The Role of MRI in Stratifying and Evaluating Chronic Liver Disease

By

Peter John Eddowes

MRCP(UK) (Gastroenterology), MBChB (with honours)

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF MEDICINE

Centre for Liver Research

Institute of Immunology and Immunotherapy

College of Medical and Dental Sciences

University of Birmingham

October 2017

ABSTRACT

Chronic liver disease is a highly prevalent condition associated with significant morbidity and mortality. There is need for clinicians to stratify chronic liver disease and for researchers to define meaningful study endpoints. Currently this is often reliant on liver biopsy histology, which is known to be a flawed gold standard. There is a need to develop novel, non-invasive techniques for the evaluation of chronic liver disease that are accurate and reliable.

In this thesis I have demonstrated that multiparametric MRI can stage hepatic fibrosis in an unselected cohort with performance comparable to existing non-invasive fibrosis markers. The assessment of fibrosis is however confounded by inflammation. The sensitivity of multiparametric MRI to inflammation allows the differentiation of simple steatosis and NASH but in a non-alcoholic fatty liver disease (NAFLD) cohort, multiparametric MRI fails to predict fibrosis stage. Evaluating NAFLD with magnetic resonance spectroscopy has shown that this technique is feasible and that lipidomic differences can be demonstrated in patients with NAFLD. Exploring the role of multiparametric MRI in primary sclerosing cholangitis (PSC) has demonstrated a characteristic pattern in the distribution of corrected T1 in PSC suggesting that multiparametric MRI may have a role in its diagnosis and evaluation.

DEDICATION

My thesis is dedicated to my wife Hannah. She has supported me tirelessly through the writing of this thesis. Without her patience and understanding this would not have been possible.

ACKNOWLEDGEMENTS

There are very many people to whom I am grateful for their advice and guidance during my research. In particular I must mention my supervisors Prof Gideon Hirschfield and Prof Phil Newsome for ensuring I remained on the right track. I am grateful to my friends in the Centre for Liver Research for their support and ensuring my serum caffeine never dipped too far. I am grateful to Chris Weston and Stuart Curbishly for teaching me how research works, Miroslava Blahova for teaching me how the lab works and James Hodson for teaching me how statistics work. My understanding of MRI is entirely dependent on Nigel Davies, Rob Flintham and Roman Wesolowski from the medical physics department at QEHB. I am enormously grateful to them as, without their help, the MRI projects would not have succeeded. There are countless other colleagues from QEHB who were continually helpful during my research. Tracy Stewart in radiology, Carmel Maguire in LOPD, Jane Herbert in MRI and Val Adkins in pathology have all gone out of their way to make my life easier. Finally I remain grateful to the patients and volunteers who have given their time to take part in my studies; without them I would have achieved nothing.

DISCLOSURES

Links to published research

The work in Chapter 3 on "cT1 as a marker of fibrosis in chronic liver disease" has been submitted for publication to Scientific Reports. Dr Peter Eddowes is a joint first author.

The work in Chapter 4 on the "evaluation of multiparametric magnetic resonance imaging for the assessment of non-alcoholic fatty liver disease" has been published in Alimentary Pharmacology and Therapeutics.[1] Dr Peter Eddowes is a joint first author.

The work in Chapter 6 on the correlation between non-invasive markers in PSC has been presented as a poster at the EASL International Liver Congress, April 2016. Dr Peter Eddowes is the first author

Funding of the work

The work described in this thesis was funded by an Innovate-UK (formally the Technology Strategy Board) grant (grant number: 101679) with additional support from the NIHR Birmingham Liver Biomedical Research Unit (now the NIHR Birmingham Biomedical Research Centre). The grant was secured jointly by The University of Birmingham, The University of Edinburgh and Perspectum Diagnostics ltd. (Oxford, UK). The purpose of the grant was to two fold. Firstly to allow Perspectum Diagnostics to develop Liver*Multiscan*™, a software product for the analysis of MRI data and secondly for the university partners to

design and run clinical studies to investigate Liver*Multiscan*™ as a biomarker of liver disease. At the time of the grant application three clinical studies were planned. These were studies of Liver*Multiscan*™ as a biomarker of liver fibrosis using liver biopsy histology (Section 3.3) and hepatectomy specimens (Section 3.4) as the reference standard and a study of the use of Liver*Multiscan*™ in primary sclerosing cholangitis (Chapter 6).

Contributions to this work

Although these studies were envisaged before his involvement in the project Dr Peter Eddowes designed, wrote the protocols and gained ethical approval for all the projects described in this this thesis. He conducted study visits at Queen Elizabeth Hospitals Birmingham (QEHB) and collected data. Study visits and data collection at the Royal Infirmary Edinburgh was conducted by Dr Natasha McDonald (clinical research fellow). Histological assessment of prepared slides was performed by Prof Stefan Hübscher, Dr Desley Neil, Dr Rachel Brown, Dr Tim Kendall and Dr Owen Cain. Collagen proportionate area (CPA) staining and analysis for the two centre study was done by Dr Natasha McDonald under the guidance of Dr T Kendall. For the explant study, histology specimens were prepared, cut and stained by Dr Gary Reynolds. CPA analysis was performed by Dr Peter Eddowes. Magnetic resonance spectroscopy data were analysed by Mr Robert Flintham (clinical scientist at University Hospitals Birmingham). Dr Eddowes performed statistical analysis of all data under the guidance of Mr James Hodson (bio-statistician at the University Hospitals Birmingham). Prof Hirschfield and Dr Fallowfield provided critical appraisal of the draft manuscript.

The role of Perspectum Diagnostics ltd.

The role of Perspectum Diagnostics ltd in the clinical studies described in this thesis was the analysis of multiparametric MRI data to generate cT1 values. During the studies MRI scans were performed at the study sites and raw data sent to Perspectum Diagnostics. Staff at Perspectum Diagnostics analysed the data using Liver*Multiscan*^{TM} and generated data that were sent back to the universities for analysis. Perspectum Diagnostics played no further part in the conduct of the studies or in the analysis or interpretation of data. This service was provided free of charge by Perspectum Diagnostics.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: GENERAL INTRODUCTION AND REVIEW OF NON-INVASIVE BIOMARKERS IN THE STAGING OF HEPATIC FIBROSIS

1.1 Preamble

Chronic liver disease is a major public health problem worldwide with significant mortality. In developed countries it is the 5th biggest killer of adults.[2] The burden of liver disease is focused on a younger age group than many other leading causes of death making liver disease the 3rd largest cause of death during working age, surpassed by ischaemic heart disease and self-harm only.[2] Not only is chronic liver disease highly prevalent, it is the only major cause of death in the UK that is increasing in incidence.[2] Figure 1.1-1 shows that from 1970 to 2005 liver disease has seen a 500% increase in its standardised mortality rate while all other disease groups saw a steady decline year on year.

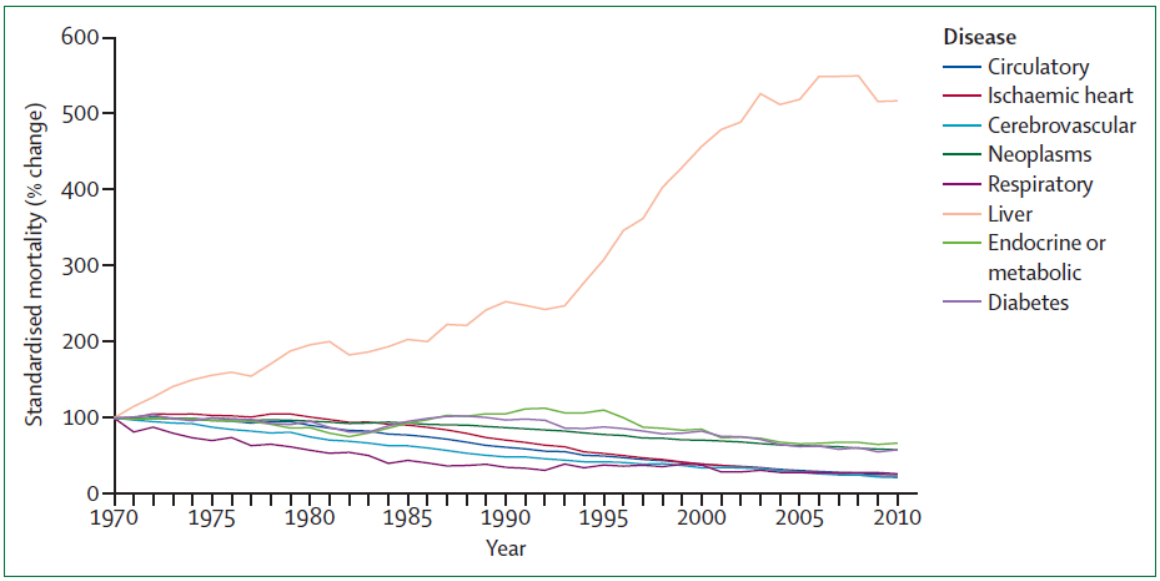


Figure 1.1-1: Rates of death from liver disease have increased since 1970 in contrast to other leading causes of death. Take from: The Lancet Commissions report: Addressing liver Disease in the UK, The Lancet, 2014.[2]

This increase in the prevalence of chronic liver disease is expected to continue over the next decade due to several factors. Firstly, the epidemic of obesity, insulin resistance and type two diabetes sweeping the developed world is leading to an increase in non-alcoholic fatty liver disease (NAFLD) and secondly, despite emerging therapies being highly effective, chronic viral hepatitis remains a public health challenge. The large pool of unrecognised infection in the community is projected to lead to an increase in cirrhosis secondary to chronic viral hepatitis until the 2020's.[3] However, the impact of NAFLD and viral hepatitis, although significant, is dwarfed by the consequences of harmful alcohol consumption. It is estimated that ¾ of all deaths from liver disease are as a result of alcohol excess and it is noteworthy that this figure does not include other harmful effects of alcohol excess including pancreatitis, injuries sustained while intoxicated, or the effects of alcohol excess on cardiovascular health and the incidence of certain cancers.[2]

Although the mechanisms of liver injury differ between the above mentioned aetiologies, the final common pathway of most chronic liver disease is progressive hepatic fibrosis. It is predominantly fibrosis that drives the clinical sequelae of chronic liver disease including portal hypertension, progressive hepatic failure and the development of hepatocellular carcinoma (HCC). Thus the presence of hepatic fibrosis is a strong predictor of morbidity and mortality.

The clear importance of fibrosis in the development of adverse clinical outcomes in liver disease means that one facet of the challenge to reduce death from liver disease is the identification, staging and management of patients with hepatic fibrosis. In this thesis,

strategies for the identification and staging of chronic liver disease will be discussed in several clinical situations with a focus on the assessment of hepatic fibrosis.

1.2 Hepatic fibrosis

1.2.1 Pathogenesis of fibrosis

Fibrosis is a pathological reaction to liver injury that occurs when severe or repeated injury overwhelms the liver's normal repair mechanisms. Microscopically the changes seen in fibrosis include hepatocyte death, liver infiltration by inflammatory cells, activation of stromal cells and expansion of the extracellular matrix (ECM).[4-6] The pathogenesis of these changes is complex but a summary of key points is outlined in Figure 1.2-1. In health, the amount of ECM present in the liver is finely controlled and there is balance between deposition and breakdown. Liver injury leads to cellular damage and the release of inflammatory cytokines such as platelet-derived growth factor and transforming growth factor beta (TGF-β). T-lymphocytes are recruited to the liver and produce further inflammatory mediators such as interferon gamma and interleukin-6. These inflammatory cytokines cause native hepatic stellate cells (HSC) to activate and migrate to sites of injury where they deposit excess ECM. In addition to increased deposition of ECM there is inhibition of the breakdown of ECM through the action of TGF-β on activated myofibroblasts. Matrix metalloproteases, which break down ECM, are inhibited by the secretion of tissue inhibitors of metalloprotease thus preventing the breakdown of ECM.

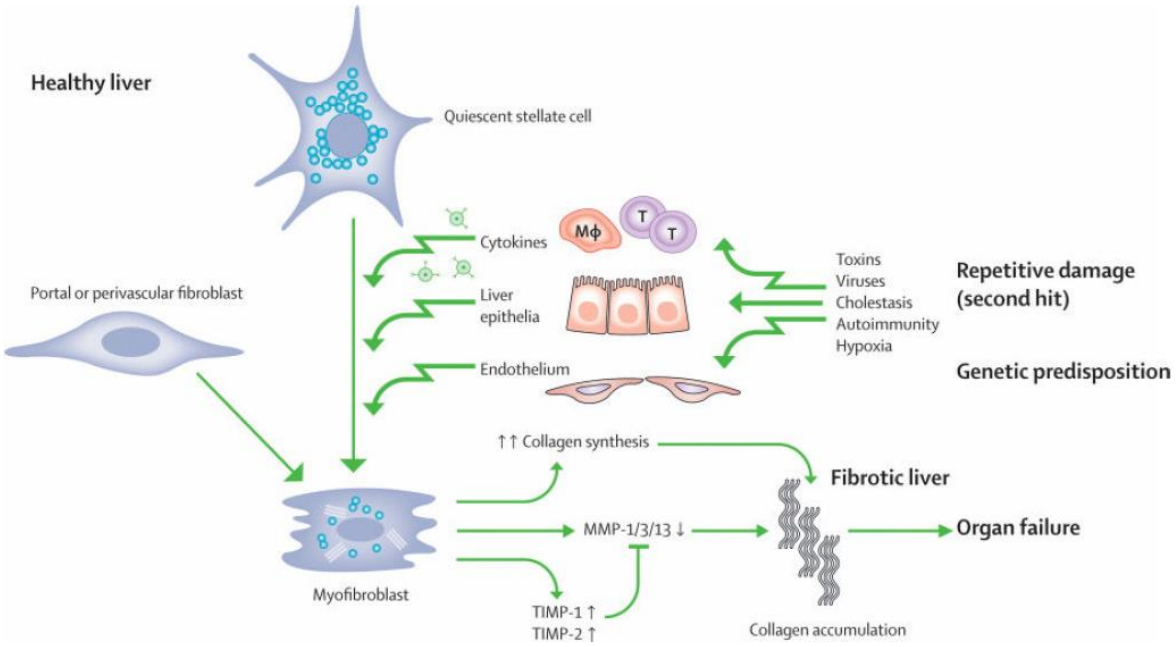


Figure 1.2-1: Schematic representation of the mechanism underlying hepatic fibrosis. Taken from Schuppan D et al, Lancet, 2008.[6]

1.2.2 Fibrosis progresses to cirrhosis

During the early stages of fibrosis there is little disruption of normal hepatic microanatomy and no impact on hepatic function. However, as the fibrogenic process outlined above continues, changes occur in the hepatic microanatomy and microcirculation. Fibrosis deposition in and around sinusoids and the space of Disse leads to 'capillarisation' of the sinusoids and porto-venous and arterio-venous shunting of blood occurs. This reduces hepatocyte perfusion and impairs hepatocyte function.[6, 7] The microcirculatory changes also increase hepatic vascular resistance, which is instrumental in the development of portal hypertension. Nodules of regenerating hepatocytes form and dysplasia within these nodules is the first step to the formation of hepatocellular carcinoma.[6] These microscopic changes are

well characterised and form the basis for the histological staging of fibrosis. For example, the system of fibrosis staging developed by Ishak et al in 1995 stages the severity of fibrosis based on the pattern of fibrosis deposition within liver tissue (Figure 1.2-2).[8] The characteristic histological appearance of cirrhotic liver seen in Figure 1.2-3, namely regenerative nodules of hepatocytes surrounded by bands of fibrous tissue that distort liver anatomy at a microscopic and macroscopic level.

Modified staging: architectural changes, fibrosis and cirrhosis

Change	Score
No fibrosis	0
Fibrous expansion of some portal areas, with or without short fibrous septa	1
Fibrous expansion of most portal areas, with or without short fibrous septa	2
Fibrous expansion of most portal areas with occasional portal to portal (P-P) bridging	3
Fibrous expansion of portal areas with marked bridging (portal to portal (P-P) as well as portal to central (P-C))	4
Marked bridging (P-P and/or P-C) with occasional nodules (incomplete cirrhosis)	5
Cirrhosis, probable or definite	6
Maximum possible score	6

Figure 1.2-2: The original description of the Ishak fibrosis staging system taken from Ishak K et al, Journal of Hepatology, 1995.

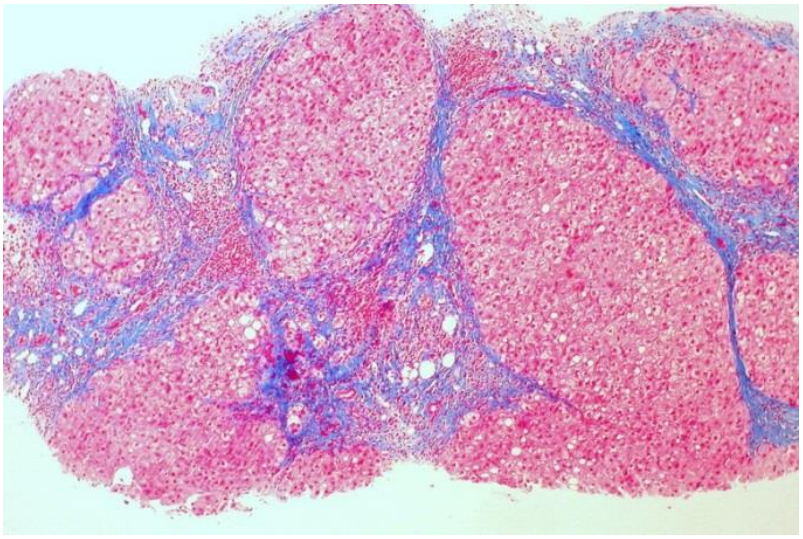


Figure 1.2-3: Trichrome stain of a liver biopsy specimen showing thick fibrous bands (blue) and regenerative nodules typical of cirrhosis. Image courtesy of Ed Uthman, via Wikimedia Commons, accessed 05 July 2016.

1.2.3 Fibrosis staging informs prognosis and guides clinical decision making

The clinical complications of chronic liver disease (portal hypertension, liver failure and hepatocellular carcinoma) occur to a large extent as a result of the effects of cirrhosis and therefore it is reasonable to expect that liver related outcomes should be concentrated in those with cirrhosis. The logical extension to this is therefore that the presence of cirrhosis has a negative impact on prognosis for patients with chronic liver disease. This assumption is supported by evidence from a number of different aetiologies. A 2010 publication from the HALT-C trial of 1050 patients with hepatitis C infection that found there was a statistically significant difference in the incidence of a first liver related outcome between those with cirrhosis (Ishak stage 5 and 6) and those with moderate (Ishak stage 4) fibrosis (Figure 1.2-4).[9] The inter group difference in incidence of a first liver related outcome between Ishak stages 2 to 4 were not statistically significant. This link between cirrhosis and negative

clinical outcomes has also been shown, for example, in patients with alcoholic liver disease.[10, 11] genetic haemochromatosis,[12] autoimmune hepatitis[13] and NAFLD.[14] In NAFLD, Angulo et al demonstrated a 4 fold increase in liver related events in patients with cirrhosis compared to those with advanced fibrosis.[14]

This identification of a group of patients at risk from liver related outcomes thus allows the targeting of screening for gastro-oesophageal varices and surveillance for HCC into high risk groups thus increasing the cost effectiveness of the intervention and avoiding procedural risk in patients unlikely to benefit.

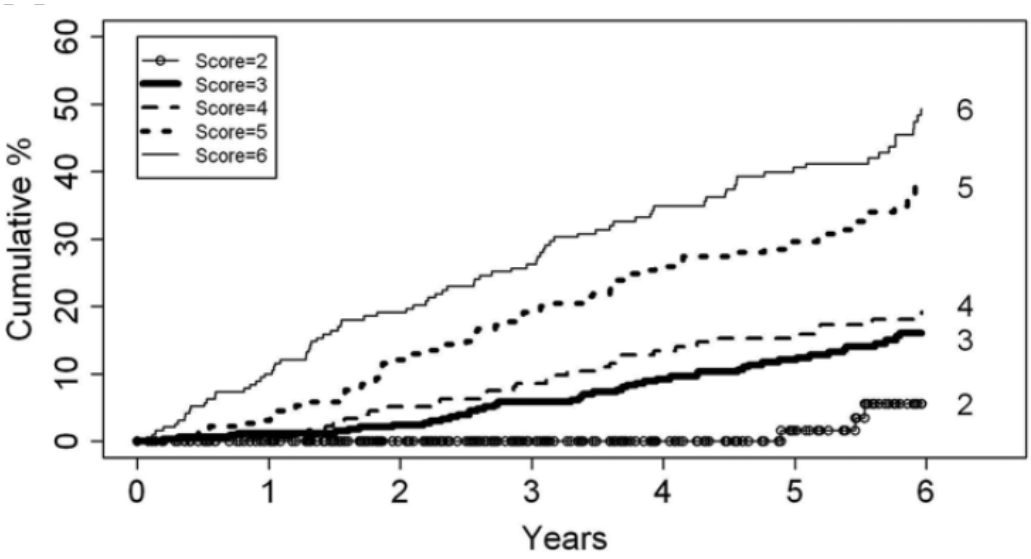


Figure 1.2-4: Cumulative risk of a first liver related outcome stratified by Ishak stage at baseline showing an increased risk and quicker progression to liver liver related events with advancing fibrosis stage. Taken from: Everhart J et al, Hepatology, 2010.[9]

While the presence of cirrhosis is clearly an important indicator of poor prognosis, evidence exists that prognosis in chronic liver disease is not simply linked to the binary distinction

between those with and those without cirrhosis. The stage of fibrosis on an index biopsy has been shown to inform the rate of progression to cirrhosis and the incidence of liver related complications in chronic hepatitis C,[9, 15, 16] alcohol related liver disease,[11, 17] primary sclerosing cholangitis (PSC),[18] primary biliary cholangitis (PBC)[19, 20] and NAFLD.[14, 21] The explanation for the link between fibrosis stage and outcomes in patients who do not have cirrhosis has not been definitively identified however it is likely that the identification of patients with fibrosis, even at an early stage, characterises a patient group with more aggressive disease that leads to greater progression of fibrosis over time and therefore an increased probability of developing complications.

The favourable prognosis for patients without significant fibrosis allows for patients to be appropriately reassured and potentially discharged from clinical follow up in a hospital setting, if there is no evidence of significant fibrosis. Similarly, those with more advanced fibrosis can be kept under close clinical follow up to monitor for progression. These patients may also benefit from entry into clinical studies. Study recruitment can be open only to those with more advanced fibrosis or may be stratified by fibrosis stage. Lastly, the timing of treatment in chronic viral hepatitis is influenced by fibrosis stage. In hepatitis C the recently introduced direct acting antiviral agents have significantly narrowed the gap in treatment efficacy between those with and those without fibrosis,[22] however, knowledge of fibrosis stage continues to influence treatment urgency and the duration of treatment in some genotypes.[22, 23]

In summary, the assessment of fibrosis is important in chronic liver disease to inform prognosis and to inform clinical decisions about follow-up, treatment and inclusion in clinical

studies. There are several available and emerging techniques for the staging of hepatic fibrosis, each with their own advantages and disadvantages. The remainder of this chapter will look at liver biopsy as the current, so-called, gold standard for fibrosis assessment as well as exploring the role of several established and emerging non-invasive techniques in the assessment of hepatic fibrosis.

