Chem 1954;208:417-29.
211. Noble PW. Hyaluronan and its catabolic products in tissue injury and repair. Matrix Biol 2002;21:25-9.
212. Arthur MJ, Iredale JP, Mann DA. Tissue inhibitors of metalloproteinases: role in liver fibrosis and alcoholic liver disease. Alcohol Clin Exp Res 1999;23:940-3.
213. Gerling B, Becker M, Waldschmidt J, et al. Elevated serum aminoterminal procollagen type-III-peptide parallels collagen accumulation in rats with secondary biliary fibrosis. J Hepatol 1996;25:79-84.
214. Koivisto H, Hietala J, Niemela O. An inverse relationship between markers of fibrogenesis and collagen degradation in patients with or without alcoholic liver disease. Am J Gastroenterol 2007;102:773-9.
215. Hakala BE, White C, Recklies AD. Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family. J Biol Chem 1993;268:25803-10.
216. Johansen JS, Moller S, Price PA, et al. Plasma YKL-40: a new potential marker of fibrosis in patients with alcoholic cirrhosis? Scand J Gastroenterol 1997;32:582-90.
217. Shackel NA, McGuinness PH, Abbott CA, et al. Novel differential gene expression in human cirrhosis detected by suppression subtractive hybridization. Hepatology 2003;38:577-88.
218. Johansen JS. Studies on serum YKL-40 as a biomarker in diseases with inflammation, tissue remodelling, fibroses and cancer. Dan Med Bull 2006;53:172-209.
219. Recklies AD, White C, Ling H. The chitinase 3-like protein human cartilage glycoprotein 39 (HC-gp39) stimulates proliferation of human connective-tissue

cells and activates both extracellular signal-regulated kinase- and protein kinase B-mediated signalling pathways. Biochem J 2002;365:119-26.

220. Johansen JS, Christoffersen P, Moller S, et al. Serum YKL-40 is increased in patients with hepatic fibrosis. J Hepatol 2000;32:911-20.
221. Rosenberg WM, Voelker M, Thiel R, et al. Serum markers detect the presence of liver fibrosis: a cohort study. Gastroenterology 2004;127:1704-13.
222. Cales P, Oberti F, Michalak S, et al. A novel panel of blood markers to assess the degree of liver fibrosis. Hepatology 2005;42:1373-81.
223. Imbert-Bismut F, Ratziu V, Pieroni L, et al. Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study. Lancet 2001;357:1069-75.
224. Adams LA, Bulsara M, Rossi E, et al. Hepascore: an accurate validated predictor of liver fibrosis in chronic hepatitis C infection. Clin Chem 2005;51:1867-73.
225. Mehta P, Ploutz-Snyder R, Nandi J, et al. Diagnostic accuracy of serum hyaluronic acid, FIBROSpect II, and YKL-40 for discriminating fibrosis stages in chronic hepatitis C. The American Journal of Gastroenterology 2008;103:928-36.
226. Guechot J, Laudat A, Loria A, et al. Diagnostic accuracy of hyaluronan and type III procollagen amino-terminal peptide serum assays as markers of liver fibrosis in chronic viral hepatitis C evaluated by ROC curve analysis. Clinical Chemistry 1996;42:558-63.
227. Walsh KM, Fletcher A, MacSween RN, et al. Basement membrane peptides as markers of liver disease in chronic hepatitis C. Journal of Hepatology 2000;32:325-30.
228. Walsh KM, Timms P, Campbell S, et al. Plasma levels of matrix metalloproteinase-2 (MMP-2) and tissue inhibitors of metalloproteinases -1 and -2 (TIMP-1 and TIMP-2) as noninvasive markers of liver disease in chronic hepatitis C: comparison using ROC analysis. Dig Dis Sci 1999;44:624-30.
229. Leroy V, Hilleret MN, Sturm N, et al. Prospective comparison of six non-invasive scores for the diagnosis of liver fibrosis in chronic hepatitis C. Journal of Hepatology 2007;46:775-82.