1.3 Liver biopsy

Indications for liver biopsy can be broadly divided into three groups: to make a diagnosis, to inform about prognosis, to guide treatment decisions.

The assessment of fibrosis is an important factor in all three of these roles. Histological assessment of tissue obtained by liver biopsy is, at present, the gold standard for the assessment of fibrosis in chronic liver disease. This status as the gold standard test for fibrosis assessment is however challenged by many newer technologies. An understanding of the benefits and limitations of liver biopsy is important to understanding its role in hepatology practice and when evaluating alternative technologies.[24]

1.3.1 Advantages of liver biopsy histology for fibrosis assessment

Liver biopsy histology is a well-established technique that is widely regarded as the gold standard for fibrosis assessment. Although non-invasive techniques are beginning to challenge this dogma, liver biopsy histology retains some clear advantages over non-invasive techniques.

Histological assessment of liver tissue allows the pattern of fibrous tissue within the liver to be evaluated. The accepted definition of cirrhosis and the semi-quantitative scoring systems for fibrosis in routine use, such as those described by Ishak et al in 1995[8] and Kleiner et al in 2005[25] are based on fibrosis pattern and distribution rather than the amount of fibrosis per se. This identification of patterns of fibrosis gives great specificity particularly in the

identification of cirrhosis. Although the qualitative assessment of fibrosis pattern is important, the amount of collagen contained within a biopsy specimen can be accurately measured. Collagen proportionate area (CPA) with digital image analysis allows the amount of collagen contained in a liver biopsy specimen to be precisely defined. This has been shown to correlate strongly with portal pressure[26] and clinical outcomes.[15]

Histology is able to differentiate between fibrosis of recent onset and more established fibrosis based on the presence or absence of elastic fibres within the fibrous bands. This information is not possible to obtain from non-invasive tests and can provide valuable insights into the disease process.

Non-invasive fibrosis tests are designed and validated for the detection of fibrosis only and do not give information on diagnosis. Histological assessment of liver tissue is able to detect aetiological factors not suspected by the clinical circumstances. For example a histological diagnosis of autoimmune hepatitis in a patient thought to have NAFLD. Histology also allows for assessment of other pathological features that add value to the assessment such as assessing the severity of steatohepatitis in NAFLD or necro-inflammation in viral hepatitis.

In summary, liver biopsy has several potential advantages over non-invasive techniques for the assessment of hepatic fibrosis and although imperfect, it remains the gold standard for fibrosis assessment in chronic liver disease.

1.3.2 Limitations of liver biopsy histology for fibrosis assessment

The limitations of liver biopsy histology to stage hepatic fibrosis are well documented and can broadly be divided into issues that relate to the biopsy procedure and issues that relate to the histological assessment of the biopsy sample.

Percutaneous biopsy is the most commonly used method of obtaining liver tissue for assessment. While the procedure is generally regarded as safe, it is invasive and has an inherent risk of serious complication. Pain is common following liver biopsy occurring to some degree in up to 84% of patients[27] with approximately 30% of patients requiring analgesia post procedure.[28] Serious complications are rare in patients undergoing non-targeted biopsies with bleeding requiring transfusion or intervention occurring in 0.35-0.5% of procedures and perforation of an adjacent viscus occurring in only 0.01-0.1% of procedures.[28] The mortality rate from liver biopsy is extremely low. A study of over 68,000 liver biopsies from multiple Italian centres revealed a mortality rate of 9 per 100,000 biopsies.[29] Despite the rarity of serious complications, the associated morbidity and potential mortality of liver biopsy reduce patient acceptance and essentially precludes repeated assessment with biopsy.

When they do occur, the complications arising from liver biopsy are most common within the first 4 hours following the procedure. This has led to the adoption of a 4-6 hour period of observation post procedure in most centres.[30] This bed occupancy adds significantly to the cost of liver biopsy and is a distinct disadvantage over non-invasive tests.

Liver biopsy is also limited by the potential for sampling error. A liver biopsy takes only 0.001-0.003% of the total liver volume and this tiny specimen may not be characteristic of the liver as a whole.[31] This variability in histological lesions across the liver has been documented in NALFD,[32] PBC[33] and PSC.[34]

Histological assessment of biopsy specimens is a subjective process and is therefore prone to inter-observer variation. Given a biopsy of ideal size, agreement about fibrosis stage is good.[35] However, agreement is adversely affected by a range of factors including biopsy size,[36, 37] degree of fragmentation and stage of fibrosis in the specimen being assessed.[38] Other features of liver disease assessed at biopsy have lower inter-observer agreement than fibrosis. For example, steatosis has poor inter-observer agreement[39] with a tendency for overestimation compared to digital image analysis of specimens.[40] Given the importance of biopsy size on the accuracy and reproducibility of fibrosis assessment it is recommended that biopsy samples are 2-3cm in length and contain at least 11 portal tracts.[30, 41]

Fibrosis progression from normal liver to cirrhosis is a continuum and it should be recognised that imposing a categorical staging system on a continuous variable is artificial and inevitably introduces error. For example, if you take the Ishak staging system, it would be wrong to think of stage 4 fibrosis as simply twice as bad as stage 2 and the transition from stage 1 to stage 2 does not have the same clinical significance as transition from stage 4 to stage 5.

Some of the above mentioned limitations of histology assessment can be mitigated by the use of collagen proportionate area (CPA) to quantify fibrosis. This technique is objective and so reduces variation in assessment and has a theoretical benefit of describing hepatic fibrosis as a

continuous variable. Digital image analysis of biopsy specimens allows the proportion of liver tissue taken up by fibrosis to be precisely calculated. CPA correlates well with semi-quantitative staging systems such as Ishak stage with published correlation coefficient (Spearman's Rho) of 0.67 ($p<0.001$).[26] It must be remembered that both techniques are assessing fibrosis in very different ways and as such should not be regarded as being directly comparable. A notable strength of CPA is that it has been shown to predict portal pressure[26] and give prognostic information.[42]

1.4 Rationale for development of a non-invasive alternative to biopsy

The issues outlined in section 1.3.2 limit the utility of liver biopsy histology for the staging of fibrosis. Under or over staging of fibrosis has a clear implication when providing prognostic information to patients and making clinical decisions based on the fibrosis stage. This lack of accuracy is also relevant to clinical trials that use liver biopsy as the reference standard. In a trial of a new therapy that uses liver biopsy histology to measure effect, inaccuracies in assessment may under or overestimate the effect of the intervention. Also in studies of novel diagnostic tests an imperfect reference test increases the risk of type 2 error.[43]

A test that is quicker and more acceptable to patients, has less cost to health systems and avoids the sampling error and inter-observer variation inherent in liver biopsy histology has clear attractions and this has led to great interest in the development of non-invasive biomarkers of liver disease. The ideal marker has yet to be discovered but a 2007 review article gives the following as the characteristics of the ideal biomarker:[44]

"*An ideal liver fibrosis marker should have the following characteristics:Liver specific, Readily available and standardised between all laboratories performing diagnostic biochemistry / haematology, Not subject to false positive results, for example due to inflammation, Identifies the stage of fibrosis*"

The characteristics of the ideal non-invasive fibrosis tests. Taken from: Rossi et al, Clin Biochem Rev, 2007.[44]

The following section will examine the currently available biomarkers of hepatic fibrosis.

1.5 Tests Currently Available

Current non-invasive technologies for assessing liver fibrosis include blood biomarker panels, novel blood biomarkers, imaging techniques assessing liver morphology, elastography techniques based on ultrasound and MRI and tissue characterisation with MRI.

1.5.1 Routine serum liver tests

'Routine'serum liver tests, often referred to as liver function tests (LFTs), include bilirubin, albumin, aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP) and gamma-glutamyl transpeptidase (gGT). These tests are inexpensive and readily available however, individually, they do not show a significant association with fibrosis stage.[44] Indeed the entire range of histological abnormalities seen in NAFLD can occur in patients with a normal ALT.[45] Clearly more complex testing is required to establish fibrosis stage.

1.5.2 Simple blood biomarker panels

The use of simple blood biomarker panels began in 1988 when Williams et al reported the use of the AST:ALT ratio to identify patients with cirrhosis.[46] Since then the field has expanded with many different panels evaluated in a range of different aetiologies. Table 1.5-1 outlines the constituents for five of the more common panels. Simple blood biomarkers can be accurate with a recent meta-analysis finding that, to detect advanced fibrosis in patients with

NAFLD, the summary AUROC values for APRI, FIB-4 and NAFLD fibrosis score (NFS) were 0.77, 0.84 and 0.84 respectively.[47]

Name	Included tests
AST:ALT ratio[46]	AST, ALT
AST:platelet ratio index (APRI)[48]	AST, platelet count
Fibrosis-4 (Fib-4)[49]	AST, ALT, platelet count, age
FibroIndex[50]	Platelet count, AST, gGT
Forns index[51]	Age, platelet count, gGT, cholesterol

Table 1.5-1: Constituents of common simple blood biomarker panels.

The research interest is perhaps unsurprising as combining routine tests and basic demographics into a biomarker panel that can predict fibrosis is an appealing way of improving the performance of readily available tests. Often the required information is collected routinely and this would make an accurate simple blood biomarker panel extremely cost effective.

However, even with the obvious advantages, simple biomarker panels are not without limitations. The panels outlines in Table 1.5-1 contain tests that do not directly measure hepatic fibrosis but rather the effects of hepatocyte injury.[52] For this reason they are sometimes referred to as 'indirect tests' and as such are prone to confounding from necro-inflammatory activity within the liver.[44] An example of this is the effect of alcohol on the liver tends to elevate AST out of proportion to ALT, which significantly confounds simple panel that include AST such as AST:ALT ratio and APRI.

The application of biomarker panels in clinical practice must also take into account the liver disease aetiology of the population in which they are being used and also the population in

which they were defined. The majority of the literature in the field of simple biomarker panels has been conducted in patients with chronic hepatitis C[53] and it does not necessarily follow that performance will be equivalent in other aetiologies. gGT has been shown to be associated with fibrosis in chronic hepatitis B but not in other aetiologies.[54] Knowledge of the characteristics of the reference population is not limited to aetiology. The gGT:platelet ratio has been found to be a useful marker of liver fibrosis in sub-Saharan Africa[55-57] but this finding was not however replicated in a large cohort of patients with chronic hepatitis B in Hong Kong.[58]

The accuracy of biomarker panels is known to vary with fibrosis stage.[59] An illustration of this is a 2013 meta-analysis in patients with hepatitis C that evaluated biomarker panels for the detection of fibrosis (Table 1.5-2) and detection of cirrhosis (Table 1.5-3).[60] It is notable from these data that, for the detection of fibrosis, the accuracy is only moderate and certainly less good than for the detection of cirrhosis. For all reported cut off values, specificity is high with low sensitivity or vice versa thus reducing the usefulness. The use of dual cut off values helps to maximise both sensitivity and specificity however this leads to a 'grey zone' for the test where the significance of the result is unclear and further testing is necessary.

For cirrhosis, the performance of all evaluated biomarker panels is superior to that for the detection of fibrosis and there is increasing confidence in the use of simple biomarker panels for the exclusion of cirrhosis however, simple blood biomarkers have insufficient sensitivity or specificity for the detection of early stage fibrosis.

Another factor that reduces the clinical usefulness of simple biomarkers is the lack of agreement between different panels. Recent work from Edinburgh in NAFLD shows poor correlation between AST:ALT ratio, APRI and Fib-4, which can lead to diagnostic confusion.[61]

Name	**Median (range) AUROC**	**Cut off value**	**Median (range) Sensitivity**	**Median (range) Specificity**
Platelet count	0.71 (0.38-0.94)	<140 to <163*	0.56 (0.28-0.89)	0.91 (0.69-1.0)
AST:ALT ratio	0.59 (0.50-0.82)	>1.0	0.35 (0.08-0.45)	0.77 (0.62-1.0)
APRI	0.77 (0.58-0.95)	>0.55	0.81 (0.29-0.98)	0.55 (0.10-0.94)
		>1.5	0.37 (0-0.72)	0.95 (0.58-1.0)
Fib-4	0.74 (0.61-0.81)	>1.45	0.64 (0.62-0.86)	0.68 (0.54-0.75)
		>3.25	0.50 (0.28-0.86)	0.79 (0.59-0.99)
FibroIndex	0.76 (0.58-0.86)	>1.25	0.94 (0.62-0.97)	0.40 (0.40-0.48)
		>2.25	0.30 (0.17-0.36)	0.97 (0.97-1.0)
Forns index	0.76 (0.60-0.86)	>4.2	0.88 (0.57-0.94)	0.52 (0.20-0.77)
		>6.9	0.36 (0.18-0.61)	0.94 (0.66-1.0)

** Cut-off value varies between studies included in the meta-analysis.*

Table 1.5-2: Common blood biomarker panels and their diagnostic performance in terms of AUROC for the detection of significant fibrosis (METAVIR ≥ stage 2). Median sensitivity and specificity of these biomarker panels are shown based on published cut off values. Table adapted from: Chou et al, An Int Med, 2013[60]

Name	Median (range) AUROC	Cut off value	Median (range) Sensitivity	Median (range) Specificity
Platelet count	0.89 (0.64-0.99)	<140 to <155*	0.78 (0.41-0.93)	0.87 (0.84-0.94
AST:ALT ratio	0.72 (0.52-0.91)	>1	0.36 (0.12-0.78)	0.92 (0.59-1.0)
APRI	0.84 (0.54-0.97)	>1	0.84 (0.54-0.97)	0.75 (0.30-0.87)
		>2	0.48 (0.17-0.76)	0.94 (0.65-0.99)
Fib-4	0.87 (0.83-0.92)	>1.45	0.90†	0.58†
		>3.25	0.55†	0.92†
FibroIndex	0.86 (0.78-0.92)	>1.9	0.70 and 0.91†	0.91 and 0.78†
Forns index	0.87 (0.85-0.91)	>4.2	0.98†	0.27†
		>6.9	0.67†	0.91†

** Cut-off value varies between studies included in the meta-analysis.*
† Results from individual studies given if fewer than 3 studies included.

Table 1.5-3: Common blood biomarker panels and their diagnostic performance in terms of AUROC for the detection of cirrhosis (METAVIR stage 4). Median sensitivity and specificity of these biomarker panels are shown based on published cut off values. Table adapted from: Chou et al, An Int Med, 2013[60]

1.5.3 Direct Serum Biomarkers

In addition to the indirect tests discussed in section 1.5.2, direct fibrosis markers have been investigated for the staging of hepatic fibrosis. Direct markers measure components of the fibrotic process either individually or combined in a panel as outlined in Table 1.5-4. These tests have a theoretical benefit over indirect tests in that they are less susceptible to confounding from hepatic inflammation. The performance of these tests is discussed individually.

Name	Included tests
Hyaluronic acid	
Osteopontin	
Cytokeratin (CK) -18	
Enhanced liver fibrosis (ELF) test [62]	hyaluronic acid, amino-terminal propeptide of type III procollagen (PIIINP) and tissue inhibitor of metalloproteinase 1 (TIMP-1), (original version also included age)
Fibrotest [63]	age, gender, α2-macroglobulin, haptoglobin, apolipoprotein A1, gGT, bilirubin
FibroMeter [64]	Age, gender, platelet count, AST, α2-macroglobulin, prothrombin time, gGT, urea

Table 1.5-4: Constituents of direct blood biomarkers.

Hyaluronic acid

Hyaluronic acid is an integral part of extracellular matrix and in some studies the AUROC of hyaluronic acid to differentiate cirrhosis from lower stages of fibrosis has been as high as 0.924.[65] This finding has not however been universally replicated. Hyaluronic acid has not been shown to correlate well with fibrosis stages below cirrhosis and it has low accuracy for the detection of early fibrosis.[66] In addition, hyaluronic acid is a component of extracellular matrix throughout the body leading to confounding from other fibrotic diseases.

Osteopontin

Osteopontin is a pro-fibrogenic extracellular matrix protein that recruits neutrophils and T-cells into the liver and promotes the deposition of collagen by hepatic stellate cells.[67] Osteopontin has been shown to predict fibrosis with high accuracy in alcoholic liver disease[68] and hepatitis C infection.[69] However, these studies are small and independent validation is lacking. The most significant obstacle to the adoption of Osteopontin as a fibrosis biomarker is that lacks specificity for hepatic fibrosis. It is elevated in patients with hepatocellular

carcinoma,[70, 71] non-hepatic carcinomas[72-74] and patients at increased risk of cardiovascular events.[75]

Cytokeratin-18

Cytokeratin-18 (CK-18) is an intracellular filament protein that is broken down and released during hepatocyte apoptosis. There is limited evidence from small, single centre studies that measurement of CK-18 fragments in serum correlates with fibrosis.[76] However, CK-18 measurements are heavily confounded by inflammation and particularly non-alcoholic steatohepatitis (NASH).[77] The main interest in CK-18 is as a marker of NASH and as such will be discussed in more detail in Chapter 3.

Enhanced liver fibrosis test

As outlined in Table 1.5-4, the enhanced liver fibrosis (ELF) test is a combination biomarker that calculates a score based on the serum concentration of several components of the fibrotic process. The original ELF was derived from a cohort of 921 patients with paired liver biopsy and serum samples and also included age in the equation. AUROC for the detection of cirrhosis was 0.887 with AUROC for detection of any fibrosis lower at 0.77.[62] A further validation study in 192 patients with NAFLD developed a simplified ELF panel that excluded age from the equation. Both original ELF and simplified ELF have high accuracy for the detection of advanced fibrosis with AUROCs of 0.89 and 0.93 respectively.[78] Meta-analysis of 9 studies including both simplified and original ELF has shown high accuracy for the identification of both advanced fibrosis and cirrhosis with median (range) AUROC 0.81 (0.72 to 0.87) and 0.88 (0.78 to 0.91) respectively.[60] An independent validation of the simplified ELF test by Lichtinghagen et al in 2013 has highlighted areas of caution in interpreting ELF

results.[79] In this large study ELF identified patients with advanced fibrosis and cirrhosis with high accuracy however; scores were significantly higher in men than women and had a trend towards being higher in the afternoon than the morning. Despite these cautions in interpreting the results, ELF is a valuable tool in the non-invasive assessment of hepatic fibrosis and is gaining widesepread acceptance. The recently published National Institute for Health and Care Excellence (NICE) guidelines on the management of NAFLD use ELF as a cost effective 1st line method of staging fibrosis and therefore guiding patient referral.[80]

Fibrotest

Fibrotest has shown high accuracy for identification of cirrhosis in hepatitis C infection,[81] hepatitis B infection[82] and alcoholic liver disease[83] with AUROC (95% CI) of 0.94 (0.86-0.98), 0.991 (0.973-1.00) and 0.86 (0.83-0.89) respectively. Similarly to other evaluated blood biomarkers, accuracy for advanced fibrosis is lower with little ability to detect early stage fibrosis.

FibroMeter

FibroMeter (Echosens, Paris, France) is a proprietary algorithm combining several direct and indirect markers of hepatic fibrosis. Direct markers are specific for fibrosis but not specific for the liver and the converse is true for indirect markers and therefore the combination of direct and indirect markers has a theoretical benefit over either approach alone. Another unique feature of the FibroMeter test is that the component blood parameters and the algorithm itself varies between viral liver disease and alcohol related liver disease. In the original work on FibroMeter by Calès et al the diagnostic accuracy (% (95% CI)) of FibroMeter for the detection of clinically significant (METAVIR ≥F2) fibrosis was 82.1 (77.7–86.5) for viral

liver disease and 92.0 (86.4–97.7) for alcohol related liver disease.[64] Diagnostic accuracy for both algorithms was reduced when used to assess the aetiology for which it was not designed and this highlights the difficulty in applying FibroMeter when aetiology is not clear or mixed. In addition FibroMeter is a proprietary algorithm and this adds to the cost of performing the test.

1.5.4 Liver morphology on routine imaging

The morphological changes in the liver that occur in cirrhosis can be identified on routine clinical imaging including ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI). Clear signs of established cirrhosis and portal hypertension such as splenomegaly, intra-abdominal varices and irregularity of the liver margin are well known and have high positive predictive value. However, negative predictive value is low and these signs occur only in established cirrhosis and portal hypertension. Several more subtle morphological changes have been identified that occur in cirrhosis in the absence of the above findings. These include elevated caudate to right lobe ratio, widening of the porta hepatis, expansion of the gall bladder fossa and reduction in the liver to abdominal area ratio (LAAR).[84-88] These features have been shown to be highly specific for cirrhosis and may help to identify patients with compensated cirrhosis who are at risk liver related adverse events.[88, 89] Despite their simplicity, many of these features are prone to inter-observer variability and although specificity is high, sensitivity is relatively low. Liver morphology has no role in the staging of fibrosis that falls short of cirrhosis.

1.5.5 Elastography Techniques

Elastography is the measurement of the elastic properties of soft tissue. This is more commonly referred to as tissue 'stiffness'. As well as the assessment of liver fibrosis, elastography has been investigated for the detection of tumours in breast and prostate, the diagnosis of muscular and tendon injury and the evaluation of venous thrombus.[90] In this section the evidence supporting the 3 most commonly used elastography techniques is examined. These techniques are vibration controlled transient elastography (VCTE), Acoustic radiation force impulse (ARFI) elastography and magnetic resonance elastography.

Vibration controlled transient elastography (FibroScan™)

Vibration controlled transient elastography (VCTE) is an ultrasound based technology upon which FibroScan™ (Echosens, Paris, France) is based. VCTE allows the non-invasive measurement of liver stiffness as a surrogate for liver fibrosis. FibroScan™ is a bedside test that can be completed in a few minutes. The handheld FibroScan™ probe generates a low frequency pulse when held again the skin on the right side of the abdomen. This pulse causes a sheer wave to propagate through the liver tissue. Multiple ultrasound pulses track the propagation of the sheer wave and allow the calculation of the speed of the wave. The speed of the wave is proportional to the stiffness of the liver tissue and the FibroScan™ machine displays a measurement of liver stiffness in kilopascals (kPa).

The advantages of VCTE over liver biopsy histology include the fact that it is non-invasive and therefore does not carry the risks associated with biopsy. The volume of tissue examined with VCTE (approximately 3 cm^3) is much greater than that assessed by liver biopsy reducing

the potential for sampling error. It is also quick to perform and, unlike blood biomarkers, provides immediate results. Although VCTE requires operator training, reliable results can be obtained after as few as 10 examinations.[91] Inter-observer and intra-observer reproducibility has been shown to be good.[92]

As with all tests there are limitations to its use. Obesity can make VCTE impossible in up to 15% of patients[93, 94] and, even when VCTE is technically possible, an increased distance from the skin to the capsule of the liver has been shown to reduce the accuracy of fibrosis assessment.[95] VCTE is an indirect test of fibrosis and liver stiffness can be confounded by several factors. Hepatic inflammation,[96] hepatic engorgement from cardiac failure[96] and increased portal blood flow from recent food intake[97] all elevate liver stiffness independently from fibrosis. It has also been reported that the degree of steatosis influences the accuracy of fibrosis assessment with severe steatosis leading to an over estimation of fibrosis.[98] However, this finding is disputed and in a large cohort of UK patients with NAFLD, it has been demonstrated that in multivariable analysis steatosis does not influence liver stiffness.[99]

Notwithstanding the above limitation, liver stiffness as measured by VCTE has been demonstrated to be a reliable surrogate for cirrhosis in multiple studies since the first report was published in 2003.[92] AUROC values for the detection of cirrhosis with VCTE vary slightly but are generally >90%.[100-104] Although VCTE can reliably diagnose cirrhosis, the accuracy for lower stages of fibrosis is less good. A prospective study of a large population of patients with chronic viral hepatitis showed AUROC for moderate fibrosis (METAVIR ≥F2) of 0.76.[105] A 2013 meta-analysis produced similar findings with sensitivity and specificity for the diagnosis of moderate fibrosis (METAVIR ≥F2) of 0.79 and 0.78.[106] The speed and ease

of use of VCTE makes it an attractive tool for screening for liver disease. A 2012 study found VCTE to be a useful tool for the identification of cirrhosis in an asymptomatic population attending for a medical check-up.[107]

VCTE has also been shown to correlate well with portal hypertension and the risk of oesophageal variceal haemorrhage.[108-111] The Baveno VI consensus report uses a cut off value of 20kPa combined with a platelet count of >150,000 to exclude clinically significant portal hypertension and thus avoid surveillance endoscopy.[112, 113] This approach has been externally validated and, in a large retrospective study, was found to misclassify only 2% of patients.[114]

Acoustic radiation force impulse (ARFI) elastography

Acoustic radiation force impulse (ARFI) elastography is an ultrasound based technique that relies on the measurement of liver stiffness as a surrogate for fibrosis. It is conceptually similar to VCTE in that sheer waves are generated in liver tissue and ultrasound pulses track the speed of sheer wave propagation. Whereas FibroScan™ requires a dedicated machine and uses a mechanical driver to generate the sheer wave, ARFI is a function built into conventional ultrasound machines and uses a 'longitudinal wave push pulse' generated by the ultrasound probe to initiate the sheer wave. As shown in Figure 1.5-1, a region of interest (ROI) is placed within the liver parenchyma and the liver stiffness is calculated within the ROI. Placing the ROI within the parenchyma under direct vision allows for major biliary and vascular structures that would increase stiffness to be avoided. ARFI is available on the majority of modern ultrasound equipment from various manufacturers under various trade names. This wide availability of ARFI elastography allows it to be performed at the same time as conventional imaging, providing valuable additional information to the clinician.

In general terms the performance of ARFI is comparable to that of VCTE for cirrhosis and advanced fibrosis[115] but lower for early fibrosis.[116] For the detection of cirrhosis AUROC across studies is consistently >0.9 with AUROC for significant fibrosis >0.8.[117-121] ARFI has also been shown to correlate with the severity of portal hypertension.[111] The failure rate of ARFI is significantly lower than VCTE even when the FibroScan™ XL probe is available.[115, 120]

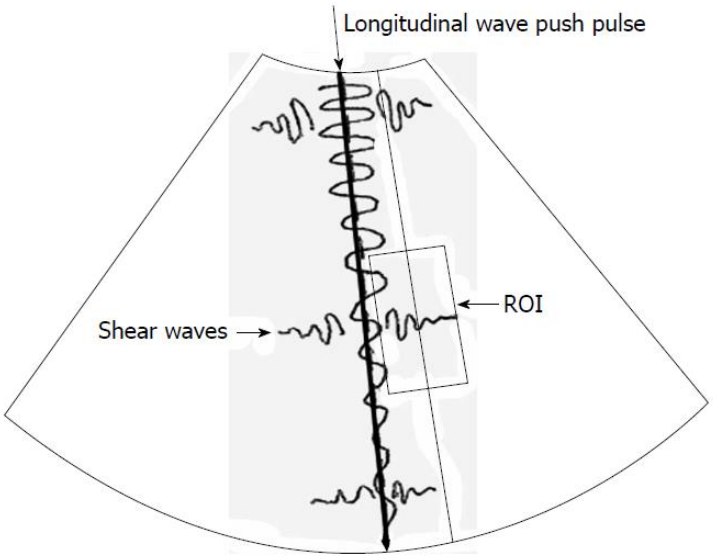


Figure 1.5-1: Schematic representation of an ARFI elastography examination. Taken from D'Onofrio et al, World J Gastroenterol, 2013.[117]

Magnetic Resonance Elastography

Magnetic Resonance Elastography (MRE) is an MRI based elastography technique and, like VCTE and ARFI, relies on the premise that the stiffness of liver tissue is proportional to the degree of fibrosis. MRE employs a mechanical driver strapped to the abdominal surface to send sheer waves through the liver. Figure 1.5-2 shows example images generated by MRE sequences. A region of interest (ROI) is placed within the liver and the mean stiffness for that ROI is expressed in kPa. For the examples in Figure 1.5-2, these values are 2.07kPa for normal liver and 9.65kPa for cirrhosis. It should be noted that despite VTCE and MRE using the same units the differences in the techniques and the use of different cut off values mean that liver stiffness measured with MRE is not directly comparable with liver stiffness measured by VTCE.

Many of the early studies into the use of MRE were small and are difficult to compare due to variation in: shear wave frequency, acquisition sequence, method of histological assessment and study population.[122] More recent work has standardised many of these parameters and MRE has emerged as a useful tool for the non-invasive assessment of liver fibrosis.[122-129] The performance of MRE is superior to other non-invasive markers including VCTE for the detection of cirrhosis [125, 130, 131] and this difference becomes more marked in early stage fibrosis where MRE shows significantly better performance.[123, 131-133] Additional value of the MRE examination in NAFLD is that MRE has also been shown to differentiate simple steatosis from the more aggressive non-alcoholic steato-hepatitis (NASH) even in those without fibrosis.[126, 134, 135] MRE has an additional advantage of VCTE in that it is not adversely affected by the presence of ascites or obesity. Overall, MRE has a significantly higher success rate than VCTE especially in obese patients.[123, 125, 136]

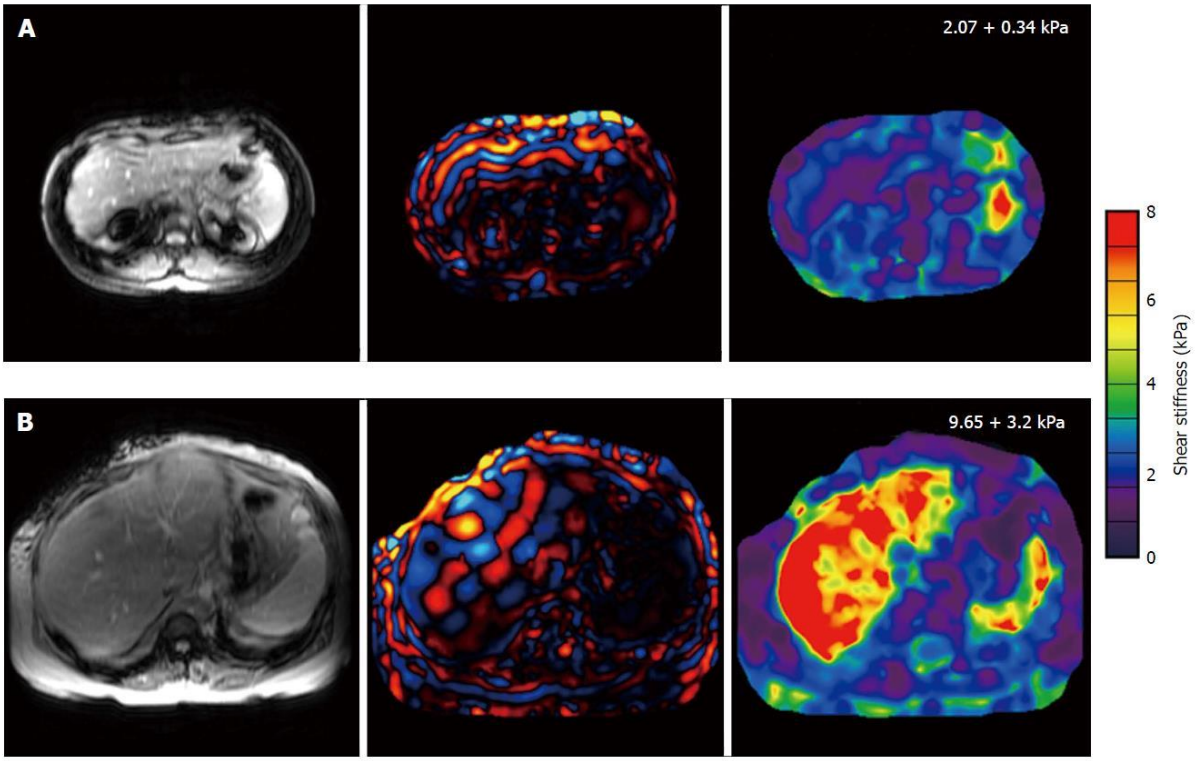


Figure 1.5-2: Example images generated by magnetic resonance elastography. From left to right the images show: localiser images ensure correct slice selection, phase contrast acquisition sequences incorporating motion encoding gradients detect the motion of liver tissue and inversion algorithms convert these measurements to colour coded maps of liver stiffness. Row A shows a patient without significant fibrosis and row B shows a patient with cirrhosis confirmed on liver biopsy. Figure taken from Low G et al, World Journal of Radiology, 2016.[137]

MRE is not without limitations. MRE suffers the same lack of specificity for fibrosis as all elastography techniques. Inflammation, hepatic congestion and postprandial hyperaemia all increase liver stiffness independent from fibrosis. In addition, the reduction in MR signal seen in patients with iron overload can make MRE impossible. [125] The major disadvantage of MRE

over VCTE is cost and availability. MRE requires additional hardware in addition to standard MRI equipment and is not widely available in the UK.

1.5.6 MRI techniques

Several MR techniques have been evaluated for their use as a measure of fibrosis in liver disease. To date the best studied techniques have been diffusion weighted imaging and perfusion imaging. In this section the evidence supporting their use is presented. Another novel MRI technique is T1 mapping, the evaluation of which is the main focus of this thesis. The technique is discussed further in Section 1.7 and Chapter 2.

Diffusion Weighted Imaging

Diffusion Weighted Imaging (DWI) is an MRI technique that uses the movement of protons within tissue to provide image contrast. The most abundant source of protons in tissue is water and, in effect, DWI measures the diffusion of water molecules. In free water, molecules move randomly by Brownian motion, however within tissues the movement of water molecules is restricted by macromolecules and cellular structures (cytoskeleton, organelles, membranes etc.). In a DWI acquisition sequence, movement of protons due to diffusion causes a loss of signal and therefore the greater the ability for water to diffuse the greater the loss of signal intensity. The loss of signal is expressed as an apparent diffusion coefficient (ADC).

DWI is a well-established technique for the identification and characterisation of focal liver lesions[138-140] and several studies have investigated the ability of DWI to stage hepatic fibrosis.

A reduction in ADC with increasing fibrosis has been shown in both experimental animal models and humans.[139, 141-145] The ability to identify early stage fibrosis is limited but some studies report AUROC (95% CI) for the identification of advanced (≥F3) fibrosis as high as 0.91 (0.84-0.99).[146] Studies are however generally small and of low quality[147] making the role of DWI to stage fibrosis in clinical practice uncertain. A 2012 meta-analysis found summary AUROC for the detection of advanced fibrosis to be reasonably high at 0.86 but the included studies were small and heterogeneous.[129] In addition, the presence of inflammation, steatosis and altered perfusion may have a role in reducing ADC and confounding fibrosis assessment.[148, 149] The final difficulty in assessing the utility of DWI is the variation in technique for calculating ADC. In particular an acquisition parameter known as the 'b-value' has been shown to effect the accuracy of fibrosis assessment[107] and this parameter varies widely between studies.[138-140] In summary, DWI has proven to be a useful technique for the characterisation of focal liver lesions but currently its efficacy for the staging of fibrosis has not been proven.

Perfusion Imaging

During the progression of fibrosis, hepatic micro-anatomical changes reduce arterial and portal venous inflow to the liver and thus alter hepatic perfusion. Perfusion imaging techniques using liver specific contrast agents can assess hepatic perfusion and therefore have potential to stage hepatic fibrosis. A 2008 study by Higiwara et al showed AUROC for the diagnosis of significant fibrosis of 0.824 [150] however, to date, studies investigating this technique have been small and heterogeneous.[150-155] Recent work by a group from Nottingham has shown the ability of a multiparametric MRI technique that includes

measurement of hepatic perfusion to predict portal hypertension.[156] There is clear potential for development of these techniques into a useful clinical tool.

1.6 Summary

Chronic liver disease is a major public health issue and its prevention, assessment and treatment is a challenge to health services worldwide. Progressive fibrosis is the common final pathway for nearly all chronic liver disease and so the detection and staging of fibrosis is an important goal for clinicians. Accurate staging of fibrosis helps to target surveillance for complications, provide prognostic information and guide treatment. The accurate assessment of fibrosis is also of importance as an endpoint for clinical studies.

Existing blood markers and imaging techniques have their strengths but often suffer from relatively low sensitivity for the detection of the early stages of fibrosis and have limited ability to measure small changes in fibrosis. Liver biopsy, the current gold standard method for assessing fibrosis, is a valuable test but is imperfect due to several important limitations.

There is a clear need for the development of a reliable alternative to existing non-invasive biomarkers.

1.7 Multiparametric MRI and Liver*Multiscan*™

A proposed biomarker of liver disease is multiparametric MRI. This technique involves running multiple acquisition sequences within a single examination to provide quantitative data about the liver. Data from these sequences is analysed with software (Liver*Multiscan*™) developed by Perspectum Diagnostics ltd (Oxford, UK). Liver*Multiscan*™ uses a proprietary algorithm to correct T1 for the confounding effects of iron and produces a corrected T1 (cT1) value. This is proposed as a novel biomarker of hepatic fibrosis. The Liver*Multiscan*™ software also provides a measure of hepatic steatosis and siderosis with modified Dixon and gradient echo sequences respectively. These techniques are well validated biomarkers of liver disease.

The sequences required for Liver*Multiscan*™ are:

- T1 mapping - Shortened modified Look Locker inversion sequence
    - A potential biomarker of hepatic fibrosis
- T2* mapping - High resolution gradient echo sequence
    - A validated measure of hepatic iron content
- Proton density fat fraction - Modified Dixon sequence
    - A validated measure of hepatic steatosis

The remainder of this thesis examines the role of multiparametric MRI analysed with Liver*Multiscan*™ to calculate cT1 and the utility of cT1 measurements in the evaluation of hepatic fibrosis.

Chapter 2 looks at the background to multiparametric MRI, the techniques used and the reproducibility of the technique.

Chapter 3 looks at the use of multiparametric MRI to stage fibrosis in a mixed cohort of patients undergoing liver biopsy and a cohort of patients on the transplant waiting list.

Chapter 4 examines the use of multiparametric MRI in the evaluation of NAFLD.

Chapter 5 assesses a novel application of magnetic resonance spectroscopy to evaluate NAFLD severity.

Chapter 6 looks at the ability of multiparametric MRI to define disease severity in PSC.

CHAPTER 2: MRI METHODS AND REPRODUCIBILITY STUDIES

2.1 Rationale for the investigation of T1 mapping as a biomarker of liver disease

In chapter 1 the importance of hepatic fibrosis assessment was discussed along with the currently available methods for fibrosis assessment. There is a need to develop new biomarkers that are accurate, cost effective and acceptable to patients. The main focus of this thesis is an investigation of the potential applications of T1 mapping as a non-invasive biomarker of liver diseases severity.

T1 mapping is a MRI technique widely used in the field of cardiac MRI to detect myocardial fibrosis[157] however, it has not been well studied in liver disease. The concept that T1 is proportional to hepatic fibrosis is not new. As early as 1983 hepatic T1 was noted to be prolonged in animal models of hepatic fibrosis[158] and this finding has subsequently reproduced in both animals and humans.[159-165] The mechanism is thought to relate to the free water content of liver tissue. T1 is a physical property of tissues that is affected by the molecular composition of that tissue. As hepatic fibrosis progresses the volume of extracellular matrix increases. Extracellular matrix is rich in free water molecules, which have a long T1. Thus, as fibrosis progresses, for any given volume of liver tissue there is an increasing abundance of free water molecules. This relative excess of free water prolongs T1. Thus it is proposed that T1 is proportional to hepatic fibrosis.

A detailed explanation of MRI physics is beyond the scope of this thesis and is not necessary to understand the rest of the work. However, a basic understand of the principles behind the techniques and a basic understanding of how T1 and T2 are derived is helpful and will be described here.

2.2 T1 relaxation

Due in part to its abundance in the body the most commonly used element for MRI is hydrogen. The nucleus of which is known as a proton. A proton can be visualised as a tiny magnet that is spinning on its axis. The poles of these 'magnets' are arranged randomly as shown in Figure 2.2-1. Within a volume of tissue the magnetic forces from each individual proton cancel out giving no net magnetic charge.

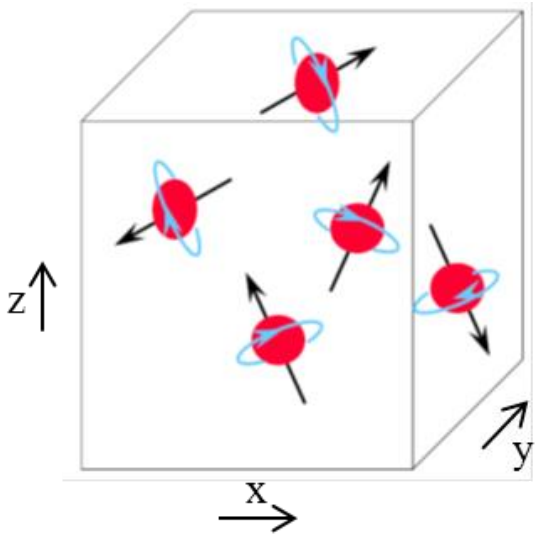


Figure 2.2-1: Protons within a volume of tissue have no net magnetic charge. Taken from: Basic MRI Physics, Evert J Blink, 2004.[166]

MRI scanners contain strong magnets that produce a static magnetic field. By convention, the axis of the static field is denoted as the *z* axis and when tissue is placed inside this strong magnetic field the protons align with *z* as shown in Figure 2.2-2.

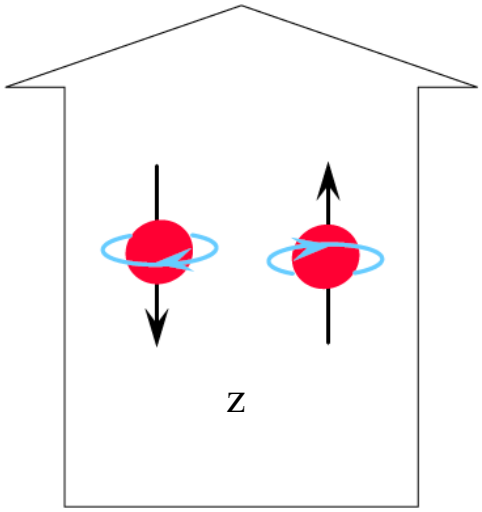


Figure 2.2-2: Protons align with the magnetic field found within MRI scanners. Taken from: Basic MRI Physics, Evert J Blink, 2004.[166]

The application of a radiofrequency (RF) pulse causes the axis around which the protons spin to 'flip' away from the *z* axis towards the *xy* plane. This process is shown in Figure 2.2-3. The angle through which the protons flip (the flip angle) can be altered by the strength and duration of the RF pulse. Once the RF pulse is removed protons relax back to align with *z*. During flipping the protons gain energy from the RF pulse and during relaxation this energy is released as RF signal and heat.

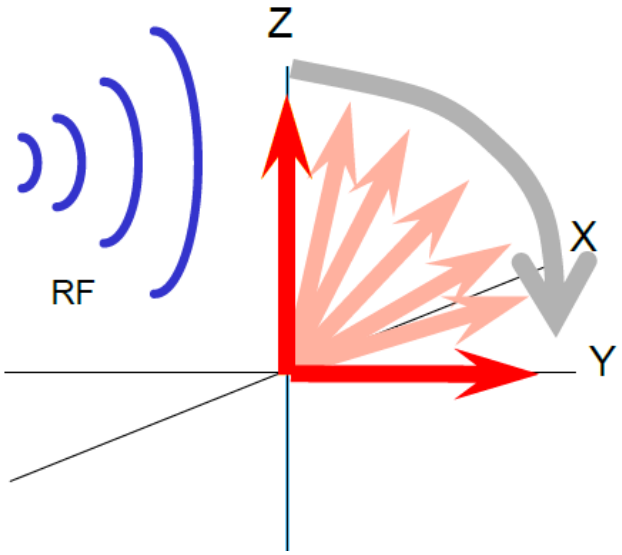


Figure 2.2-3: Application of a radio frequency pulse causes the axis of protons to 'flip'. In this example the axis has moved through 90 degrees and so has a flip angle of 90 degrees. Taken from: Basic MRI Physics, Evert J Blink, 2004.[166]

Although all protons in a given volume of tissue will start to relax at the same time the rate at which protons relax is governed by the chemical bonds in which each individual proton is involved. For example the tightly bound protons within fat molecules relax much more quickly than the more loosely bound protons in water molecules. The rate of relaxation for a volume of tissue can be described by the T1 relaxation curve as shown in Figure 2.2-4. T1 is defined as the time in milliseconds for 63% of the protons within a volume of tissue to align with *z*. In different tissues the T1 relaxation curve is shifted either up or down thus giving a different value for T1.

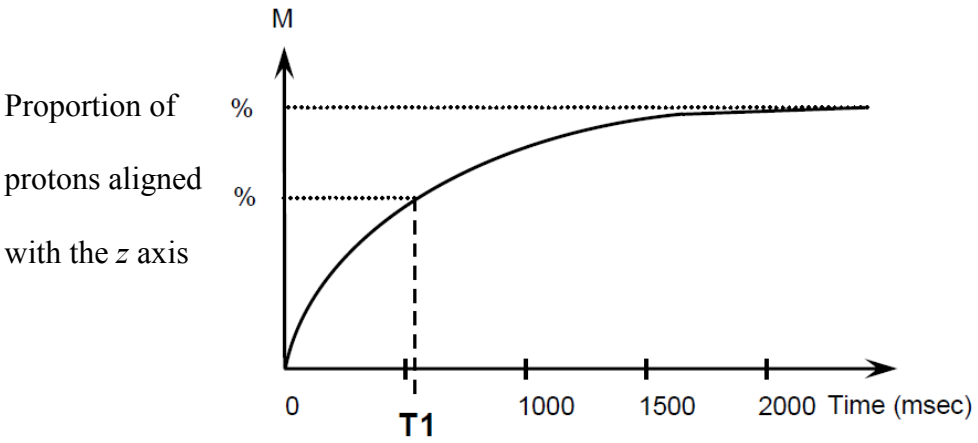


Figure 2.2-4: T1 relaxation curve. Taken from: Basic MRI Physics, Evert J Blink, 2004.[166]

2.3 T2 relaxation

T2 relaxation occurs simultaneous with T1 relaxation but is an entirely separate process. Protons wobble on their axis as they spin. This is known as precession. When exposed to the RF pulse not only does the axis of the spins flip (as described above) but the protons start to precess in phase. This is represented schematically in Figure 2.3-1.