230. Sebastiani G, Vario A, Guido M, et al. Stepwise combination algorithms of non-invasive markers to diagnose significant fibrosis in chronic hepatitis C. Journal of Hepatology 2006;44:686-93.
231. Wai CT, Greenson JK, Fontana RJ, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology 2003;38:518-26.
232. Fontana RJ, Kleiner DE, Bilonick R, et al. Modeling hepatic fibrosis in African American and Caucasian American patients with chronic hepatitis C virus infection. Hepatology 2006;44:925-35.
233. Castera L, Le Bail B, Roudot-Thoraval F, et al. Early detection in routine clinical practice of cirrhosis and oesophageal varices in chronic hepatitis C: comparison of transient elastography (FibroScan) with standard laboratory tests and non-invasive scores. Journal of Hepatology 2009;50:59-68.
234. Islam S, Antonsson L, Westin J, et al. Cirrhosis in hepatitis C virus-infected patients can be excluded using an index of standard biochemical serum markers. Scandinavian Journal of Gastroenterology 2005;40:867-72.
235. Fontana RJ, Goodman ZD, Dienstag JL, et al. Relationship of serum fibrosis markers with liver fibrosis stage and collagen content in patients with advanced chronic hepatitis C. Hepatology 2008;47:789-98.
236. Sebastiani G, Vario A, Guido M, et al. Performance of noninvasive markers for liver fibrosis is reduced in chronic hepatitis C with normal transaminases. Journal of Viral Hepatitis 2008;15:212-8.
237. Forns X, Ampurdanes S, Llovet JM, et al. Identification of chronic hepatitis C patients without hepatic fibrosis by a simple predictive model. Hepatology 2002;36:986-92.
238. Koda M, Matunaga Y, Kawakami M, et al. FibroIndex, a practical index for predicting significant fibrosis in patients with chronic hepatitis C. Hepatology 2007;45:297-306.
239. Lok AS, Ghany MG, Goodman ZD, et al. Predicting cirrhosis in patients with hepatitis C based on standard laboratory tests: results of the HALT-C cohort. Hepatology 2005;42:282-92.
240. Vallet-Pichard A, Mallet V, Nalpas B, et al. FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. Hepatology 2007;46:32-6.

241. Adams LA, Bulsara M, Rossi E, et al. Hepascore: an accurate validated predictor of liver fibrosis in chronic hepatitis C infection. Clinical Chemistry 2005;51:1867-73.
242. Becker L, Salameh W, Sferruzza A, et al. Validation of hepascore, compared with simple indices of fibrosis, in patients with chronic hepatitis C virus infection in United States. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2009;7:696-701.
243. Halfon P, Bacq Y, De Muret A, et al. Comparison of test performance profile for blood tests of liver fibrosis in chronic hepatitis C. Journal of Hepatology 2007;46:395-402.
244. Patel K, Gordon SC, Jacobson I, et al. Evaluation of a panel of non-invasive serum markers to differentiate mild from moderate-to-advanced liver fibrosis in chronic hepatitis C patients. Journal of Hepatology 2004;41:935-42.
245. Zaman A, Rosen HR, Ingram K, et al. Assessment of FIBROSpect II to detect hepatic fibrosis in chronic hepatitis C patients. The American journal of medicine 2007;120:280 e9-14.
246. Rosenberg WMC, Voelker M, Thiel R, et al. Serum markers detect the presence of liver fibrosis: A cohort study. Gastroenterology 2004;127:1704-1713.
247. Parkes J, Guha IN, Roderick P, et al. Enhanced Liver Fibrosis (ELF) test accurately identifies liver fibrosis in patients with chronic hepatitis C. Journal of Viral Hepatitis 2011;18:23-31.
248. Cobbold JFL, Crossey MME, Colman P, et al. Optimal combinations of ultrasound-based and serum markers of disease severity in patients with chronic hepatitis C. Journal of Viral Hepatitis 2009;17:537-545.
249. Guha IN, Parkes J, Roderick P, et al. Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: Validating the European Liver Fibrosis Panel and exploring simple markers. Hepatology 2007;47:455-460.
250. Adams LA, George J, Bugianesi E, et al. Complex non-invasive fibrosis models are more accurate than simple models in non-alcoholic fatty liver disease. Journal of Gastroenterology and Hepatology 2011;26:1536-43.
251. Ratziu V, Massard J, Charlotte F, et al. Diagnostic value of biochemical markers (FibroTest-FibroSURE) for the prediction of liver fibrosis in patients with non-alcoholic fatty liver disease. BMC Gastroenterology 2006;6:6.

252. Angulo P, Hui JM, Marchesini G, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. Hepatology 2007;45:846-54.
253. Shah AG, Lydecker A, Murray K, et al. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2009;7:1104-12.
254. Cales P, Laine F, Boursier J, et al. Comparison of blood tests for liver fibrosis specific or not to NAFLD. Journal of Hepatology 2009;50:165-73.
255. Harrison SA, Oliver D, Arnold HL, et al. Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. Gut 2008;57:1441-7.
256. McPherson S, Stewart SF, Henderson E, et al. Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. Gut 2010;59:1265-9.
257. Hannivoort RA, Hernandez-Gea V, Friedman SL. Genomics and proteomics in liver fibrosis and cirrhosis. Fibrogenesis Tissue Repair 2012;5:1.
258. Ladaru A, Balanescu P, Stan M, et al. Candidate proteomic biomarkers for non-alcoholic fatty liver disease (steatosis and non-alcoholic steatohepatitis) discovered with mass-spectrometry: a systematic review. Biomarkers 2016;21:102-14.
259. Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. J Hepatol 2008;48:835-47.
260. Castera L, Foucher J, Bernard PH, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. Hepatology 2010;51:828-35.
261. Festi D, Schiumerini R, Marzi L, et al. Review article: the diagnosis of non-alcoholic fatty liver disease -- availability and accuracy of non-invasive methods. Aliment Pharmacol Ther 2013;37:392-400.
262. Castera L, Vergniol J, Foucher J, et al. Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. Gastroenterology 2005;128:343-50.