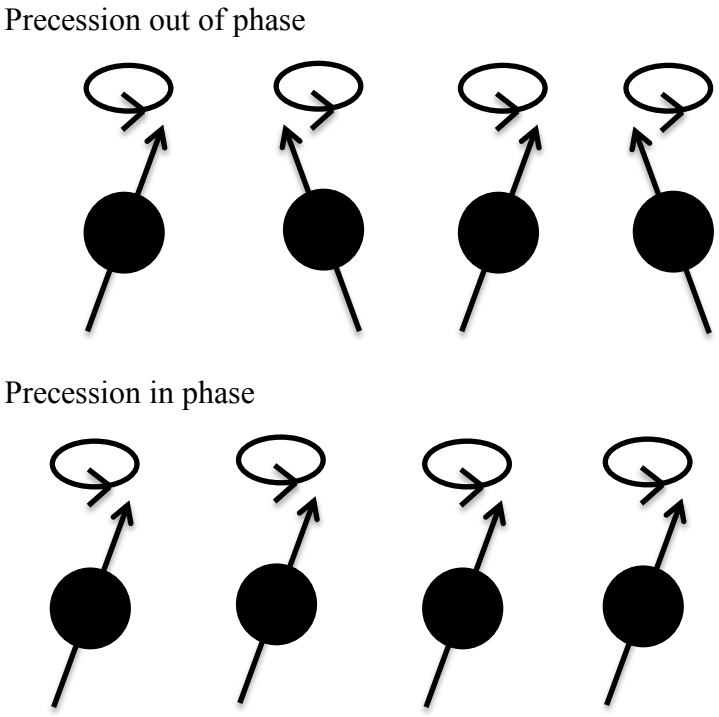


Figure 2.3-1: Schematic representation of precession in and out of phase. Adapted from: Basic MRI Physics, Evert J Blink, 2004.[166]

When the RF pulse is removed, interactions between protons (known as spin-spin interaction) cause the precession to de-phase. The rate at which a proton goes out of phase is again dependent on the bonds in which the proton is involved. For example, tightly bound protons

within fat molecules de-phase much more quickly than the more loosely bound protons in water molecules.

The rate of de-phasing for a given volume of tissue can be described by the T2 relaxation curve as shown in Figure 2.3-2. T2 is defined as the time in milliseconds for the spins to de-phase to 37% of the original value.

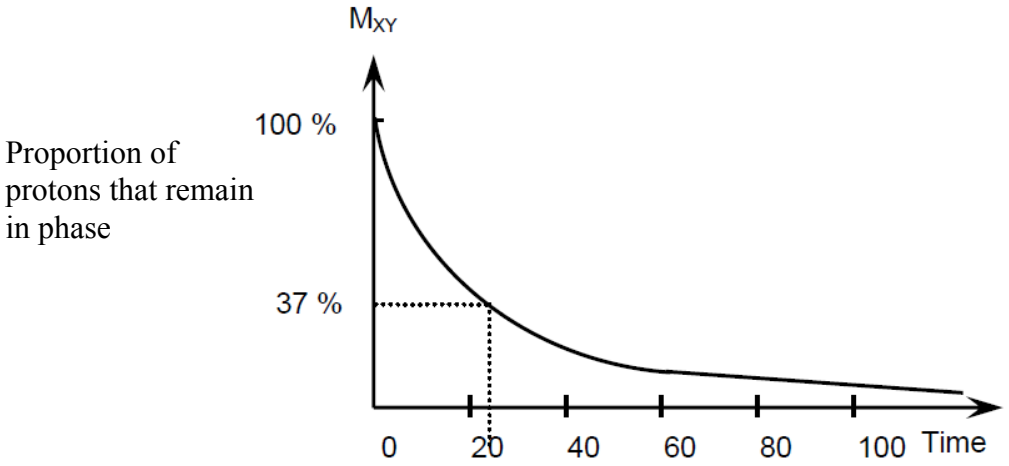


Figure 2.3-2: T2 relaxation curve. Taken from: Basic MRI Physics, Evert J Blink, 2004.[166]

2.4 T2* relaxation

In reality the observed loss of phase coherence occurs more quickly than would be expected from T2 effects alone. Tiny imperfections in the construction of the magnet and the mass of the patient within the scanner distort the magnetic field and cause it to become inhomogeneous. This inhomogeneity causes the spins to de-phase more quickly than would occur from spin-spin interaction alone. The time taken for 63% of the spins to de-phase is the T2* time.

T2* is particularly relevant to this work as the presence of paramagnetic substances within the magnetic field increases the magnetic field inhomogeneity and thus shorten T2*. Within the liver the only paramagnetic substance in a significant concentration is iron and so T2* can be used as a measure of liver iron content.[167]

2.5 T1 and T2* mapping

The terms 'T1 mapping' and 'T2* mapping' refer to MRI techniques that generates images with contrast derived from the T1 or T2* value for each pixel of the image. For example, to generate the T1 map shown in Figure 2.5-1 the T1 value for each pixel of the image has been calculated and displayed as a greyscale image. The brighter the pixel the higher the T1 value for that pixel. The black circle on the image is a region of interest (ROI) which can be placed on the image. The software will display the mean T1 value within that ROI.

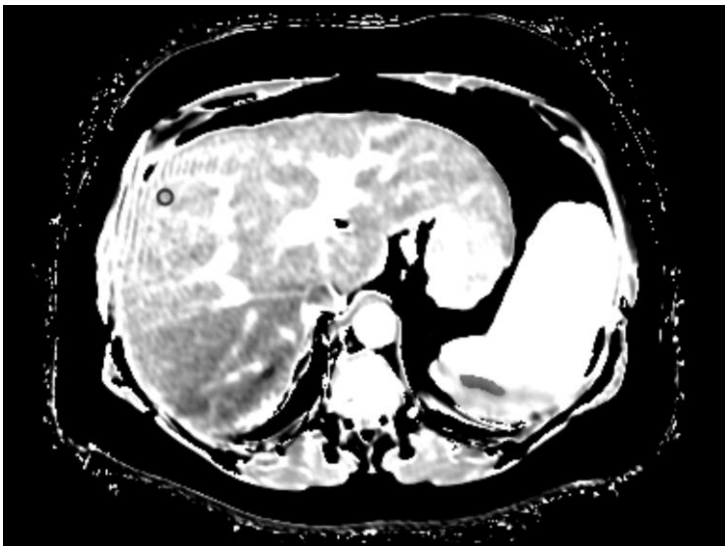


Figure 2.5-1: Representative T1 map with region of interest (black circle) for T1 measurement.

2.5.1 Known confounding effect of iron on T1 mapping

The presence of extracellular water within the liver tissue is not the only factor that influences T1. The presence of hepatic iron overload is known to reduce hepatic T1 due to its paramagnetic effects.[164, 167] This effect can be clearly seen in Figure 2.5-2 taken from the

work of Hoad et al, NMR Biomed, 2015. Patients with high levels of iron deposition had a shorter T1 than those without iron overload. This effect reached statistical significance in those with early stage fibrosis. This confounding effect of iron would have a detrimental effect on the ability of T1 to act as a reliable surrogate of hepatic fibrosis.

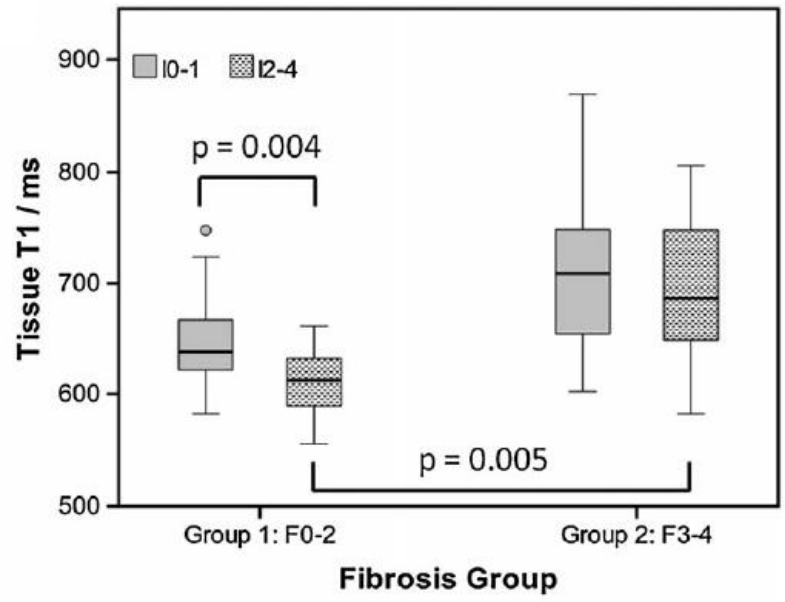


Figure 2.5-2: Box plot demonstrating that T1 is lower in patients with iron overload (hatched bars) than without iron overload (solid bars) across all fibrosis stages. Taken from: Hoad et al, NMR Biomed, 2015.[164]

2.6 Corrected T1

A novel approach to using MRI to stage hepatic fibrosis was taken by Banerjee R. and colleagues in a 2013 paper in the Journal of Hepatology.[163] In this work, measurement of T2* with a gradient echo sequence was used to quantify hepatic iron content. The measured T1 value was then corrected for the effects of iron excess with a patented algorithm to produce a corrected T1 (cT1). This eliminates the confounding effect of iron excess on T1 measurement and so cT1 can be thought of as the T1 value that would have been measure if liver iron was normal.

In a clinical study assessing this new technique 79 patients having a standard of care liver biopsy and 7 healthy volunteers underwent MRI scan and calculation of cT1. Liver biopsies were staged according to the Ishak system by expert pathologists blinded to the MRI findings. In this study cT1 was found to have a strong association with histologic fibrosis stage. This relationship is shown in

Figure 2.6-1 and there were statistically significant differences between all groups except Ishak 1-2 and Ishak 3-4.Perhaps the he most notable finding with this work was the ability of their MRI protocol to identify those with early fibrosis. AUROC for the differentiation of those with any liver fibrosis (Ishak ≥1) from those without fibrosis (Ishak 0 or healthy volunteer) was 0.94. This identification of early fibrosis had not previously been demonstrated with a non-invasive technique.

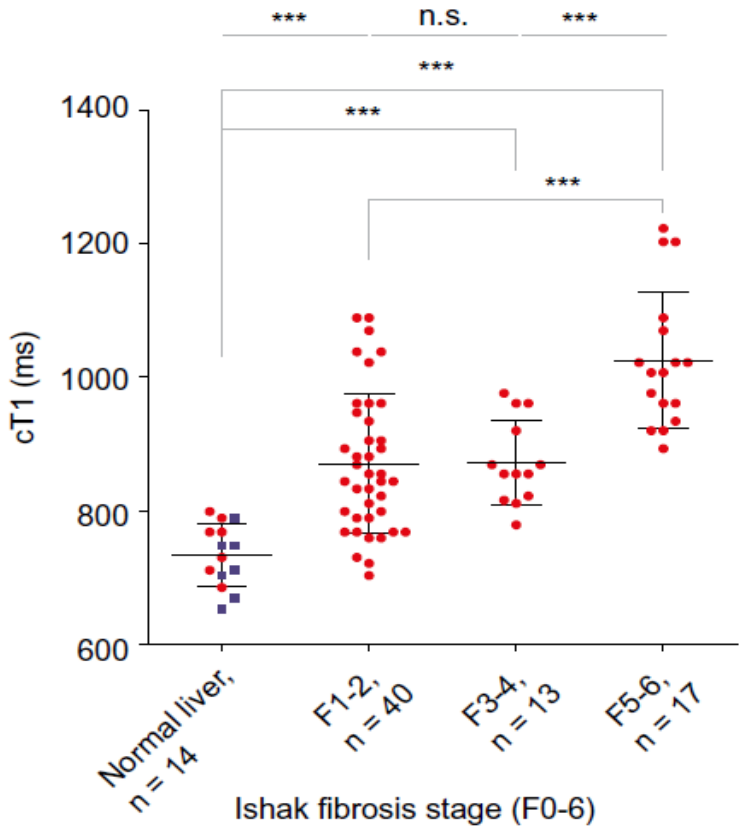


Figure 2.6-1: Scatter plot showing that cT1 correlates strongly with fibrosis stage (red circles – patients, blue squares – healthy volunteers). *** denotes significance at the $p<0.05$ level for inter-group differences by one way ANOVA with Bonferroni's correction. Taken from: Banerjee et al, J Hep, 2014.[163]

2.7 Liver*Multiscan*™

On the strength of the Banerjee et al paper the authors founded a University of Oxford spin out company called Perspectum Diagnostics Ltd. The algorithm used to generate cT1 from T1 and T2* measurements was patented and the development of this technology into a commercially available product for the assessment of hepatic fibrosis was started. This software product is called Liver*Multiscan*™ and is now CE marked and food and drug administration (FDA) approved. Liver*Multiscan*™ is a software programme that takes MRI data from freely available MRI sequences and processes these data to quantify cT1, fat fraction and liver iron. It also generates a colour coded cT1 map of the liver. Figure 2.7-1 shows examples of the cT1 maps generated by Liver*Multiscan*™.

The grant that is the main source of funds for the work contained in this thesis was awarded jointly to The University of Birmingham, The University of Edinburgh and Perspectum diagnostics to develop Liver*Multiscan*™ and validate to Banerjee paper and further investigate Liver*Multiscan*™ and cT1 as a potential biomarker of hepatic fibrosis.

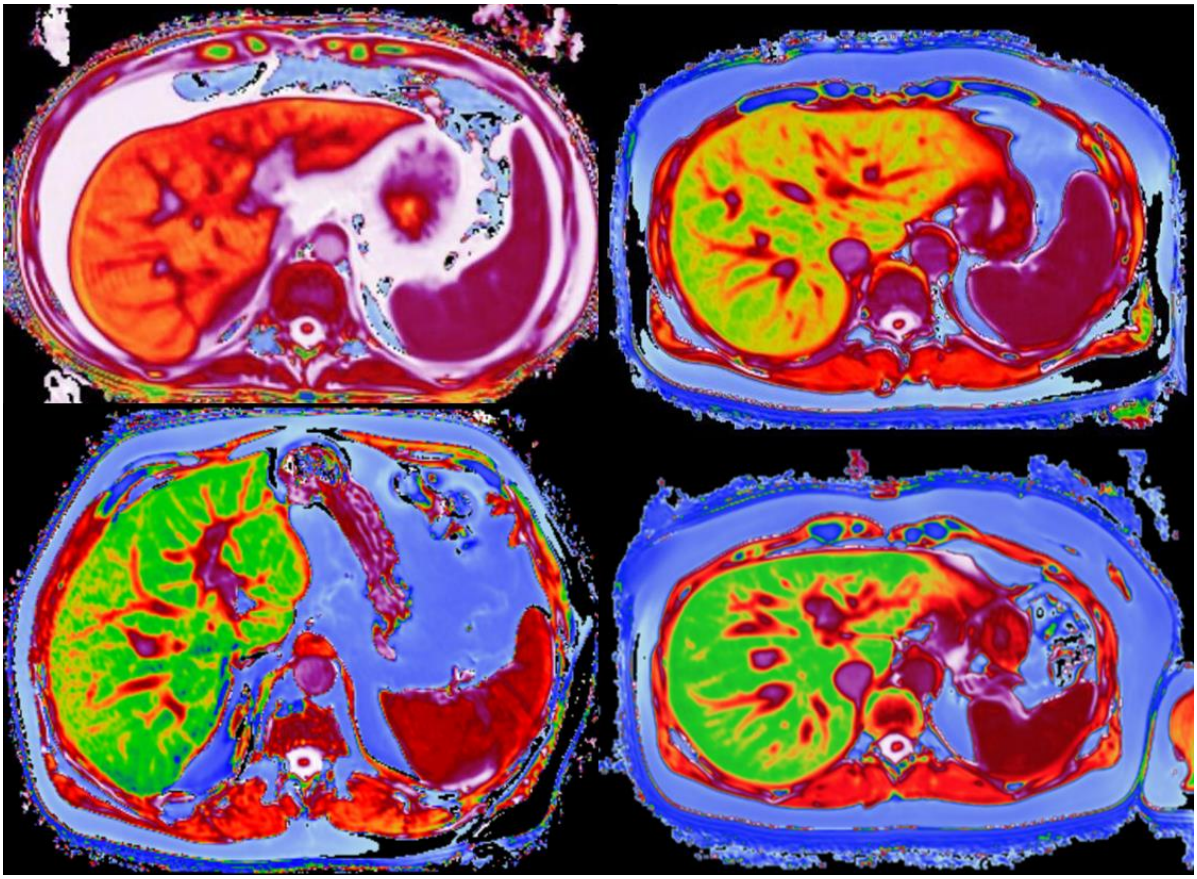


Figure 2.7-1: Sample cT1 maps generated by Liver*Multiscan*^{TM}. The liver parenchyma is colour coded according to the calculated cT1 value of each pixel. Low cT1 is represented by green increasing to yellow, orange and red for the highest cT1 values seen in the liver.

2.8 MRI data acquisition

In this work the shortened modified Look Locker inversion (ShMoLLI) sequence (5, 1, 1) is used to measure T1 and a high resolution Gradient Recalled Echo (GRE) sequence is used to measure T2*. The well-established Dixon sequence is used to measure proton density fat fraction (PDFF).

2.8.1 Patient preparation

All MRI scans were performed at 3 Tesla on Siemens Verio MRI scanners (Siemens Healthcare GMBH, Erlangen, Germany). The MRI protocol does not require intravenous contrast. The participant lies supine with 3 lead ECG for cardiac gating. A combination of body matrix and spine matrix coil elements was used to acquire data. Following localisers and shimming, the sequences include: ShMOLLI recovery sequence (T1 mapping), multi-gradient-echo sequence (T2* mapping) and modified Dixon sequence. All data were acquired during diastole with breath held in expiration to minimise movement artefact. Maps were acquired in a transverse plane through the liver hilum using the same slice position for each sequence.

2.8.2 T1 mapping

The ShMoLLI sequence calculates T1 by sampling the signal intensity at set intervals following 3 inversion pulses (5 samples after the 1st, 1 after the 2nd and 1 after the 3rd). These

measured signal intensities are plotted against time and the T1 relaxation curve fitted to the measured points. A schematic representation of this is shown in Figure 2.8-1.

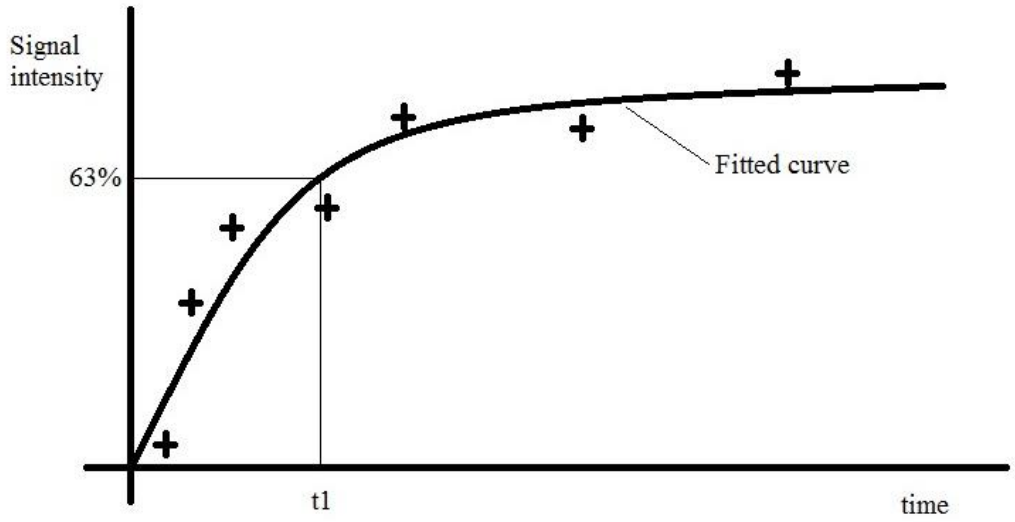


Figure 2.8-1: Schematic representation of the fitting of the T1 relaxation curve to measured points.

The ShMoLLI sequence has an advantage over other T1 mapping sequences due to the shorter duration of the required breath holds and the increased accuracy in patients with higher heart rates, which is a particular disadvantage of the MoLLI sequence.[168]

2.8.3 T2* mapping

The GRE sequence used in this work is a well-established technique for T2* mapping and, as with T1 mapping relies on the fitting of a T2* relaxation curve to a number of measured signal intensities at known time points.

2.8.4 Dixon sequence

The Dixon sequence is a well validated technique to measure liver fat and has been shown to correlate strongly with histological assessment of steatosis.[169-172] Quantification of liver fat with the Dixon sequence relies on the fact that fat and water molecules precess at different rates. This means that fat and water de-phase at different rates and the acquisition of images at slightly different times generates images that have fat and water molecules in and out of phase with each other. Combining these mages allows the fat content to be calculated.[173]

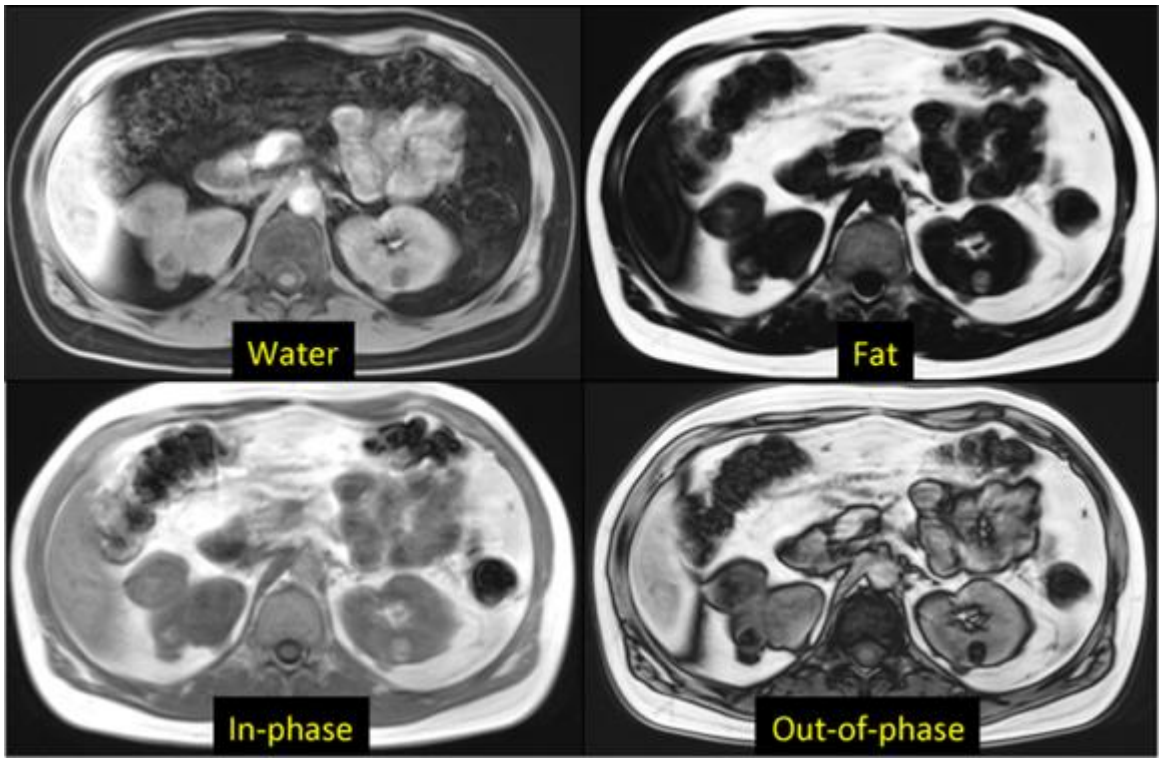


Figure 2.8-2: The 4 images generated by the Dixon sequence. Image courtesy of Allen D. Elster, MRIquestions.com.