263. Ziol M, Handra-Luca A, Kettaneh A, et al. Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with chronic hepatitis C. Hepatology 2005;41:48-54.
264. Arena U, Vizzutti F, Abraldes JG, et al. Reliability of transient elastography for the diagnosis of advanced fibrosis in chronic hepatitis C. Gut 2008;57:1288-93.
265. Lupsor M, Badea R, Stefanescu H, et al. Analysis of histopathological changes that influence liver stiffness in chronic hepatitis C. Results from a cohort of 324 patients. J Gastrointestin Liver Dis 2008;17:155-63.
266. Wang JH, Changchien CS, Hung CH, et al. FibroScan and ultrasonography in the prediction of hepatic fibrosis in patients with chronic viral hepatitis. J Gastroenterol 2009;44:439-46.
267. Degos F, Perez P, Roche B, et al. Diagnostic accuracy of FibroScan and comparison to liver fibrosis biomarkers in chronic viral hepatitis: a multicenter prospective study (the FIBROSTIC study). J Hepatol 2010;53:1013-21.
268. Zarski JP, Sturm N, Guechot J, et al. Comparison of nine blood tests and transient elastography for liver fibrosis in chronic hepatitis C: the ANRS HCEP-23 study. J Hepatol 2012;56:55-62.
269. Yoneda M, Mawatari H, Fujita K, et al. Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with nonalcoholic fatty liver disease (NAFLD). Dig Liver Dis 2008;40:371-8.
270. Nobili V, Vizzutti F, Arena U, et al. Accuracy and reproducibility of transient elastography for the diagnosis of fibrosis in pediatric nonalcoholic steatohepatitis. Hepatology 2008;48:442-8.
271. Lupsor M, Badea R, Stefanescu H, et al. Performance of unidimensional transient elastography in staging non-alcoholic steatohepatitis. J Gastrointestin Liver Dis 2010;19:53-60.
272. Wong VW, Vergniol J, Wong GL, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. Hepatology 2010;51:454-62.
273. Gaia S, Carenzi S, Barilli AL, et al. Reliability of transient elastography for the detection of fibrosis in non-alcoholic fatty liver disease and chronic viral hepatitis. J Hepatol 2011;54:64-71.

274. Petta S, Di Marco V, Camma C, et al. Reliability of liver stiffness measurement in non-alcoholic fatty liver disease: the effects of body mass index. Aliment Pharmacol Ther 2011;33:1350-60.
275. Myers RP, Pomier-Layrargues G, Kirsch R, et al. Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients. Hepatology 2012;55:199-208.
276. Wong VW, Vergniol J, Wong GL, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. Am J Gastroenterol 2012;107:1862-71.
277. Rouviere O, Yin M, Dresner MA, et al. MR elastography of the liver: preliminary results. Radiology 2006;240:440-8.
278. Shi KQ, Fan YC, Pan ZZ, et al. Transient elastography: a meta-analysis of diagnostic accuracy in evaluation of portal hypertension in chronic liver disease. Liver Int 2013;33:62-71.
279. Sandahl TD, McGrail R, Moller HJ, et al. The macrophage activation marker sCD163 combined with markers of the Enhanced Liver Fibrosis (ELF) score predicts clinically significant portal hypertension in patients with cirrhosis. Aliment Pharmacol Ther 2016;43:1222-31.
280. Thabut D, Imbert-Bismut F, Cazals-Hatem D, et al. Relationship between the Fibrotest and portal hypertension in patients with liver disease. Aliment Pharmacol Ther 2007;26:359-68.
281. Robic MA, Procopet B, Metivier S, et al. Liver stiffness accurately predicts portal hypertension related complications in patients with chronic liver disease: a prospective study. J Hepatol 2011;55:1017-24.
282. Ngo Y, Munteanu M, Messous D, et al. A prospective analysis of the prognostic value of biomarkers (FibroTest) in patients with chronic hepatitis C. Clin Chem 2006;52:1887-96.
283. Irvine KM, Wockner LF, Shanker M, et al. The Enhanced liver fibrosis score is associated with clinical outcomes and disease progression in patients with chronic liver disease. Liver Int 2016;36:370-7.
284. Parkes J, Guha IN, Roderick P, et al. Enhanced Liver Fibrosis (ELF) test accurately identifies liver fibrosis in patients with chronic hepatitis C. J Viral Hepat 2011;18:23-31.

285. Corpechot C, Gaouar F, El Naggar A, et al. Baseline values and changes in liver stiffness measured by transient elastography are associated with severity of fibrosis and outcomes of patients with primary sclerosing cholangitis. Gastroenterology 2014;146:970-9; quiz e15-6.
286. Vergniol J, Boursier J, Coutzac C, et al. Evolution of noninvasive tests of liver fibrosis is associated with prognosis in patients with chronic hepatitis C. Hepatology 2014;60:65-76.
287. Poynard T, Ngo Y, Munteanu M, et al. Biomarkers of liver injury for hepatitis clinical trials: a meta-analysis of longitudinal studies. Antivir Ther 2010;15:617-31.
288. Lichtinghagen R, Pietsch D, Bantel H, et al. The Enhanced Liver Fibrosis (ELF) score: normal values, influence factors and proposed cut-off values. J Hepatol 2013;59:236-42.
289. Fagan KJ, Pretorius CJ, Horsfall LU, et al. ELF score >/=9.8 indicates advanced hepatic fibrosis and is influenced by age, steatosis and histological activity. Liver Int 2015;35:1673-81.
290. Castera L. Non-invasive assessment of liver fibrosis in chronic hepatitis C. Hepatol Int 2011;5:625-34.
291. Sebastiani G, Halfon P, Castera L, et al. SAFE biopsy: a validated method for large-scale staging of liver fibrosis in chronic hepatitis C. Hepatology 2009;49:1821-7.
292. Sebastiani G, Vario A, Guido M, et al. Stepwise combination algorithms of non-invasive markers to diagnose significant fibrosis in chronic hepatitis C. J Hepatol 2006;44:686-93.
293. Petta S, Vanni E, Bugianesi E, et al. The combination of liver stiffness measurement and NAFLD fibrosis score improves the noninvasive diagnostic accuracy for severe liver fibrosis in patients with nonalcoholic fatty liver disease. Liver Int 2015;35:1566-73.
294. Hernaez R, Lazo M, Bonekamp S, et al. Diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver: a meta-analysis. Hepatology 2011;54:1082-90.
295. Poynard T, Ratziu V, Naveau S, et al. The diagnostic value of biomarkers (SteatoTest) for the prediction of liver steatosis. Comp Hepatol 2005;4:10.

296. Bedogni G, Bellentani S, Miglioli L, et al. The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. BMC Gastroenterol 2006;6:33.
297. Lee JH, Kim D, Kim HJ, et al. Hepatic steatosis index: a simple screening tool reflecting nonalcoholic fatty liver disease. Dig Liver Dis 2010;42:503-8.
298. Bedogni G, Kahn HS, Bellentani S, et al. A simple index of lipid overaccumulation is a good marker of liver steatosis. BMC Gastroenterol 2010;10:98.
299. Otgonsuren M, Estep MJ, Hossain N, et al. Single non-invasive model to diagnose non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH). J Gastroenterol Hepatol 2014;29:2006-13.
300. Kotronen A, Peltonen M, Hakkarainen A, et al. Prediction of non-alcoholic fatty liver disease and liver fat using metabolic and genetic factors. Gastroenterology 2009;137:865-72.
301. Oshima RG. Apoptosis and keratin intermediate filaments. Cell Death Differ 2002;9:486-92.
302. Feldstein AE, Wieckowska A, Lopez AR, et al. Cytokeratin-18 fragment levels as noninvasive biomarkers for nonalcoholic steatohepatitis: a multicenter validation study. Hepatology 2009;50:1072-8.
303. Kwok R, Tse YK, Wong GL, 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;39:254-69.
304. Cusi K, Chang Z, Harrison S, et al. Limited value of plasma cytokeratin-18 as a biomarker for NASH and fibrosis in patients with non-alcoholic fatty liver disease. J Hepatol 2014;60:167-74.
305. Feldstein AE, Canbay A, Angulo P, et al. Hepatocyte apoptosis and fas expression are prominent features of human nonalcoholic steatohepatitis. Gastroenterology 2003;125:437-43.
306. Tamimi TI, Elgouhari HM, Alkhouri N, et al. An apoptosis panel for nonalcoholic steatohepatitis diagnosis. J Hepatol 2011;54:1224-9.
307. Poynard T, Ratziu V, Charlotte 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.

308. Campos GM, Bambha K, Vittinghoff E, et al. A clinical scoring system for predicting nonalcoholic steatohepatitis in morbidly obese patients. Hepatology 2008;47:1916-23.
309. Palekar NA, Naus R, Larson SP, et al. Clinical model for distinguishing nonalcoholic steatohepatitis from simple steatosis in patients with nonalcoholic fatty liver disease. Liver Int 2006;26:151-6.
310. Machado MV, Coutinho J, Carepa F, et al. How adiponectin, leptin, and ghrelin orchestrate together and correlate with the severity of nonalcoholic fatty liver disease. Eur J Gastroenterol Hepatol 2012;24:1166-72.
311. Manousou P, Kalambokis G, Grillo F, et al. Serum ferritin is a discriminant marker for both fibrosis and inflammation in histologically proven non-alcoholic fatty liver disease patients. Liver Int 2011;31:730-9.
312. Feldstein AE, Lopez R, Tamimi TA, et al. Mass spectrometric profiling of oxidized lipid products in human nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. J Lipid Res 2010;51:3046-54.
313. Alkhouri N, Berk M, Yerian L, et al. OxNASH score correlates with histologic features and severity of nonalcoholic fatty liver disease. Dig Dis Sci 2014;59:1617-24.
314. Sumida Y, Yoneda M, Hyogo H, et al. A simple clinical scoring system using ferritin, fasting insulin, and type IV collagen 7S for predicting steatohepatitis in nonalcoholic fatty liver disease. J Gastroenterol 2011;46:257-68.
315. Marchesini G, Bugianesi E, Forlani G, et al. Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. Hepatology 2003;37:917-23.
316. Rafiq N, Bai C, Fang Y, et al. Long-term follow-up of patients with nonalcoholic fatty liver. Clin Gastroenterol Hepatol 2009;7:234-8.
317. Matteoni CA, Younossi ZM, Gramlich T, et al. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. Gastroenterology 1999;116:1413-9.
318. Brunt EM. Pathology of nonalcoholic steatohepatitis. Hepatol Res 2005;33:68-71.
319. Adams LA, Angulo P, Lindor KD. Nonalcoholic fatty liver disease. CMAJ 2005;172:899-905.
320. Ekstedt M, Franzen LE, Mathiesen UL, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. Hepatology 2006;44:865-73.