2.9 MRI data analysis

MRI data were acquired at the study sites and recorded in the industry standard DICOM file format. The DICOM files were anonymised and sent to Perspectum Diagnostics by secure electronic transfer. Once at Perspectum, T1 and T2* and Dixon maps were analysed using Liver*Multiscan*^{TM} software by employees of Perspectum Diagnostics. The operator analysing the scans was blinded to the clinical findings and biopsy results. To generate cT1 values, a region of interest (ROI) of approximately 1.4cm^3 was placed in a representative area of the right lobe of the liver avoiding vascular and biliary structures on both the T1, T2* and Dixon maps. Liver*Multiscan*^{TM} software then calculates a value for cT1, iron content and fat fraction.

2.10 Reproducibility of cT1 mapping

The use of cT1 to stage liver fibrosis is an emerging technique and little is known of the reproducibility of the test or the stability of hepatic cT1. Some existing non-invasive techniques for assessment of hepatic fibrosis such as VCTE and MRE are known to be less accurate in the post prandial state. We undertook studies with healthy volunteers to test the reproducibility of cT1 measurement and the stability of hepatic cT1 over time and in response to eating.

Healthy volunteers were recruited from colleagues and students at the University of Birmingham. To ensure that volunteers did not have undiagnosed liver disease the following exclusion criteria were used: history of liver disease, significant medical co-morbidity (as judged by the study team), presence of features of the metabolic syndrome, body mass index (BMI) >30kg/m^2 and alcohol consumption in excess of 21 units/week for men or 14 units/week for women.

2.10.1 Test, re-test

In this study 5 volunteers were scanned as described in Chapter 2 and then taken out of the scanner. They were immediately taken back into the scanner and re-scanned using the same acquisition protocol. This immediate test and re-test study assesses the reproducibility of the acquisition protocol and data analysis. It is assumed that hepatic cT1 would not change in the short time between scans.

Of the 5 volunteers, all were Caucasian and 3 were female. Mean (±SD) age was 38.8 (±9.8) years. Mean (±SD) BMI was 23.7 (±2.5) Kg/m^2. No volunteer was a current smoker. One volunteer was an ex-smoker with an 8 pack year history. All volunteers consumed alcohol, with a range of intake from 1-16 UK units/week. No volunteer consumed alcohol outside of recommended limits. Mean (±SD) time between the scans was 43 (±3) minutes.

Mean (±SD) cT1 in scan 1 was 768.7 (±71.8) msec and in scan 2 was 793.6 (±35.6) msec. A paired sample t-test shows the difference in cT1 between first and second scans to be non-significant (p=0.307). The coefficient of variation between the first and second scan for each individual was calculated and the mean (±SD) for these values was 3.7 (±3.5) %. The absolute change in cT1 between scans for each individual volunteer is shown in Figure 2.10-1. This demonstrates that although overall there was no statistically significant difference in cT1, one volunteer does have a striking increase in cT1 from 657.8 msec to 754.8 msec. This is a 14.7% change in cT1 between the 2 scans, which is markedly higher than the mean (±SD) % change in cT1 for all volunteers (5.4 (±5.4) %) and the remaining 4 volunteers alone (3.0 (±1.7) %).

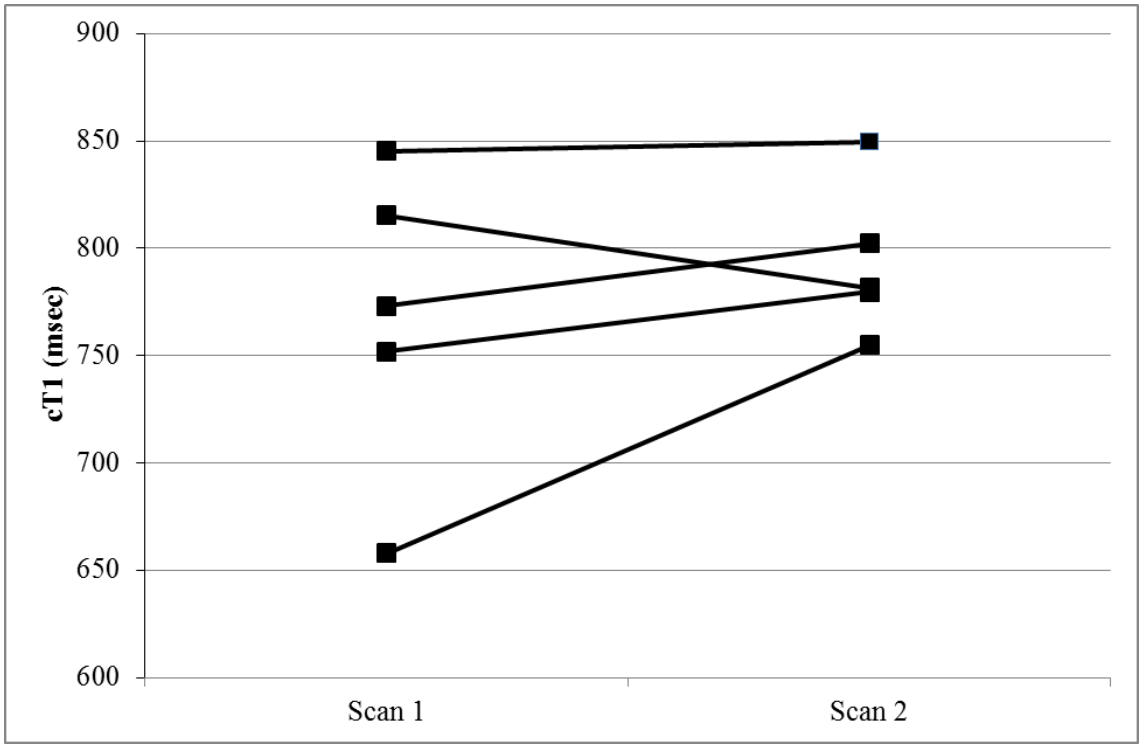


Figure 2.10-1: Absolute cT1 values in the test, re-test study demonstrating the change in cT1 for individual volunteers.

This study suggests that measurement of cT1 is reproducible and has low inter-test variability with mean coefficient of variation of 3.7%. One volunteer however had a noticeably greater rise in cT1 between scans than the other 4 volunteers. The reason for the larger change in cT1 is not clear. The quality of the acquisition and particularly the T2* map can degrade the accuracy of measurement, however all MRI acquisitions in this study were of good quality. The most likely source of variation is the placement of the ROI used to take the T1 or T2* measurement. Although the operator was blinded to the circumstances of the scan, this procedure is operator dependent and therefore a potential source of variation. The use of automated systems to place the ROI or analysis of the whole liver slice should reduce operator dependence and improve reproducibility further.

A limitation of this study is the small sample size used in this study. It is difficult to know how the distribution in a small sample such as this reflects the population distribution. It may be that a change in cT1 of 14.7% is very unusual in the population as a whole or it may be closer to the norm. A repeat of this study with a larger sample size and more consistent ROI placement could be conducted to better define the reproducibility of cT1 measurement.

2.10.2 Fasted and fed

The portal venous and hepatic arterial flow increases greatly in the fed state compared to fasting.[174] This increase in flow causes a change in the liver parenchyma known as postprandial hyperaemia.[175, 176] These changes are known to adversely affect the accuracy of assessment of fibrosis with elastography techniques such as VCTE and MRE.[97, 175, 176] We sought to determine the changes in hepatic cT1 that occur following eating with a view to establishing if fasting before cT1 measurement was necessary to ensure reliable results are obtained.

8 volunteers were scanned between 09:30 and 10:00 following an overnight fast. Drinking water was allowed at any time. After the first scan, volunteers continued with their usual daily activities and returned for a second scan between 14:00 and 15:30 the same day. It has been shown that the increase in liver blood flow due to eating occurs quickly and lasts for up to 180 minutes.[97] To ensure that scans were completed within this window, volunteers started their lunch no more than 60 minutes prior to their MRI scan. The meal consumed was not standardised. For this study, volunteers were included with BMI outside the range specified for other reproducibility studies.

Of the 8 volunteers, all were Caucasian and 5 were female. Mean (±SD) age was 34.9 (±4.4) years. Mean (±SD) BMI was 25.9 (±4.6) kg/m^2. All volunteers were never smokers. All volunteers consumed alcohol, with a range of intake from 1-15 UK units/week. No volunteer consumed alcohol outside of recommended limits. Mean (±SD) time between the scans was 4 hours 58 minutes (±20 minutes). All volunteers reported eating their typical lunch 45-60 minutes before the scan.

Mean (±SD) cT1 in the fasted scan was 771.1 (±47.1) msec and in the fed scan was 778.7 (±47.6) msec. A paired sample t-test shows this difference to be non-significant (p=0.347). The coefficient of variation between the fasted and fed scan for each individual was calculated and the mean (±SD) for these values was 1.2 (±1.5)%. Of note, the mean coefficient of variation in this study is lower than in the test, re-test study (3.7 (±3.5)%). The absolute change in cT1 for individual volunteers is shown in Figure 2.10-2. This shows that although overall there was no statistically significant difference in mean cT1, one volunteer does have a marked increase in cT1 from 778.6 msec to 835.5 msec. This was not the same volunteers as discussed in the test, re-test study.

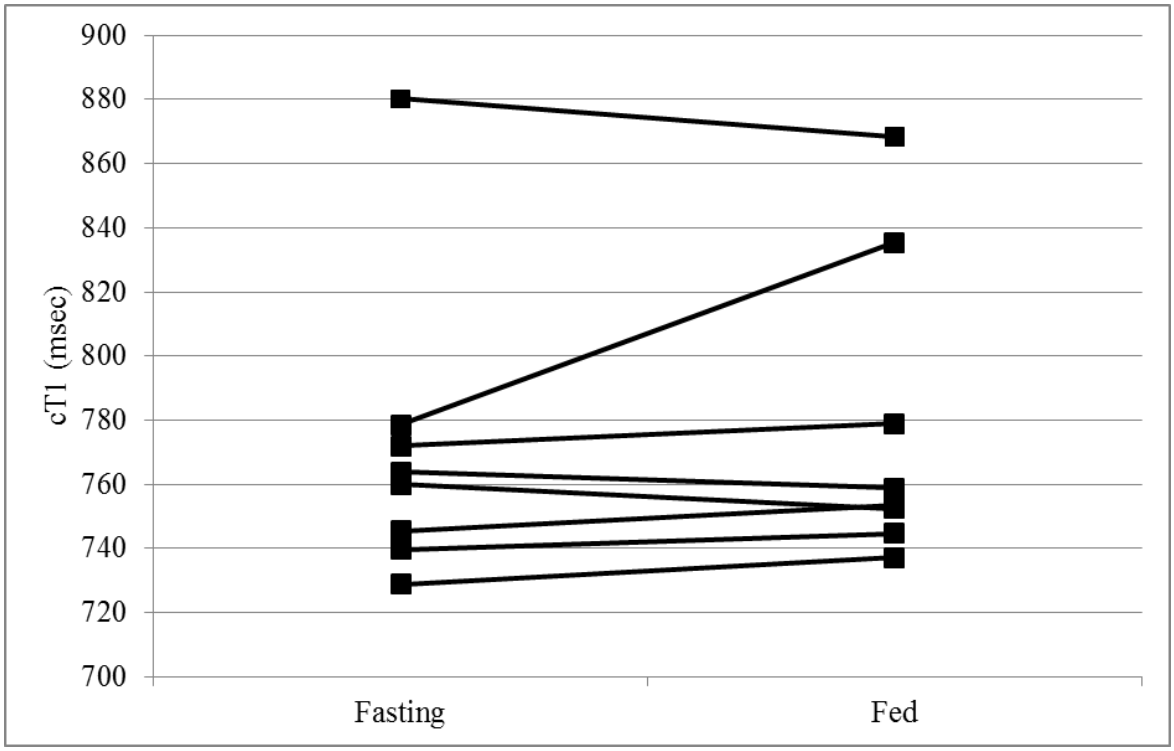


Figure 2.10-2: Absolute cT1 values in the fasted and fed study demonstrating the change in cT1 for individual volunteers.

These data indicate that cT1 in the fed state is not different to cT1 in the fasted state and it implies that fasting is not required before measurement of cT1. This conclusion should however be made with caution as one patient did have a marked increase in cT1 in the fed state. It is again not clear why this should have occurred for this one individual. The same explanation for variation may apply as in the test re-test study with changes in the ROI placement explaining the variation.

A weakness of this study is the lack of meal standardisation. The effects of the fed state on cT1 could be underestimated by some individuals eating small or low calorie meals. The one volunteer in whom a marked increase in cT1 was observed may have eaten a particularly high calorie meal. Another factor that may be having an effect is the baseline condition of the liver.

The effects of a standardised test meal on FibroScan measurements became more pronounced as underling fibrosis stage increased.[175] As the volunteers in this study were not assessed for the presence of liver fibrosis with any validated test, it is possible that underlying fibrosis stage may have influenced the result.

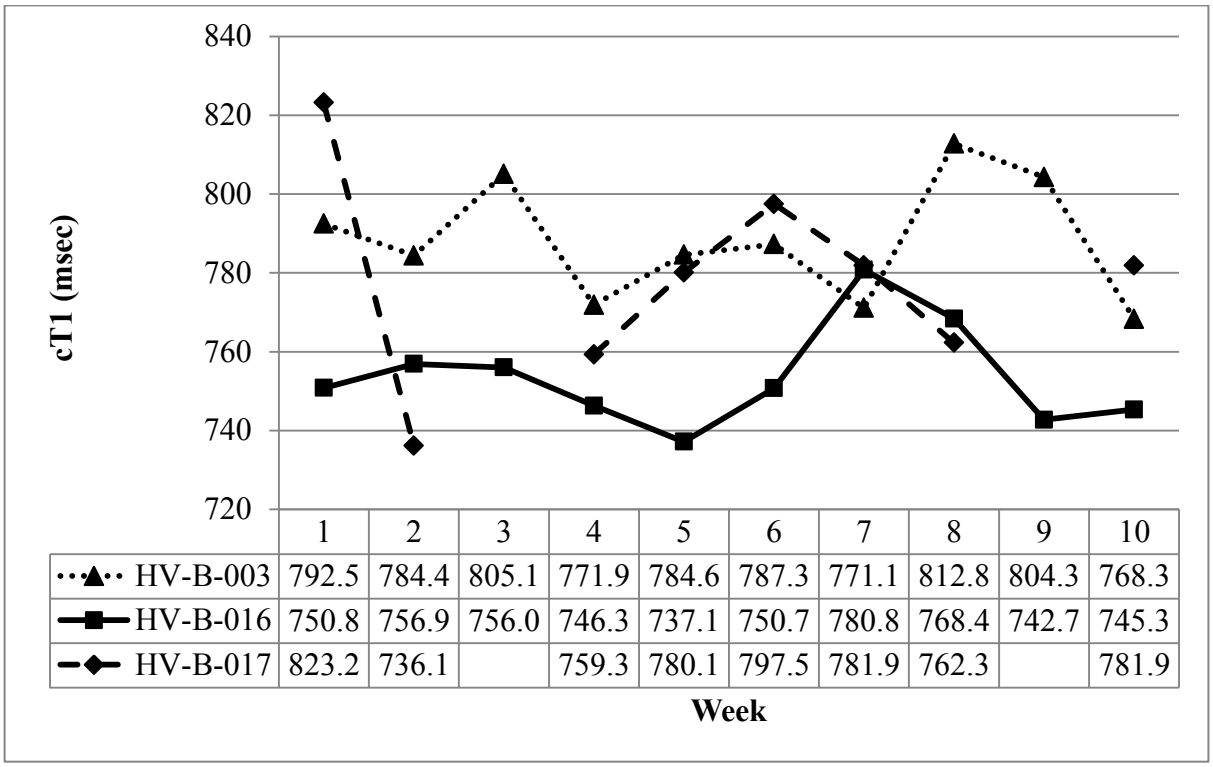
2.10.3 Stability of cT1 over time

In sections 2.10.1 and2.10.2, it has been demonstrated that the inter-test variation of cT1 measurement is low. However, the chronological variation in hepatic cT1 has not to my knowledge been studied. cT1 may vary from day to day and this raises a question about the interpretation of small changes in cT1 on follow-up scans.

To establish the natural variation in cT1 over time, 3 volunteers underwent cT1 mapping on a weekly basis for 10 weeks. These scans were performed as described in Chapter 2 and, for consistency, occurred in the early afternoon in each case. As shown above the fasting status of the volunteer should not be expected to alter the measured cT1.

The 3 volunteers were all male with mean (±SD) age of 33 (±3.6) years. Mean (±SD) BMI was 25.4 (±1.0) Kg/m^2. All volunteers were never smokers. All volunteers consumed alcohol, with a range of intake from 5-15 UK units/week. No volunteer consumed alcohol outside of recommended limits. Due to unforeseen circumstances, HV-B-017 was unable to have a scan on weeks 3 and 9. No volunteer experienced any significant intercurrent illness or had any change to BMI or alcohol intake during the 10 week period.

Mean (±SD) cT1 for HV-B-003, HV-B-016 and HV-B-017 was 788.2 (±14.7), 753.5 (±12.3) and 777.8 (±24.5) msec respectively. The variation in the absolute cT1 values over the 10 week period is shown in Figure 2.10-3.



	1	2	3	4	5	6	7	8	9	10
HV-B-003	792.5	784.4	805.1	771.9	784.6	787.3	771.1	812.8	804.3	768.3
HV-B-016	750.8	756.9	756.0	746.3	737.1	750.7	780.8	768.4	742.7	745.3
HV-B-017	823.2	736.1		759.3	780.1	797.5	781.9	762.3		781.9

Figure 2.10-3: Absolute cT1 values for 3 volunteers over a 10 week period demonstrating the natural variation in cT1 over time.

The coefficient of variation for HV-B-003, HV-B-016 and HV-B-017 was 2.0%, 1.7% and 3.4% respectively. Mean (±SD) coefficient of variation was 2.4 (±0.9) %. This study demonstrates that variation in cT1 over time is small. The coefficient of variation for all three volunteers is within the mean coefficient of variation in the test, re-test study.

2.10.4 Discussion

In Summary, these studies have shown that T1 measurement and cT1 calculation with our technique is reproducible, stable over time and does not need to be performed fasted. These results are relevant to the application of this test in clinical and research practice. These studies test the whole process of producing a cT1 value including both measurement and data analysis. It is likely that the variation seen in two of the test subjects is introduced during data analysis and relates to the position of the ROI. The next step in assessing the reproducibility of this test would be to repeat these studies with standardised ROI placement. An increase in the sample size would also increase the reliability of this conclusion.

2.11 cT1 for healthy individuals

cT1 mapping for staging liver disease is a little studied technique and the range of values expected in normal individuals is currently documented in one published study only. In the 2014 Banerjee et al paper the mean (±SD) cT1 for the 7 healthy volunteers was 717 (±48) msec and 7 patients with Ishak 0 fibrosis was 750 (±42).[163] It is clear that this is an insufficient sample upon which to reliably define a reference range for this technique. To my knowledge there are no other published studies outlining a reference range for hepatic cT1 or indeed for uncorrected T1 at 3T. The field strength is important as T1 increases with the strength of the standing magnetic field. This means that the value for healthy volunteers in the Banerjee study is not comparable to two published studies looking at hepatic T1 in healthy individuals at 1.5T. As would be expected for a lower field strength, these studies found a shorter mean (±SD) T1 of 678 (±45) msec (n=31)[177] and 645 (±44) msec (n=14).[164] Although field strength is important, the fact that these studies do not use the iron correction algorithm incorporated into Liver*Multiscan*™ should not affect the comparability of the results. In a patient with normal liver iron, there should be no correction from with algorithm and therefore T1 should equal cT1.

2.11.1 Current upper reference limit for cT1

The upper limit of the reference range for cT1 used by Perspectum is 822 msec. This value is derived from the mean (±SD) cT1 of the 14 patient in the 'normal liver' group (healthy volunteers and patients with Ishak 0 fibrosis) in the 2014 Banerjee et al paper (Personal

communication: Dr M. Kelly, Head of Innovation, Perspectum Diagnostics ltd, 07 Feb 2018). The mean (±SD) cT1 values for this combined group are not given in the paper.

2.11.2 Defining a new upper reference limit for cT1

In the studies described in this thesis and the initial development of the protocol a total of 103 scans were performed on 34 individual healthy volunteers. These healthy volunteers were recruited for a range of reasons including: development of the study protocol, the reproducibility studies discussed in section 2.10 and as controls for the studies described in Chapters 3 and 4. There were also a number of volunteers recruited from the University of Edinburgh for similar reasons. The criteria for defining a 'healthy' volunteer, the MRI acquisition protocol and the method of data processing was identical at the two sites. In addition the scanners were calibrated on the same MRI phantom prior to the start of the study. This means that data from the two sites should be directly comparable.

Healthy volunteers were recruited from colleagues and students at the University of Birmingham and University of Edinburgh. To ensure that volunteers did not have undiagnosed liver disease the following exclusion criteria were used: history of liver disease, significant medical co-morbidity (as judged by the study team), presence of features of the metabolic syndrome, body mass index (BMI) >30kg/m^2 and alcohol consumption in excess of 21 units/week for men or 14 units/week for women. The number of scans per healthy volunteer ranged from 1 to 11 scans. The majority of healthy volunteers had 2 scans with the frequency of multiple scans shown in Table 2.11-1.

Number of scans per healthy volunteer	Frequency
1	6
2	21
3-4	2
8-11	5

Table 2.11-1: The distribution of the number of scans per healthy volunteer.

Taking the healthy volunteer cohort as a whole, the distribution of cT1 was not normal. Median (IQR) cT1 for healthy volunteers was 771.1 (750.7-805.4) msec. Minimum and maximum values were 657.8 and 951.5 msec respectively and can be seen in Figure 2.11-1.

In this cohort, individual healthy volunteers having multiple scans has had an impact on the distribution of cT1. For example HV-E-010 has a hepatic cT1 at the upper end of the observed range in the majority of their scans. These values are highlighted in red in Figure 2.11-1. The reason for the high cT1 values in this individual is not clear. They are not an outlier in terms of their demographics or clinical parameters. Other measures of liver health such as biochemical tests or transient elastography were not performed. Whatever the reason for their relatively high cT1 the fact that this individual underwent 11 scans (the most of any volunteer) influenced the distribution.