321. Thampanitchawong P, Piratvisuth T. Liver biopsy:complications and risk factors. World J Gastroenterol 1999;5:301-304.
322. Miller MH, Ferguson MA, Dillon JF. Systematic review of performance of non-invasive biomarkers in the evaluation of non-alcoholic fatty liver disease. Liver Int 2011;31:461-73.
323. Guha IN, Parkes J, Roderick P, et al. Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: Validating the European Liver Fibrosis Panel and exploring simple markers. Hepatology 2008;47:455-60.
324. Younossi ZM, Jarrar M, Nugent C, et al. A novel diagnostic biomarker panel for obesity-related nonalcoholic steatohepatitis (NASH). Obes Surg 2008;18:1430-7.
325. Montazeri G, Estakhri A, Mohamadnejad M, et al. Serum hyaluronate as a non-invasive marker of hepatic fibrosis and inflammation in HBeAg-negative chronic hepatitis B. BMC Gastroenterol 2005;5:32.
326. Murphy FR, Issa R, Zhou X, et al. Inhibition of apoptosis of activated hepatic stellate cells by tissue inhibitor of metalloproteinase-1 is mediated via effects on matrix metalloproteinase inhibition: implications for reversibility of liver fibrosis. J Biol Chem 2002;277:11069-76.
327. Ling H, Recklies AD. The chitinase 3-like protein human cartilage glycoprotein 39 inhibits cellular responses to the inflammatory cytokines interleukin-1 and tumour necrosis factor-alpha. Biochem J 2004;380:651-9.
328. Wieckowska A, Zein NN, Yerian LM, et al. In vivo assessment of liver cell apoptosis as a novel biomarker of disease severity in nonalcoholic fatty liver disease. Hepatology 2006;44:27-33.
329. Jaskiewicz K, Rzepko R, Sledzinski Z. Fibrogenesis in fatty liver associated with obesity and diabetes mellitus type 2. Dig Dis Sci 2008;53:785-8.
330. Brunt EM. Nonalcoholic steatohepatitis: pathologic features and differential diagnosis. Semin Diagn Pathol 2005;22:330-8.
331. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837-45.
332. Horslev-Petersen K, Bentsen KD, Halberg P, et al. Connective tissue metabolites in serum as markers of disease activity in patients with rheumatoid arthritis. Clin Exp Rheumatol 1988;6:129-34.

333. Yilmaz Y, Dolar E, Ulukaya E, et al. Soluble forms of extracellular cytokeratin 18 may differentiate simple steatosis from nonalcoholic steatohepatitis. World J Gastroenterol 2007;13:837-44.
334. Verhaegh P, Bavalia R, Winkens B, et al. Non-invasive Tests Do Not Accurately Differentiate Non-alcoholic Steatohepatitis From Simple Steatosis: A Systematic Review and Meta-Analysis. Clin Gastroenterol Hepatol 2017.
335. Fan JG, Zhu J, Li XJ, et al. Prevalence of and risk factors for fatty liver in a general population of Shanghai, China. J Hepatol 2005;43:508-14.
336. Chou R, Wasson N. Blood tests to diagnose fibrosis or cirrhosis in patients with chronic hepatitis C virus infection: a systematic review. Ann Intern Med 2013;158:807-20.
337. Suzuki A, Angulo P, Lymp J, et al. Hyaluronic acid, an accurate serum marker for severe hepatic fibrosis in patients with non-alcoholic fatty liver disease. Liver Int 2005;25:779-86.
338. Ratziu V, Giral P, Charlotte F, et al. Liver fibrosis in overweight patients. Gastroenterology 2000;118:1117-23.
339. Bonacini M, Hadi G, Govindarajan S, et al. Utility of a discriminant score for diagnosing advanced fibrosis or cirrhosis in patients with chronic hepatitis C virus infection. Am J Gastroenterol 1997;92:1302-4.
340. Cales P, Laine F, Boursier J, et al. Comparison of blood tests for liver fibrosis specific or not to NAFLD. J Hepatol 2009;50:165-73.
341. Shah AG, Lydecker A, Murray K, et al. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2009;7:1104-12.
342. Lambert J, Halfon P, Penaranda G, et al. How to measure the diagnostic accuracy of noninvasive liver fibrosis indices: the area under the ROC curve revisited. Clin Chem 2008;54:1372-8.
343. Tanwar S, Ellis E, Parkes J, et al. 125 Changes In ELF Score During Treatment With Pegylated Interferon And Silymarin Are Associated With Progression Of Histological Liver Disease In Non-Responder Patients With Hepatitis C. J Hepatol;54:S55.
344. Hagstrom H, Nasr P, Ekstedt M, et al. Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. J Hepatol 2017.

345. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.
346. Marshall RJ. The Predictive Value of Simple Rules for Combining Two Diagnostic Tests. Biometrics 1989;45:1213-1222.
347. Macaskill P, Walter SD, Irwig L, et al. Assessing the gain in diagnostic performance when combining two diagnostic tests. Stat Med 2002;21:2527-46.
348. Tanwar S, Trembling PM, Thorburn D, et al. Direct serum markers are more accurate than simple marker panels for the detection of fibrosis in nonalcoholic fatty liver disease (NAFLD). J Hepatol 2012;56:S423.
349. Poynard T, Muntenau M, Morra R, et al. Methodological aspects of the interpretation of non-invasive biomarkers of liver fibrosis: a 2008 update. Gastroenterol Clin Biol 2008;32:8-21.
350. Tillmann HL. Hepatitis C infection and presence of advanced fibrosis: wait or treat? Why wait? There is no time to lose, is there? J Hepatol 2013;58:412-4.
351. Regev A, Berho M, Jeffers LJ, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. Am J Gastroenterol 2002;97:2614-8.
352. Ocker M, Neureier D, Schuppan D, et al. The combination of silymarin and pegylated interferon alpha2b attenuates the progression of liver fibrosis in prior nonresponders and relapsers with chronic hepatitis C infection - the PROFI C trial. Z Gastroenterol 2008;46:P4_07.
353. Churg J, Prado A. A rapid Mallory trichrome stain (chromotrope-aniline blue). AMA Arch Pathol 1956;62:505-6.
354. Barikbin R, Neureiter D, Wirth J, et al. Induction of heme oxygenase 1 prevents progression of liver fibrosis in Mdr2 knockout mice. Hepatology 2012;55:553-62.
355. Daniel HD, Grant PR, Garson JA, et al. Quantitation of hepatitis C virus using an in-house real-time reverse transcriptase polymerase chain reaction in plasma samples. Diagn Microbiol Infect Dis 2008;61:415-20.
356. Stuyver L, Wyseur A, van Arnhem W, et al. Second-generation line probe assay for hepatitis C virus genotyping. J Clin Microbiol 1996;34:2259-66.
357. Davidson F, Simmonds P, Ferguson JC, et al. Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5' non-coding region. J Gen Virol 1995;76 (Pt 5):1197-204.

358. Sheth SG, Flamm SL, Gordon FD, et al. AST/ALT ratio predicts cirrhosis in patients with chronic hepatitis C virus infection. Am J Gastroenterol 1998;93:44-8.
359. Snyder N, Gajula L, Xiao SY, et al. APRI: an easy and validated predictor of hepatic fibrosis in chronic hepatitis C. J Clin Gastroenterol 2006;40:535-42.
360. Report to the World Health Organisation, 1957-1960. Sydney, N.S.W.: V.C.N. Blight, Govt. Printer, 1961.
361. Cales P, Halfon P, Batisse D, et al. Comparison of liver fibrosis blood tests developed for HCV with new specific tests in HIV/HCV co-infection. J Hepatol 2010;53:238-44.
362. Patel K, Gordon SC, Jacobson I, et al. Evaluation of a panel of non-invasive serum markers to differentiate mild from moderate-to-advanced liver fibrosis in chronic hepatitis C patients. J Hepatol 2004;41:935-42.
363. McHutchison JG, Blatt LM, de Medina M, et al. Measurement of serum hyaluronic acid in patients with chronic hepatitis C and its relationship to liver histology. Consensus Interferon Study Group. J Gastroenterol Hepatol 2000;15:945-51.
364. Leroy V, Hilleret MN, Sturm N, et al. Prospective comparison of six non-invasive scores for the diagnosis of liver fibrosis in chronic hepatitis C. J Hepatol 2007;46:775-82.
365. Veillon P, Gallois Y, Moal V, et al. Assessment of new hyaluronic acid assays and their impact on FibroMeter scores. Clin Chim Acta 2011;412:347-52.
366. Davis GL, Albright JE, Cook SF, et al. Projecting future complications of chronic hepatitis C in the United States. Liver Transpl 2003;9:331-8.
367. Wise M, Bialek S, Finelli L, et al. Changing trends in hepatitis C-related mortality in the United States, 1995-2004. Hepatology 2008;47:1128-35.
368. Cales P, Boursier J, Bertrais S, et al. Optimization and robustness of blood tests for liver fibrosis and cirrhosis. Clin Biochem 2010;43:1315-22.
369. Castera L. Transient elastography and other noninvasive tests to assess hepatic fibrosis in patients with viral hepatitis. J Viral Hepat 2009;16:300-14.
370. Castera L. Acoustic radiation force impulse imaging: a new technology for the noninvasive assessment of liver fibrosis? J Gastrointestin Liver Dis 2009;18:411-2.