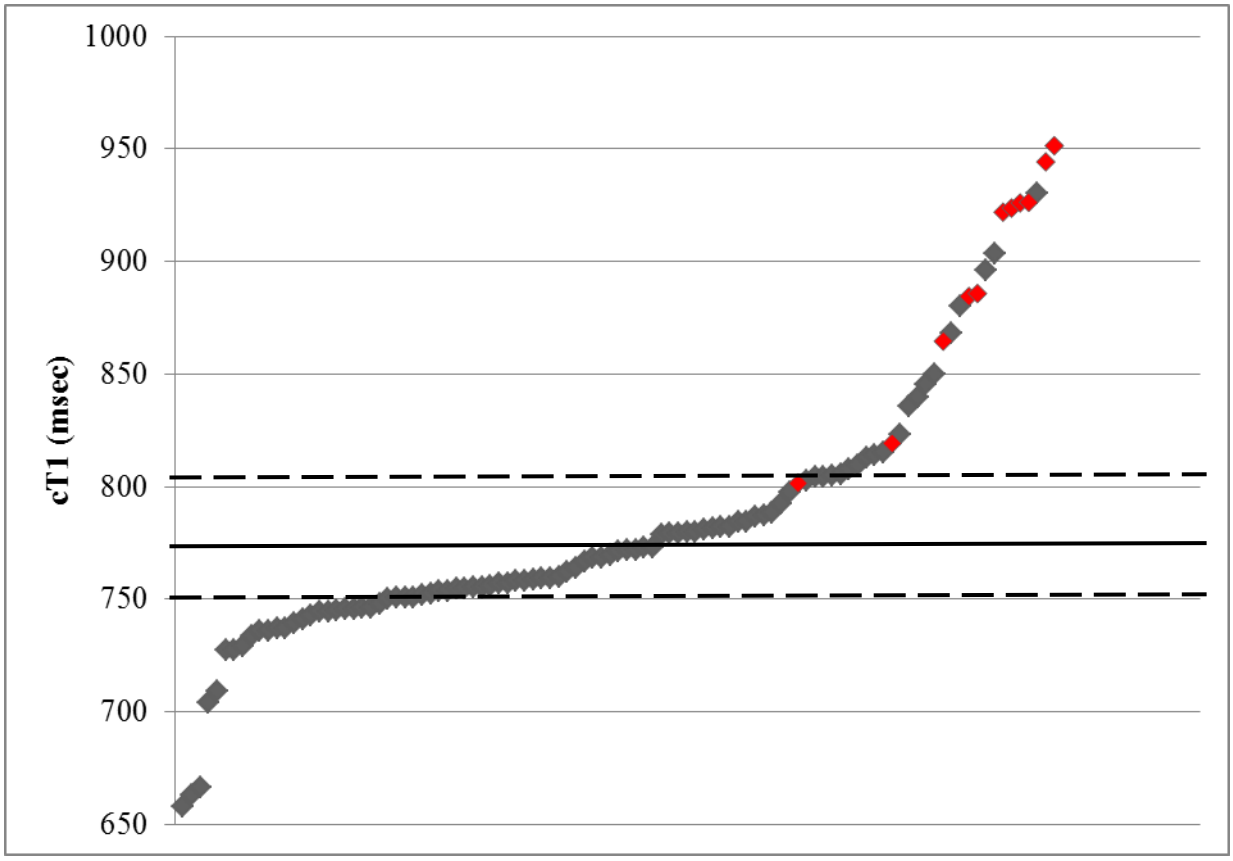


Figure 2.11-1: Scatter plot of cT1 values from the studied healthy volunteers. The points in red are from a single individual (HV-E-010). Solid line: median, dashed lines 25th and 75th centiles.

Excluding the values from HV-E-010 gives cT1 in healthy volunteers a distribution that approximates to normal as seen in Figure 2.11-2. Mean (±SD) cT1 in this group is 772.9 (±45.8). Using the typically defined normal range of 1.96 standard deviations from the mean [178] makes the upper limit of the reference range 862.6 msec.

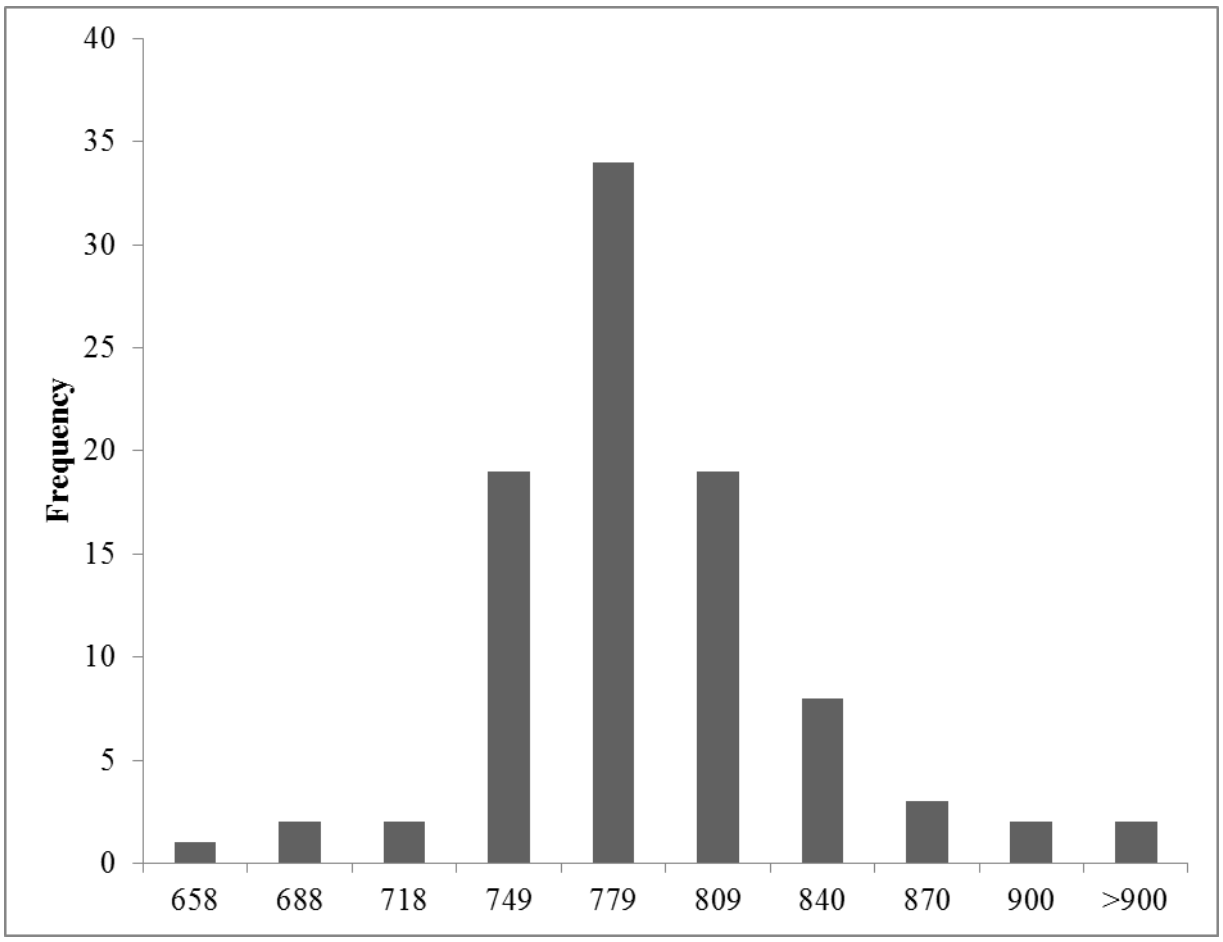


Figure 2.11-2: Histogram demonstrating the distribution of cT1 in the healthy volunteer population with HV-E-010 excluded. The values on the x-axis are the upper border of the group.

2.11.3 Discussion

The upper limit of the reference range for cT1 defined in this study (863 msec) is clearly different to the one previously defined using data from Banerjee et al (822 msec). This study has a much larger sample size and so would be expected to provide a more reliable reference range. It is noticeable that the cT1 values seen in healthy volunteers in the Banerjee paper are low compared to the values seen in this cohort. The reason for this is not clear. There may be

unidentified differences in data acquisition but it may simply reflect the small sample size of the Banerjee paper. In addition to the small sample size the 822 msec cut-off may be unreliable as it is defined from a group where 7 of the 14 were not healthy volunteers but patients with Ishak 0 fibrosis. Healthy volunteers and patients with Ishak 0 fibrosis are clearly not the same and I do not think they should be analysed together.

This study has some limitations. Repeated scans give too much weight to the cT1 values of individuals with more scans. This causes the distribution to skew and may have altered the reference value defined by this study. In addition, healthy volunteers were not all assessed for signs of liver disease by established methods. Efforts were made to recruit volunteers without risk factors for liver disease but the majority did not undergo blood tests or alternative fibrosis markers to 'prove' that they did not have undiagnosed liver disease.

However, limitations notwithstanding, the large sample size used in this study provides an important contribution to the available knowledge of cT1 in healthy individuals.

2.12 Proton Magnetic Resonance Spectroscopy

Proton Magnetic Resonance Spectroscopy (^{1}H-MRS) is a MR technique that can non-invasively identify substances within living tissue and measure their relative concentration. This ability to identify different substances is investigated in Chapter 5 as a potential biomarker in non-alcoholic fatty liver disease.

2.12.1 Background

The spin of protons and in particular the precession of their axis has previously been discussed in terms of T2 relaxation. Within a static magnetic field, the rate of precession of a proton (known in the field of MR spectroscopy as the 'resonant frequency') is proportional to two factors. The first is the strength of the magnetic field and the second is the chemical milieu in which a proton exists. For example, within a static magnetic field, the protons within a water molecule ($\mathbf{H}_2O$) have a different resonant frequency to those in a methane molecule ($C\mathbf{H}_4$) and within methanol (CH_3OH) the protons in the methyl ($-C\mathbf{H}_3$) group have a different resonant frequency to the proton in the hydroxyl ($-O\mathbf{H}$) group. This is due to the different bonding atom involved, the bond length and the adjacent bonds. This difference in resonant frequency between different protons is the principle on which ^{1}H-MRS is based. During a proton magnetic resonance spectroscopy (^{1}H-MRS) sequence, the magnetic field within the examined volume of tissue is homogeneous and thus the only factor influencing resonant frequency are the bonds in which protons are involved.

The strength of the static magnetic field will vary slightly between different scanners due to slight variations in the manufacturing of the magnet coil. Due to the fact that resonant frequency is dependent on field strength, it is not possible to directly compare ^{1}H-MRS results from different scanners without first standardising for field strength. For this purpose the concept of 'chemical shift' is used. Chemical shift is the resonant frequency of the sample normalised to the resonant frequency of a standard molecule (usually tetramethylsilane) and is expressed in parts per million (ppm). The acquired signal from a ^{1}H-MRS sequence is displayed as a histogram showing the relative abundance of protons on the y axis (arbitrary units) against chemical shift on the x axis (ppm). Figure 2.12-1 shows example spectra from patients with high and normal liver fat.

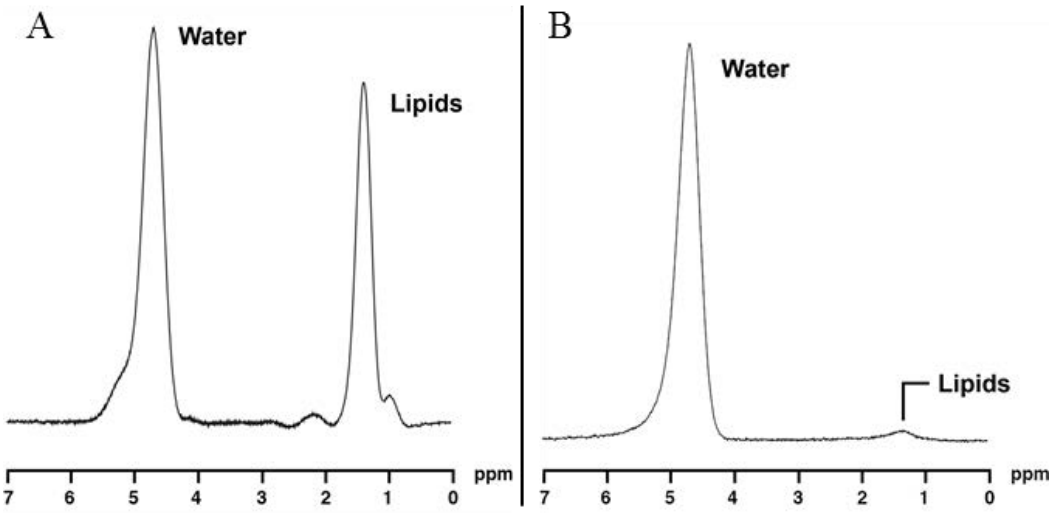


Figure 2.12-1: Simplified ^{1}H-MRS spectra showing: A) A patient with high hepatic fat content and B) normal hepatic fat content. Image adapted from Radiological Society of North America (RSNA) press release, 16/07/2013.[179]

The abundance of protons of a particular resonant frequency and therefore the concentration of a particular molecule within the examined volume of tissue is calculated by measuring the area under each peak. As the y axis on the spectrum is in arbitrary units the area under each peak cannot be directly compared between patients and must be normalised before comparison. It is conventional to normalise to the water peak and so the amount of lipid in a sample is found with the equation:

$$lipid\ fraction = \frac{area\ under\ lipid\ peak}{area\ under\ lipid\ peak\ +\ area\ under\ water\ peak}$$

2.12.2 Water suppression

Within liver tissue the majority of protons are contained in water with lipid making up the majority of the remainder. For patients with relatively low levels of liver fat the water signal is so large that the signal from the protons in fat is difficult to quantify. For this reason water suppression pulses are used during signal acquisition to de-phase water molecules and therefore prevent the acquisition of signal from water. This allows the fat signal to be more clearly seen. However, water suppression is not precisely directed at water and may also reduce the signal acquired from molecules with a similar resonant frequency. In practice this may reduce the size of peaks that are adjacent to the water peak and introduce error to the measurements.

An example for water suppressed and non-water-suppressed MR spectra are shown in Figure 2.12-2. This figure shows the water peak and the peaks from different groups within fat molecules numbered 1 to 7. The origin of these peaks is shown in Table 2.12-1.

Peak number (Figure 2.12-2)	**Chemical shift (ppm)**	**Group name**	**Group structure**	**Protons per group**
1	5.3	Olefinic	-C**H**=C**H**-	2
Water	4.7*	Water	$\mathbf{H}_2O$	2
2†	2.8	Diacyl	-CH=CH-C$\mathbf{H}_2$-CH=CH-	2
3	2.2	α-Carboxyl	-COOH-C$\mathbf{H}_2$-CH_2-	2
4	2.0	α-Olefinic	-C$\mathbf{H}_2$-CH=CH-C$\mathbf{H}_2$-	4
5	1.6	β-Carboxyl	COOH-CH_2-C$\mathbf{H}_2$-	2
6	1.3	Methylene	-C$\mathbf{H}_2$-	2
7	0.9	Methyl	-C$\mathbf{H}_3$	3

** Measured as the sum of two peaks(narrow and broad) to provide more accurate fitting*
† Peak often difficult to identify and is not well seen in Figure 2.12-2

Table 2.12-1: Identification of peaks in ^{1}H-MRS spectra of the liver. The first column refers to the labelled peaks in Figure 2.12-2. The protons highlighted in bold in the column headed "Group structure" shows the protons contributing to the magnitude of the peak. The number of protons per group is from a customised basis set based on the work of Hamilton et al.[180] The magnitude of each peak is based on the abundance of protons with a particular chemical shift, dividing the peak magnitude by the number of protons in each group gives the number of groups and the number of groups that occur only once per FA chain (eg. methyl groups) gives the number of chains.

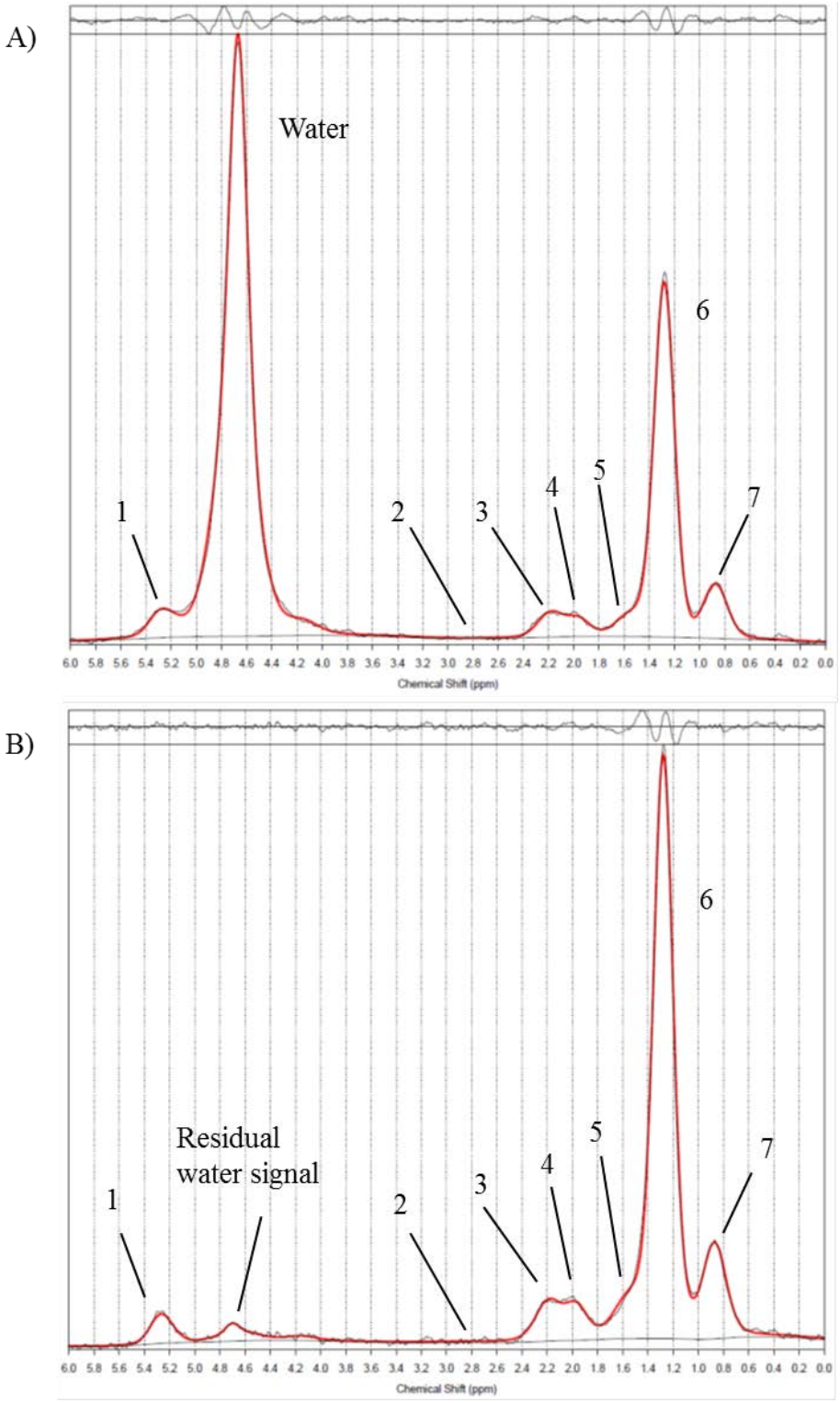


Figure 2.12-2: MR spectra from the same voxel showing A) Non-water-suppressed spectrum and B) Water-suppressed spectrum. The black line is the measured spectrum with the overlying red line showing the fitted spectrum. For key to peaks see Table 2.12-1.

2.12.3 ^{1}H-MRS data acquisition and processing

^{1}H-MRS measurements were performed on a 3 Tessla Siemens Verio MRI scanner (Siemens Healthcare GMBH, Erlangen, Germany). High order, automated shimming was performed and localisers in 3 planes were used to place a 2 x 2 x 2 cm voxel (volume in which ^{1}H-MRS is measured) in the right lobe of the liver avoiding large biliary and vascular structures. ^{1}H-MRS data from water-suppressed (WS) and non-water-suppressed (WREF) Stimulated Echo Acquisition Mode (STEAM) acquisitions (Repeat time 3 sec, echo time 20 msec; 5 measurements of 1 signal average for both) were processed and analysed by Mr Robert Flintham (RBF) (medical physicist with experience of in vivo MRS) using TARQUIN software.[181] TARQUIN performs automated phase and frequency correction, pre-processing, and fitting of the fat/liver spectrum. Visual quality control of fitted spectra was performed by RBF and poorly fitted spectra were excluded from the analysis. TARQUIN was used to select the individual peaks in turn to measure the area under the peak. An exception to this is that the 1.6 and 2.2 ppm peaks (protons α and β to carboxyl groups) were fitted and measured simultaneously. An example of the fitting of the peaks at 1.3 and 0.9 ppm is shown in Figure 2.12-3 the area under the fitted curve measured by TARQUIN gives the relative concentration of protons with that resonant frequency.

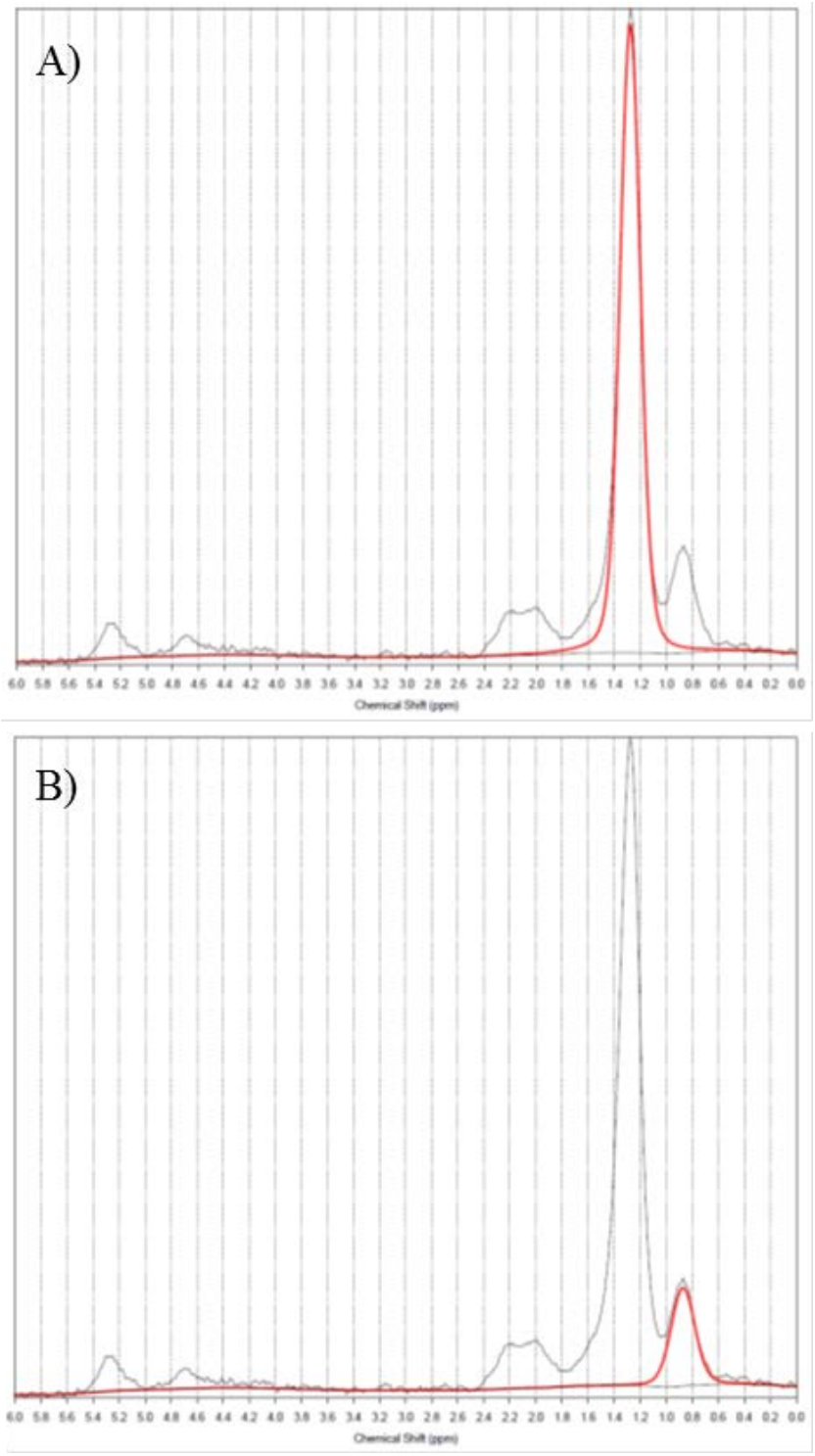


Figure 2.12-3: Example of the fitting individual peaks in a water-suppressed spectrum. A) 1.3 ppm (methylene) peak. B) 0.9 ppm (methyl) peak.