371. Huwart L, Sempoux C, Salameh N, et al. Liver fibrosis: noninvasive assessment with MR elastography versus aspartate aminotransferase-to-platelet ratio index. Radiology 2007;245:458-66.
372. Bolanos-Meade J, Lopez-Arvizu C. Histologic improvement of fibrosis in patients with hepatitis C and sustained response to interferon therapy. Ann Intern Med 2000;133:312.
373. Boigk G, Stroedter L, Herbst H, et al. Silymarin retards collagen accumulation in early and advanced biliary fibrosis secondary to complete bile duct obliteration in rats. Hepatology 1997;26:643-9.
374. Lieber CS, Leo MA, Cao Q, et al. Silymarin retards the progression of alcohol-induced hepatic fibrosis in baboons. J Clin Gastroenterol 2003;37:336-9.
375. Jia JD, Bauer M, Cho JJ, et al. Antifibrotic effect of silymarin in rat secondary biliary fibrosis is mediated by downregulation of procollagen alpha1(I) and TIMP-1. J Hepatol 2001;35:392-8.
376. Obuchowski NA. Estimating and comparing diagnostic tests' accuracy when the gold standard is not binary. Acad Radiol 2005;12:1198-204.
377. Tanwar S, Trembling PM, Hogan BJ, et al. Noninvasive markers of liver fibrosis: on-treatment changes of serum markers predict the outcome of antifibrotic therapy. Eur J Gastroenterol Hepatol 2017;29:289-296.
378. Schuppan D, Pinzani M. Anti-fibrotic therapy: lost in translation? J Hepatol 2012;56 Suppl 1:S66-74.
379. Fontana RJ, Dienstag JL, Bonkovsky HL, et al. Serum fibrosis markers are associated with liver disease progression in non-responder patients with chronic hepatitis C. Gut 2010;59:1401-9.
380. Cales P, Zarski JP, Marc Chapplain J, et al. Fibrosis progression under maintenance interferon in hepatitis C is better detected by blood test than liver morphometry. J Viral Hepat 2012;19:e143-53.
381. Poynard T, Ngo Y, Marcellin P, et al. Impact of adefovir dipivoxil on liver fibrosis and activity assessed with biochemical markers (FibroTest-ActiTest) in patients infected by hepatitis B virus. J Viral Hepat 2009;16:203-13.
382. Ishibashi K, Kashiwagi T, Ito A, et al. Predictive factors for efficacy of interferon therapy in chronic hepatitis type C. Hepatogastroenterology 1995;42:535-41.

383. Nojgaard C, Johansen JS, Krarup HB, et al. Effect of antiviral therapy on markers of fibrogenesis in patients with chronic hepatitis C. Scand J Gastroenterol 2003;38:659-65.
384. Patel K, Benhamou Y, Yoshida EM, et al. An independent and prospective comparison of two commercial fibrosis marker panels (HCV FibroSURE and FIBROSpect II) during albinterferon alfa-2b combination therapy for chronic hepatitis C. J Viral Hepat 2009;16:178-86.
385. Martinez SM, Fernandez-Varo G, Gonzalez P, et al. Assessment of liver fibrosis before and after antiviral therapy by different serum marker panels in patients with chronic hepatitis C. Aliment Pharmacol Ther 2011;33:138-48.
386. Poynard T, Zoulim F, Ratziu V, et al. Longitudinal assessment of histology surrogate markers (FibroTest-ActiTest) during lamivudine therapy in patients with chronic hepatitis B infection. Am J Gastroenterol 2005;100:1970-80.
387. Poynard T, Colombo M, Bruix J, et al. Peginterferon alfa-2b and ribavirin: effective in patients with hepatitis C who failed interferon alfa/ribavirin therapy. Gastroenterology 2009;136:1618-28 e2.
388. Ryder SD, Irving WL, Jones DA, et al. Progression of hepatic fibrosis in patients with hepatitis C: a prospective repeat liver biopsy study. Gut 2004;53:451-5.
389. Parkes J, Roderick P, Harris S, et al. Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. Gut 2010;59:1245-51.
390. Schuppan D, Kim YO. Evolving therapies for liver fibrosis. J Clin Invest 2013;123:1887-901.
391. Yoon YJ, Friedman SL, Lee YA. Antifibrotic Therapies: Where Are We Now? Semin Liver Dis 2016;36:87-98.
392. Choo QL, Kuo G, Weiner AJ, et al. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 1989;244:359-62.
393. Hoofnagle JH, Seeff LB. Peginterferon and ribavirin for chronic hepatitis C. N Engl J Med 2006;355:2444-51.
394. Strader DB, Wright T, Thomas DL, et al. Diagnosis, management, and treatment of hepatitis C. Hepatology 2004;39:1147-71.
395. Fried MW, Shiffman ML, Reddy KR, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med 2002;347:975-82.

396. Manns MP, McHutchison JG, Gordon SC, et al. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. Lancet 2001;358:958-65.
397. Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 2006;43:1317-25.
398. D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. J Hepatol 2006;44:217-31.
399. Gurusamy KS, Tsochatzis E, Xirouchakis E, et al. Antiviral therapy for recurrent liver graft infection with hepatitis C virus. Cochrane Database Syst Rev 2010:CD006803.
400. Shiffman ML, Di Bisceglie AM, Lindsay KL, et al. Peginterferon Alfa-2a and ribavirin in patients with chronic hepatitis C who have failed prior treatment. Gastroenterology 2004;126:1015-1023.
401. Tanwar S, Trembling PM, Guha IN, et al. Validation of terminal peptide of procollagen III for the detection and assessment of nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease. Hepatology 2013;57:103-11.
402. Nielsen MJ, Nedergaard AF, Sun S, et al. The neo-epitope specific PRO-C3 ELISA measures true formation of type III collagen associated with liver and muscle parameters. Am J Transl Res 2013;5:303-15.
403. Leeming DJ, Grove JI, Kaye P, et al. THU-351 - Estimation of serum “true collagen type III formation” (Pro-C3) levels as a marker of non-alcoholic steatohepatitis in a prospective cohort. J Hepatol 2017;66:S154.
404. Tanwar S, Trembling PM, Thorburn D, et al. 1076 Direct Serum Markers Are More Accurate Than Simple Marker Panels For The Detection Of Fibrosis In Nonalcoholic Fatty Liver Disease (NAFLD). J Hepatol;56:S423.
405. Hagstrom H, Nasr P, Ekstedt M, et al. Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. J Hepatol 2017;67:1265-1273.
406. Tanwar S. An algorithm combining direct and indirect liver fibrosis tests enhances diagnostic performance for the detection of advanced fibrosis (F3-4) in NAFLD minimising the need for liver biopsy Hepatology 2012;56:35A-88A.

407. Srivastava A, Jong S, Gola A, et al. A probabilistic decision model using non-invasive fibrosis markers in Primary Care NAFLD pathways predicts increased cirrhosis detection rates and reduced overall healthcare expenditure. Hepatology 2015;62.
408. Srivastava A, Gailer R, Demma S, et al. Primary Care Sequential Use of FIB-4 and the Enhanced Liver Fibrosis Test to Stratify Patients with Non-alcoholic Fatty Liver Acide Doubles Cirrhosis Detection and Reduces Referrals of Patients with Mild Disease. J Hepatol;64:S474-S475.
409. Newsome PN, Cramb R, Davison SM, et al. Guidelines on the management of abnormal liver blood tests. Gut 2018;67:6-19.
410. Boursier J, de Ledinghen V, Leroy V, et al. A stepwise algorithm using an at-a-glance first-line test for the non-invasive diagnosis of advanced liver fibrosis and cirrhosis. J Hepatol 2017;66:1158-1165.
411. Sanyal A, Harrison S, Ratziu V, et al. Changes in fibrosis, but not the NAFLD Activity Score (NAS), are associated with disease progression in patients with nonalcoholic steatohepatitis (NASH) and advanced fibrosis. J Hepatol;66:S2-S3.
412. Vergniol J, Foucher J, Terrebonne E, et al. Noninvasive tests for fibrosis and liver stiffness predict 5-year outcomes of patients with chronic hepatitis C. Gastroenterology 2011;140:1970-9, 1979 e1-3.
413. Selby PJ, Banks RE, Gregory W, et al. A Multi-Centre Programme into Methods for the Evaluation of Biomarkers Suitable for Use in Patients with Kidney and Liver Diseases and a Randomised Controlled Trial for Their Use, 2017.
414. Kennedy OJ, Parkes J, Tanwar S, et al. The Enhanced Liver Fibrosis (ELF) Panel: Analyte Stability Under Common Sample Storage Conditions Used in Clinical Practice. The Journal of Applied Laboratory Medicine: An AACC Publication 2017;1:720-728.
415. Islam S, Antonsson L, Westin J, et al. Cirrhosis in hepatitis C virus-infected patients can be excluded using an index of standard biochemical serum markers. Scand J Gastroenterol 2005;40:867-72.
416. Kelleher TB, Mehta SH, Bhaskar R, et al. Prediction of hepatic fibrosis in HIV/HCV co-infected patients using serum fibrosis markers: the SHASTA index. J Hepatol 2005;43:78-84.