2.12.4 Calculation of fatty acid characteristics

When looking at fatty acid composition it is necessary to identify the number of fatty acid chains within a sample and use this to normalise the measurements. With knowledge of the basic structure of a FA (Figure 2.12-4) it can be seen that the methyl, α-carboxyl and β-carboxyl groups occur only once in any given FA and so the abundance of these groups can be used to estimate the total number of FA chains present.

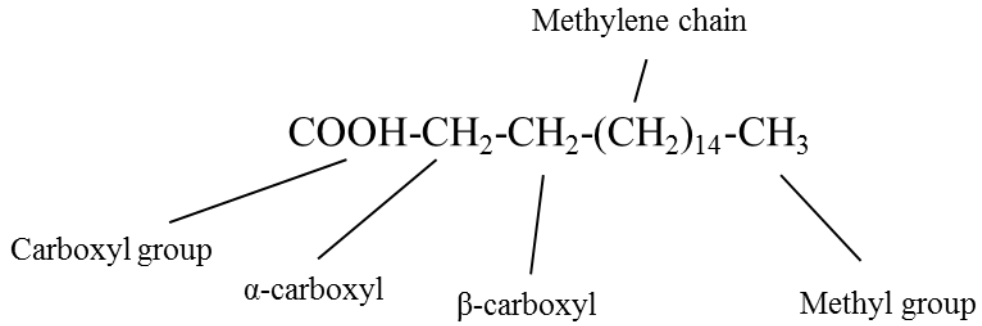


Figure 2.12-4: Basic structure of a simple fatty acid; in this case stearic acid.

The relative magnitude of each spectral peak is based on the number of protons rather than the number of groups. A basis set was customised based on Hamilton et al[180] to define the number of proton within each chemical group (Table 2.12-1) and thus define the relative number of groups present.

Thus, the number of fatty acid chains to act as a denominator can be calculated from the formula:

$$number\ of\ chains = \frac{area\ under\ 0.9, 1.6\ and\ 2.2\ ppm\ peaks}{(3 + 2 + 2)}$$

With a known denominator, the required parameters can be calculated with the following formulae:

$$Mean\ chain\ length\ (mCL)\ = \frac{area\ under\ all\ fat\ peaks}{number\ of\ chains}$$

$$PUFA = \frac{(area\ under\ 2.8\ ppm\ peak\ \div 2)}{number\ of\ chains}$$

$$Double\ bonds\ per\ chain\ (nDB) = \frac{(area\ under\ 5.3\ ppm\ peak\ \div 2)}{number\ of\ chains}$$

CHAPTER 3: MAGNETIC RESONANCE T1 MAPPING FOR THE STAGING OF HEPATIC FIBROSIS

3.1 Disclosure

This data has been written up and submitted for publication in Scientific Reports. Dr Peter Eddowes is a joint first name author on this paper. The analysis presented here was performed by Dr Peter Eddowes and is independent of the submitted manuscript.

3.2 Aims

As outlined in Chapter 2, T1 mapping has potential utility in the staging of hepatic fibrosis. Liver*Multiscan*™ (Perspectum Diagnostics, Oxford, UK) has been developed to overcome the known confounding effects of hepatic iron overload. In this chapter data are presented that aim to prospectively evaluate the ability of Liver*Multiscan*™ to identify and stage hepatic fibrosis.

In the first study (Section 3.3) Liver*Multiscan*™ is evaluated using liver biopsy histology as the reference standard. As a secondary aim, the performance of Liver*Multiscan*™ for staging hepatic fibrosis will be compared to existing non-invasive biomarkers.

Liver biopsy histology has the potential for sampling error and this is a well-known limitation of using liver biopsy histology as the reference standard. The second experiment (Section 3.4) aims to control for this limitation by using multiple samples from hepatectomy specimens as the reference standard.

3.3 cT1 as a marker of fibrosis in chronic liver disease: a prospective validation study of Liver*Multiscan*™

3.3.1 Methods

This study conforms to the research ethics guidelines of the 1975 declaration of Helsinki. It has received approval from the national research ethics service (14/WM /0010) and local research and development offices at the study sites. The study is registered with the ISRCTN registry (ISRCTN39463479) and the National Institute of Health Research (NIHR) portfolio (15912).

Participants

All adult patients booked for a non-targeted liver biopsy for any indication at Queen Elizabeth Hospital Birmingham and Royal Infirmary of Edinburgh between February 2014 and September 2015 were invited to take part in the study. Exclusion criteria were: biopsy of a distinct focal lesion, inability to give fully informed consent and any contraindication to MRI. All patients who took part gave fully informed written consent and then underwent research MRI, FibroScan™, blood sampling and collection of demographic data in the two weeks prior to their liver biopsy.

At the Queen Elizabeth Hospital Birmingham a total of 652 patients were invited to take part and 80 patients were consented. The total number of patients invited to take part at the Royal Infirmary of Edinburgh was not recorded and 82 patients were consented. The flow chart for recruitment is shown in Figure 3.3-1.

Healthy volunteers were recruited from colleagues and students at the University of Birmingham and University of Edinburgh. To reduce the chance that volunteers had undiagnosed liver disease they were excluded if they had a history of liver disease, any significant medical co-morbidity, features of the metabolic syndrome, BMI >30kg/m^2 or alcohol consumption in excess of 21 units/week for men or 14 units/week for women.

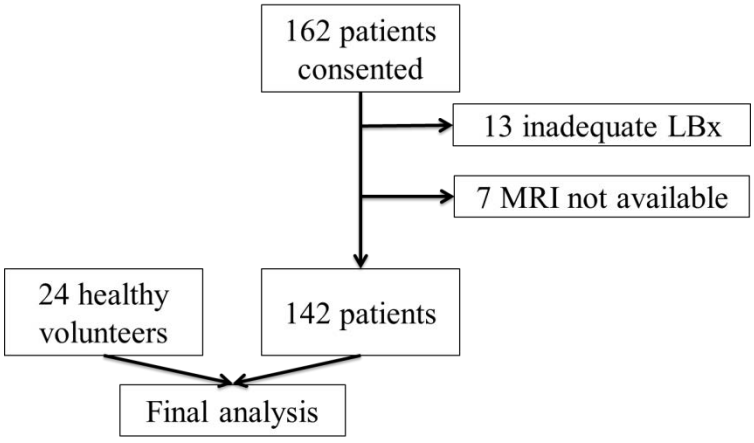


Figure 3.3-1: Flow chart for recruitment.

Study Interventions

MRI scans were conducted as described in Chapter 2. VCTE was measured by trained operators in accordance with manufacture's guidelines. Examinations were regarded as 'possible' if at least 10 valid readings could be recorded and 'reliable' if they contained at least 10 valid readings and had interquartile range (IQR) to median ratio ≤30%.[182] The decision on using the M probe or XL probe was made on the recommendation from the FibroScan™ machine based on the automated skin to liver capsule distance measurement.

Blood samples were analysed for routine markers of liver disease including aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), bilirubin and albumin. Simple blood biomarker panels including AST:ALT ratio, AST/platelet ratio index (APRI) and Fib-4 were calculated according to published formulae.[46, 48, 49] Serum samples were also sent to iQur ltd. (London, UK) for analysis to determine the enhanced liver fibrosis (ELF) score.

Histological assessment

Liver biopsies samples were taken with 16 gauge biopsy needles. Histology was assessed by experienced liver histopathologists blinded to the MRI findings. Biopsies that were less than 15mm in length or that contained less than 11 portal tracts were regarded as inadequate for histological assessment and were therefore excluded.[30, 41] Fibrosis was staged according to Ishak et al.[8] Liver biopsy samples were also assessed for collagen proportionate area by Dr Natasha McDonald at the University of Edinburgh. The method used was according to Calvaruso et al.[26]

As multiple pathologists were involved in the scoring of liver biopsies in this study there is potential for inter-observer variation in fibrosis assessment to reduce the reliability of the results. To assess if this has had a significant effect the interobserver agreement between pathologists was assessed. A random sample of cases was shared between the University of Birmingham and the University of Edinburgh. These cases were rescored and weighted kappa statistics used to assess the agreement between pathologists.

Statistical analysis

Statistical analysis was carried out using IBM SPSS Statistics for Windows version 22 (IBM Corp, Armonk, NY). Variables are summarised with mean and standard deviation (SD) if normally distributed and with median and range if not normally distributed. Correlation between continuous variables was determined with Spearman's correlation coefficient (Rho). Association between normally distributed variables was assessed with t-tests or ANOVA as appropriate. Non-normally distributed variables were assessed with the Mann-Whitney U test, the Kruskal-Wallis test or the Jonckheere-Terpstra test as appropriate. For all tests a p-value <0.05 was taken to indicate statistical significance. Diagnostic performance was compared by calculation of the receiver operating characteristic and determination of the area under the curve (AUROC) with 95% confidence intervals (CI).

3.3.2 Results

Out of 162 patients consented to take part in the study 13 were excluded due to an inadequate liver biopsy specimen. A further 7 patients were excluded due to MRI scans being unavailable for analysis (technical failure/breakdown: 3, withdrawal of consent: 2, unable to tolerate scan: 2), which left a total of 142 patients for analysis. 24 volunteers were recruited and underwent research MRI scans. The acquisition protocol for volunteers was the same as for patients and is described in Chapter 2. All recruited volunteers had complete data for analysis.

Characteristics of the participants and histology

Baseline demographics of the study cohort are outlined in Table 3.3-1. Of the 142 patients included in the final analysis the mean (±SD) biopsy length was 25.3 (±6.2) mm. The post-

biopsy diagnoses are shown in Table 3.3-2 and the distribution of patients in each Ishak stage is shown in Table 3.3-3.

Median (range) collagen proportionate area was 5.64 (0.25-43.19) %. The relationship between CPA and Ishak stage is highly significant by the Jonckheere-Terpstra test ($p<0.0001$) (Figure 3.3-2).

	Patients n=142	**Healthy Volunteers n=24**	**p-value**
Recruited from Birmingham	73 (51.4%)	11 (45.8%)	0.663
Age (years)	52 (18-77)	37 (22-67)	0.001
Male	84 (59.2%)	13 (54.2%)	0.660
Caucasian	125 (88%)	24 (100%)	0.297
Asian	13 (9.2%)	0 (0%)	
Afro-Caribbean	4 (2.8%)	0 (0%)	
BMI (Kg/m^2)	29.8 (±6.7)	22.6 (±3.0)	<0.001
Post-transplant	34 (23.9%)	n/a	n/a
Current alcohol consumption	42 (29.6%)	24 (100%)	<0.001
Median intake*	7.5 (1-140)	7.5 (1-17)	0.995

Data presented as n (%), median (range) or mean (±SD) as applicable.
p-values from Fisher's exact test, Mann-Whitney U test or t-tests as appropriate.
** Of those who consume alcohol*

Table 3.3-1: Baseline demographics of the study population separated into patients and healthy volunteers. Volunteers were significantly younger, had a lower BMI and were more likely to consume alcohol. Median intake of drinkers was not significantly different.

Post-biopsy diagnosis	**n**	**%**
Non-alcoholic fatty liver disease (NAFLD)	50	35.2%
Autoimmune liver disease*	27	19.0%
Chronic hepatitis B and C	18	12.7%
Normal	13	9.2%
Transplant related complications (rejection or vascular problems)	11	7.7%
Alcohol related liver disease	6	4.2%
Drug induced liver injury	4	2.8%
Others including haemochromatosis, α1-antitrypsin deficiency, sarcoid liver disease, ductal plate malformation and nodular regenerative hyperplasia	13	9.2%

** Includes autoimmune hepatitis, primary biliary cholangitis and primary sclerosing cholangitis.*

Table 3.3-2: Post-biopsy diagnosis for the 142 patients included in the final analysis.

Ishak Stage	**n**	**%**
0	29	20.4%
1	26	18.3%
2	22	15.5%
3	32	22.5%
4	9	6.3%
5	7	4.9%
6	17	12.0%

Table 3.3-3: Distribution of Ishak fibrosis stages in the cohort.

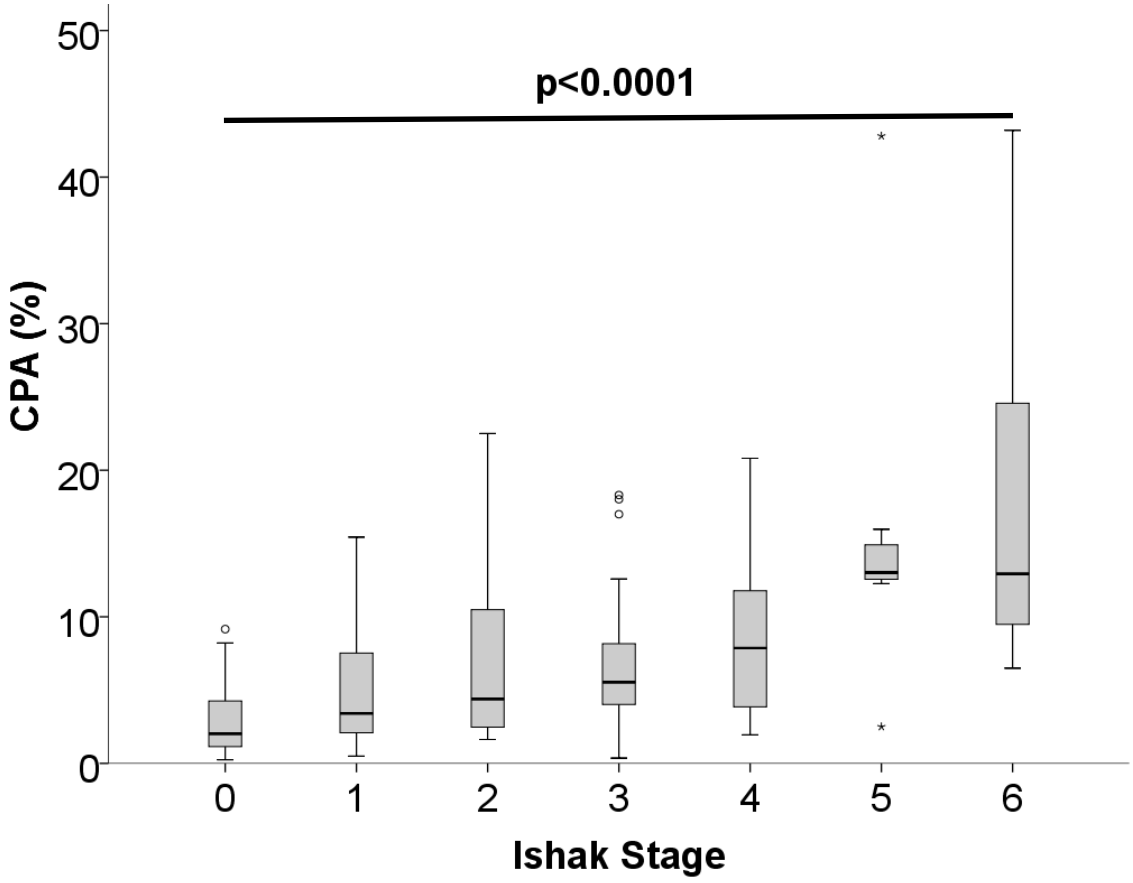


Figure 3.3-2: Box plot showing the relationship between CPA and Ishak stage in the 142 patients included in the final analysis. The overall relationship is highly significant ($p<0.0001$) by the Jonckheere-Terpstra test.

Inter-observer agreement between pathologists

The assessment of fibrosis by Ishak stage showed good agreement between the 4 pathologists assessing biopsies with weighted kappa of 0.66.

Relationship between cT1 and fibrosis

cT1 had a moderate correlation with Ishak stage with Rho = 0.432 ($p<0.001$). Mean cT1 for healthy volunteers (HV) was 768.9 msec and mean cT1 for Ishak stages 0, 1-2, 3-4 and 5-6 were 851.7, 901.9, 953.5 and 995.6 msec respectively. This relationship is shown in Figure

3.3-3 and was statistically significant by the Kruskal-Wallis test with $p<0.001$. Statistically significant pairwise differences are seen between groups HV and 1-2 ($p<0.001$), HV and 3-4 ($p<0.001$), HV and 5-6 ($p<0.001$), 0 and 3-4 ($p=0.002$), 0 and 5-6 ($p<0.001$) and 1-2 and 5-6 ($p=0.006$). cT1 had a weak but statistically significant positive correlation with CPA with Spearman's Rho of 0.333 ($p<0.001$).

Using the cT1 cut-off value of 822 msec that is proposed by Perspectum Diagnostics as the upper limit of normal (see section 2.11.1) multiparametric MRI had sensitivity of 0.82, specificity of 0.45, positive predictive value (PPV) of 0.85 and negative predictive value (NPV) of 0.39 for the detection of liver fibrosis (≥ Ishak stage 1). Using the upper limit of normal defined in section 2.11.2 (863 msec) multiparametric MRI had sensitivity of 0.78, specificity of 0.59, PPV of 0.88 and NPV of 0.40 for the detection of liver fibrosis (≥ Ishak stage 1).

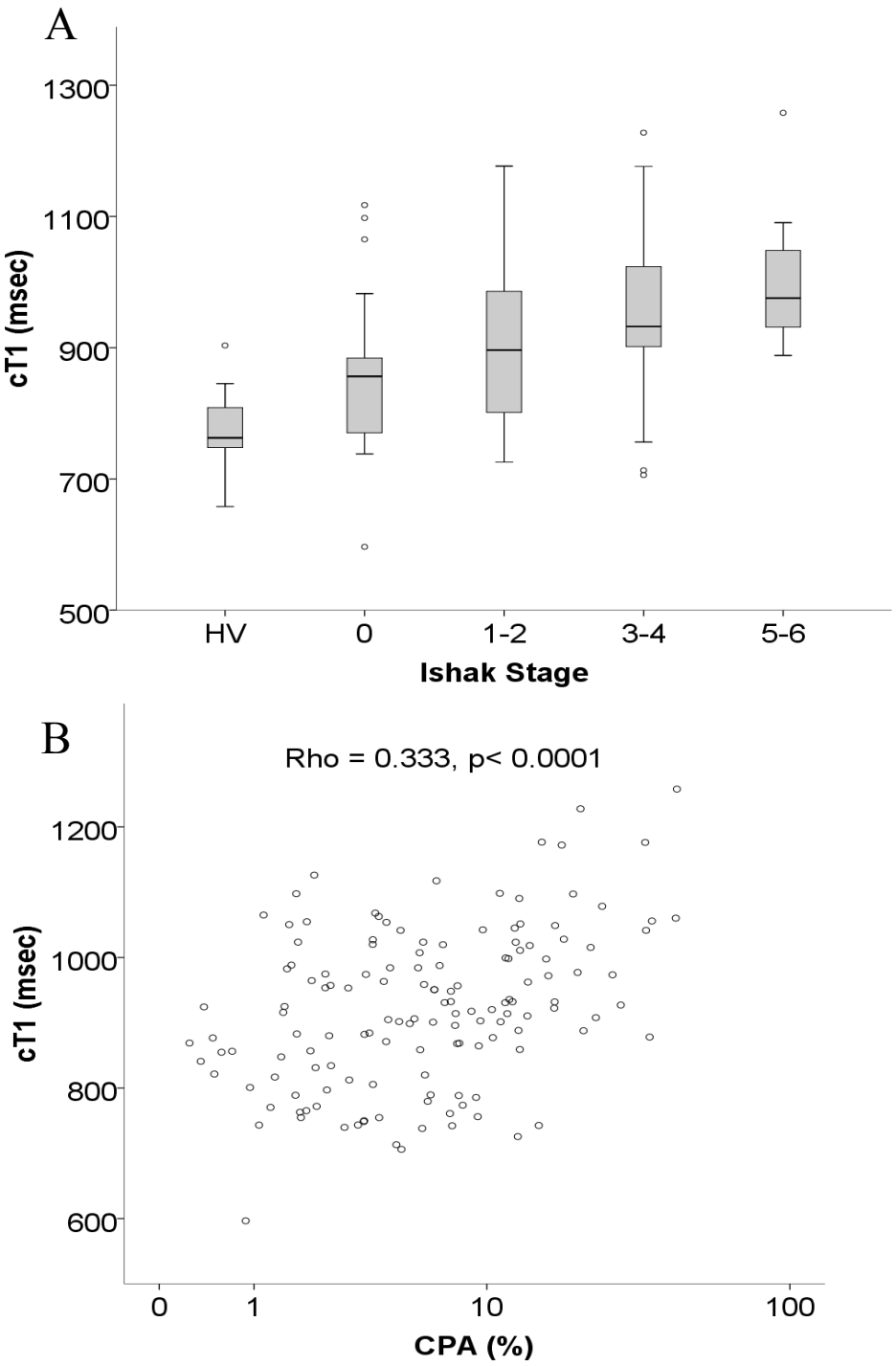


Figure 3.3-3: A) Box plot showing the relationship between cT1 and Ishak stage for patients undergoing liver biopsy. HV did not undergo liver biopsy and are presented as a separate group. $p<0.001$ by the Kruskal-Wallis test. See text for significant intergroup differences by post hoc tests. B) Scatter plot showing the correlation between cT1 and CPA for patients only. Spearman's Rho = 0.333, $p<0.001$.

Relationship between established biomarkers and fibrosis

Volunteers did not undergo blood sampling or VCTE so established biomarkers of fibrosis are available for patients only. VCTE was possible in 136/142 (95.8%) patients and reliable by Boursier's criteria[182] in 130/142 (91.5%) patients. Only reliable VCTE was included in further analysis. As shown in Figure 3.3-4, all assessed biomarkers had a statistically significant association with Ishak stage by the Jonckheere-Terpstra test. AST:ALT ratio ($p=0.001$), APRI ($p=0.006$), FIB-4 ($p<0.001$), hyaluronic acid ($p<0.001$), ELF test ($p<0.001$) and liver stiffness from reliable VCTE examinations ($p<0.001$).

As shown in Figure 3.3-5, CPA showed a moderately strong positive correlation with ELF (Rho=0.401, $p<0.001$) and VCTE (Rho=0.432, $p<0.001$). There was a weak but statistically significant positive correlation with AST:ALT ratio (Rho=0.272 , $p=0.001$), FIB-4 (Rho=0.269, $p=0.001$) and hyaluronic acid (Rho=0.390, $p<0.001$). There was no statistically significant correlation with APRI (Rho=0.146, $p=0.085$).

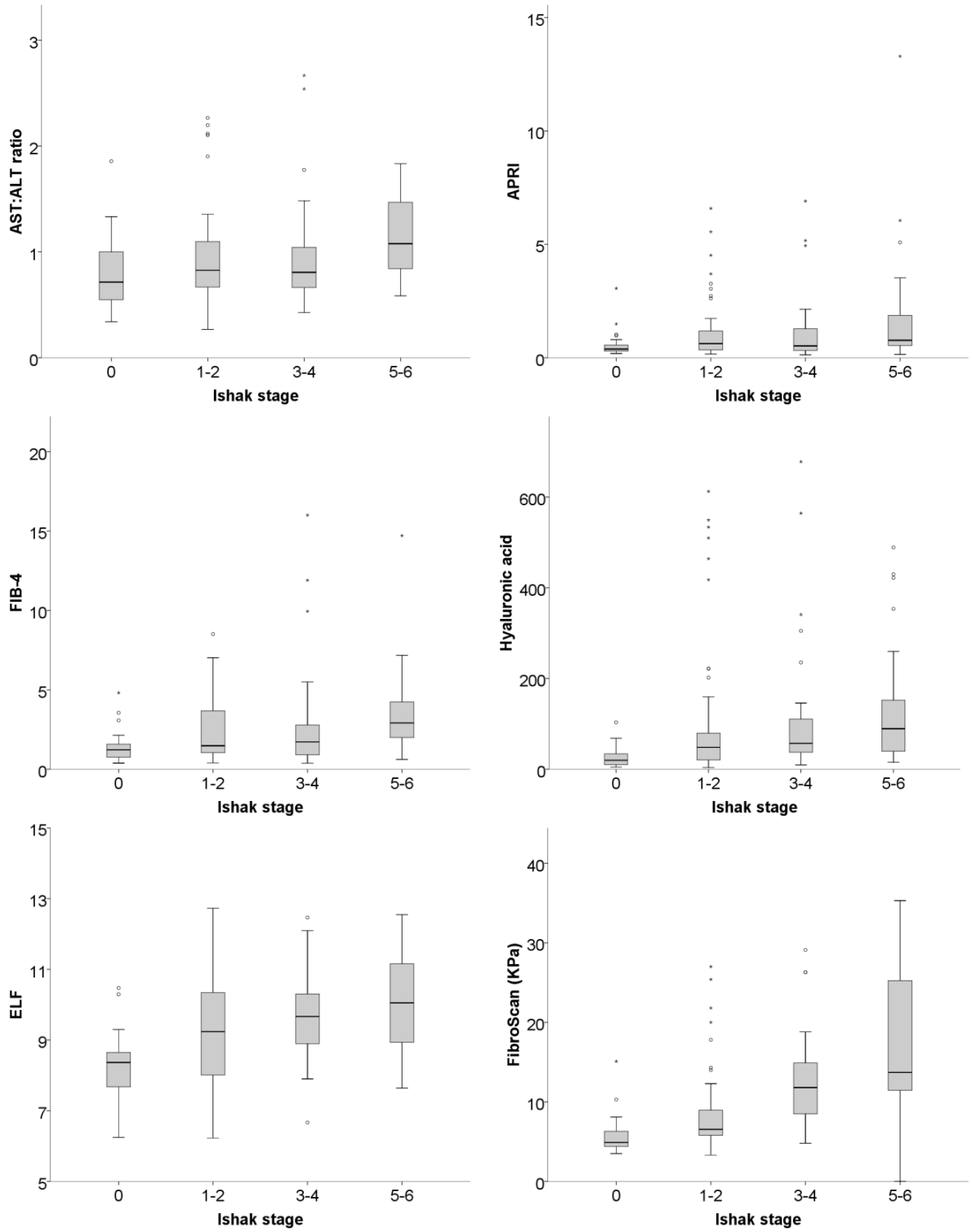


Figure 3.3-4: Box plots showing the relationship between Ishak stage and non-invasive biomarkers of fibrosis. All relationships significant by the Jonckheere-Terpstra test (n=130 for FibroScan.

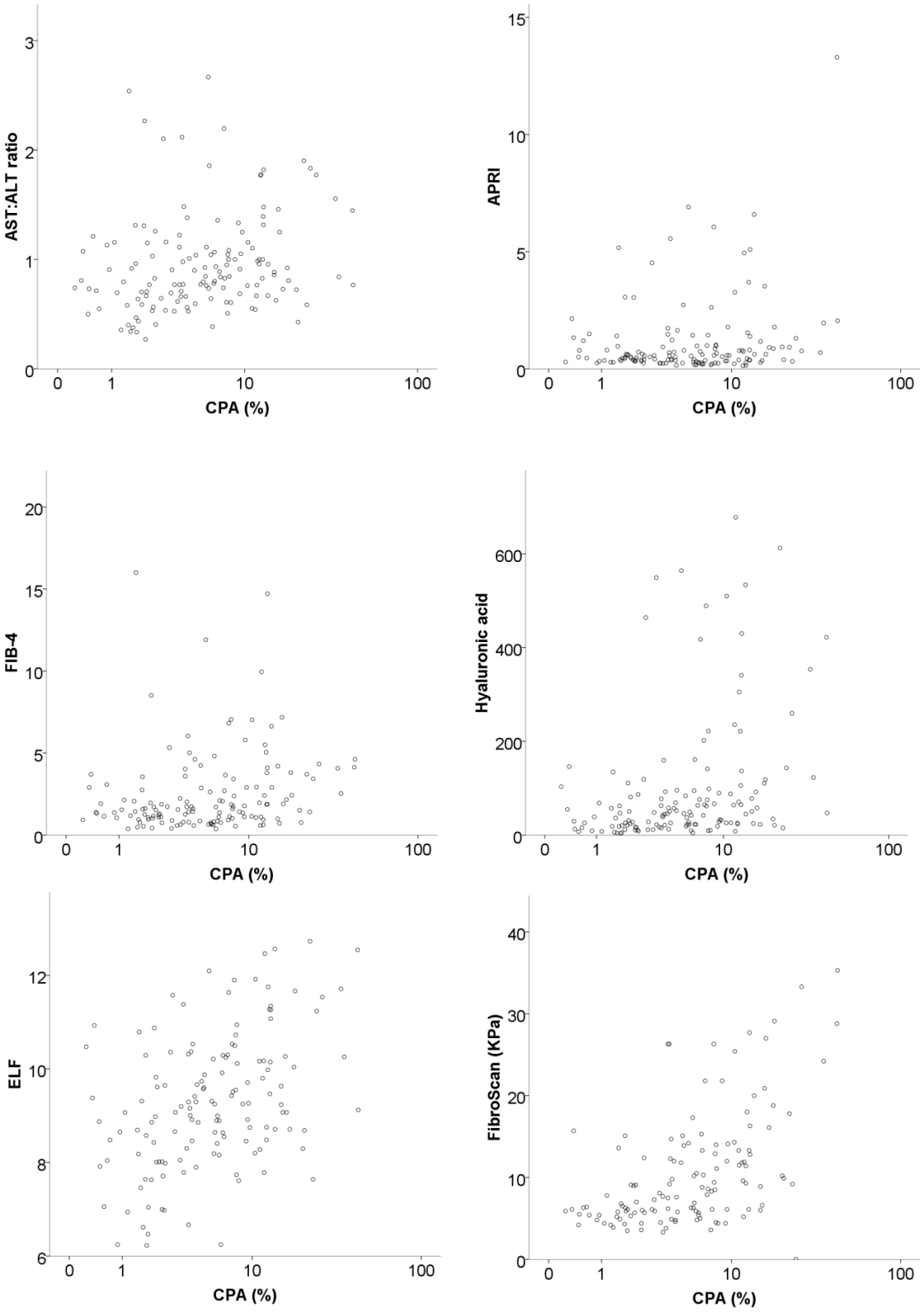


Figure 3.3-5: Scatter plots showing the correlation between CPA and non-invasive biomarkers of fibrosis.

Comparative performance of cT1 compared to other non-invasive biomarkers for the staging of fibrosis

For the detection of any fibrosis (Ishak ≥1), advanced fibrosis (Ishak ≥4) and cirrhosis (Ishak ≥5) the area under the receiver operating characteristic (AUROC) with 95% confidence interval is shown in Table 3.3-4. The cT1 values to maximise the Youden index are also given along with the sensitivity and specificity at this cut off.

	Any fibrosis (Ishak ≥1)	**Advanced fibrosis (Ishak ≥4)**	**Cirrhosis (Ishak ≥5)**
cT1	0.71 (0.59-0.82)	0.73 (0.64-0.82)	0.71 (0.61-0.81)
Cut-off value to maximise Youden index	888.0 msec	888.0 msec	888.0 msec
Sensitivity	0.72	0.97	1.00
Specificity	0.79	0.50	0.47
Liver stiffness *	0.83 (0.75-0.91)	0.83 (0.75-0.91)	0.81 (0.71-0.92)
ELF	0.79 (0.71-0.88)	0.66 (0.55-0.77)	0.66 (0.54-0.78)
Hyaluronic acid	0.78 (0.70-0.87)	0.69 (0.58-0.79)	0.69 (0.58-0.80)
FIB-4	0.68 (0.58-0.79)	0.67 (0.56-0.78)	0.73 (0.62-0.84)
APRI	0.66 (0.55-0.76)	0.60 (0.49-0.72)	0.65 (0.53-0.78)
AST:ALT ratio	0.63 (0.51-0.75)	0.60 (0.48-0.72)	0.68 (0.57-0.80)

** n=130 for liver stiffness (Reliable FibroScan examinations only) n=142 for all other tests*

Table 3.3-4: AUROC (95% CI) values for the identification of any fibrosis, advanced fibrosis and cirrhosis.

Influence of inflammation of cT1

During data analysis it was noted that there were several patients with an elevated cT1 without significant fibrosis and it was also noted that cT1 is significantly higher in patients with non-alcoholic fatty liver disease (NAFLD) than in other aetiologies.

It is proposed that this increase in cT1 is caused by hepatic inflammation. This is supported by previous work by Hoad et al, which suggested that inflammation leads to a significant increase in hepatic T1 independent of fibrosis stage.[164] The correction algorithm to calculate cT1 from T1 corrects for iron overload but not inflammation.

We sought to assess how hepatic inflammation influences cT1 in this cohort. Hepatic inflammation is difficult to quantify in a mixed cohort such as this due to the absence of a single validated histological score for inflammation across aetiologies. Different aetiologies have different patterns of liver injury and therefore a system that describes the histological appearance of inflammation in one condition is not valid for other conditions. For this reason a complex system to describe inflammation was inappropriate. To define the presence of inflammation, histology reports were reviewed by investigators blinded to the MRI findings. Biopsy samples were divided into those with and those without inflammation by consensus agreement. For patients with and without inflammation on liver biopsy, mean cT1 was 942.1 and 884.9 msec respectively. The difference in cT1 was statistically significant by the Mann-Whitney U test ($p=0.004$) and is shown in Figure 3.3-6.

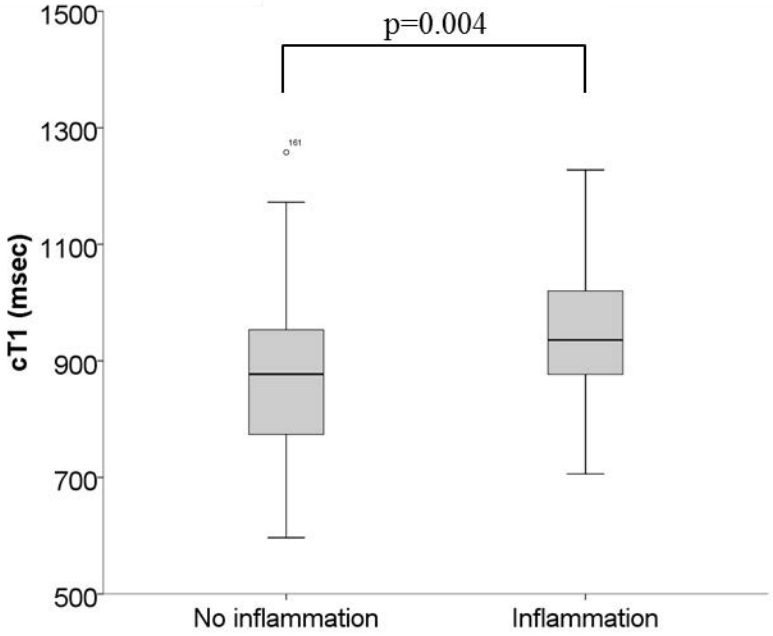


Figure 3.3-6: Box plot showing the statistically significant difference in cT1 between patients with no inflammation (n=93) and those with inflammation (n=49) on liver biopsy.

As demonstrated in Figure 3.3-7. This difference in cT1 is not seen in patients with fibrosis and approaches statistical significance in those without fibrosis suggesting that there may be a ceiling effect. cT1 appears to reflect ‘fibro-inflammation’ within the liver.

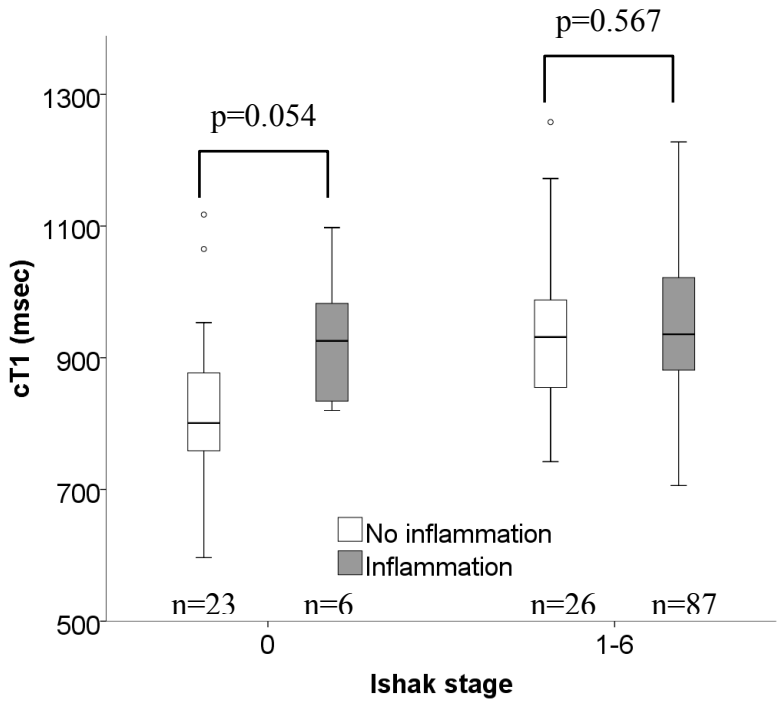


Figure 3.3-7: Box plot showing the difference in cT1 between patients stratified by fibrosis stage and by the presence of inflammation.

Exclusion of 'fibro-inflammatory' liver disease

23/142 patients had no fibrosis and no more than minimal inflammation on biopsy. Mean (±SD) cT1 was statistically significantly lower in this group compared to those with either fibrosis, inflammation or both (831 (±111) msec vs 940 (±113) msec, $p<0.001$) (Figure 3.3-8). cT1 had AUROC (95% CI) for the differentiation of these groups of 0.768 (0.658-0.879) A cut off value of 822 msec is proposed as the upper limit of normal by Perspectum diagnostics.[163] Using this value gave cT1 a PPV and NPV of 0.90 and 0.36 respectively for the exclusion of fibro-inflammatory liver disease.

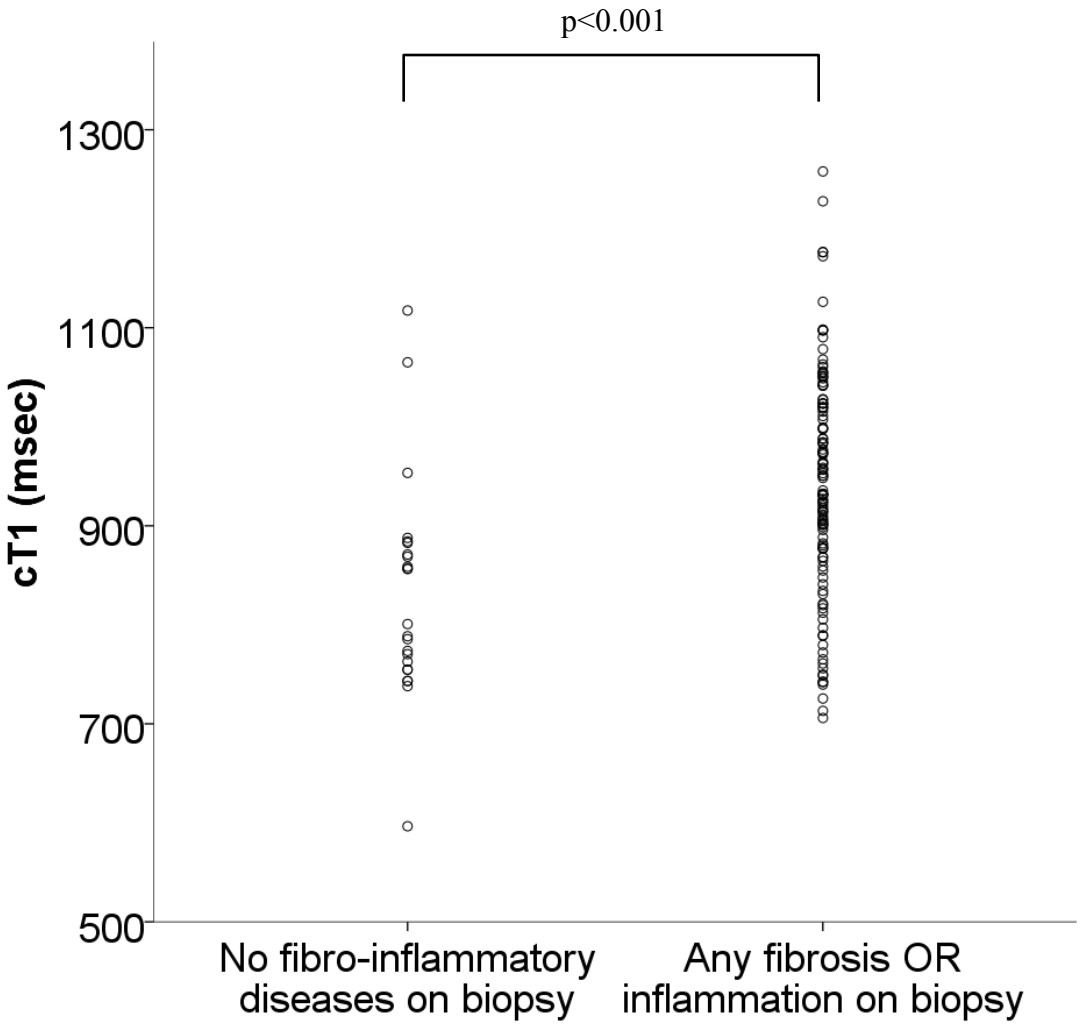


Figure 3.3-8: Distribution of cT1 values in patients with either fibrosis or inflammation compared to those with neither. P-value from independent samples t-test.

Confounding effect of aetiology

If the 34 liver transplant recipients are excluded from the analysis, the correlation between cT1 and Ishak stage for the remaining 108 patients showed a stronger correlation (Rho=0.53, p<0.001) than in the whole cohort (Rho = 0.432, p<0.001). The same stronger correlation was seen when fibrosis was assessed with CPA. With transplant patients were excluded Rho=0.41 (p<0.001) and in the whole cohort Rho=0.333 (p<0.001). For the 50 patients with NAFLD,

the correlations between cT1 and both Ishak stage and CPA are non-significant with Rho=0.23 (p=0.115) and Rho=0.18 (p=0.209) respectively.

As demonstrated in Figure 3.3-9, there is a statistically significant elevation of cT1 in NAFLD compared to other aetiologies. As seen in Figure 3.3-10, this difference in cT1 is not explained by the distribution of fibrosis or inflammation.

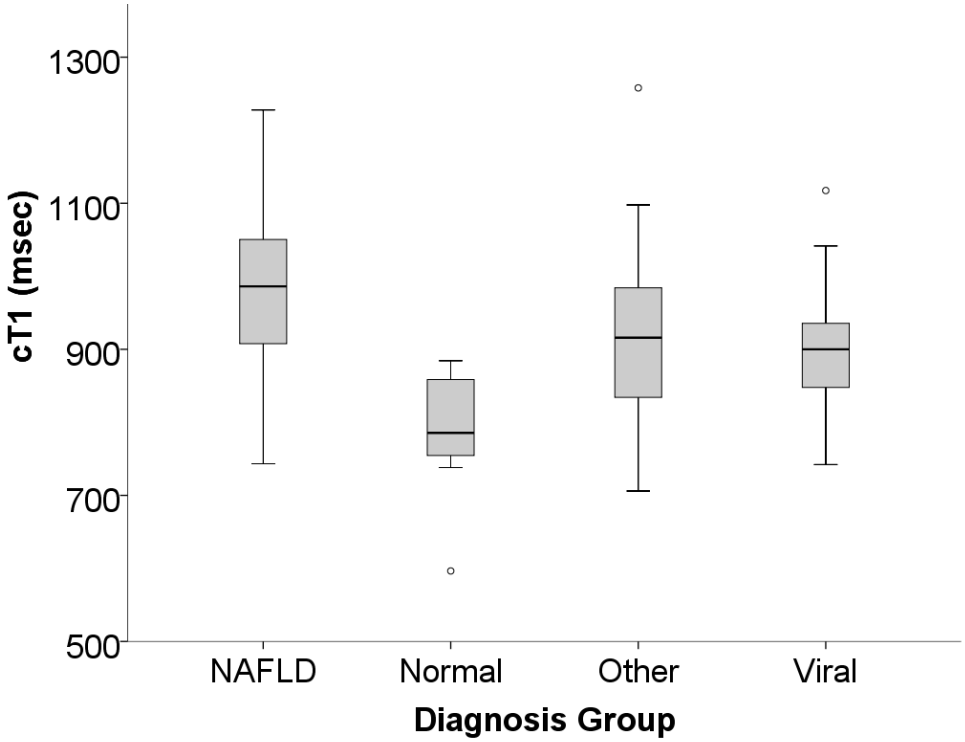


Figure 3.3-9: Boxplot demonstrating that cT1 varies by diagnosis. p<0.001 by the Kruskal-Wallis test. Post hoc tests show significant inter-group differences between NAFLD and all other groups and between normal biopsies and 'other'. There is no statistically significant difference between those in the viral and other groups.

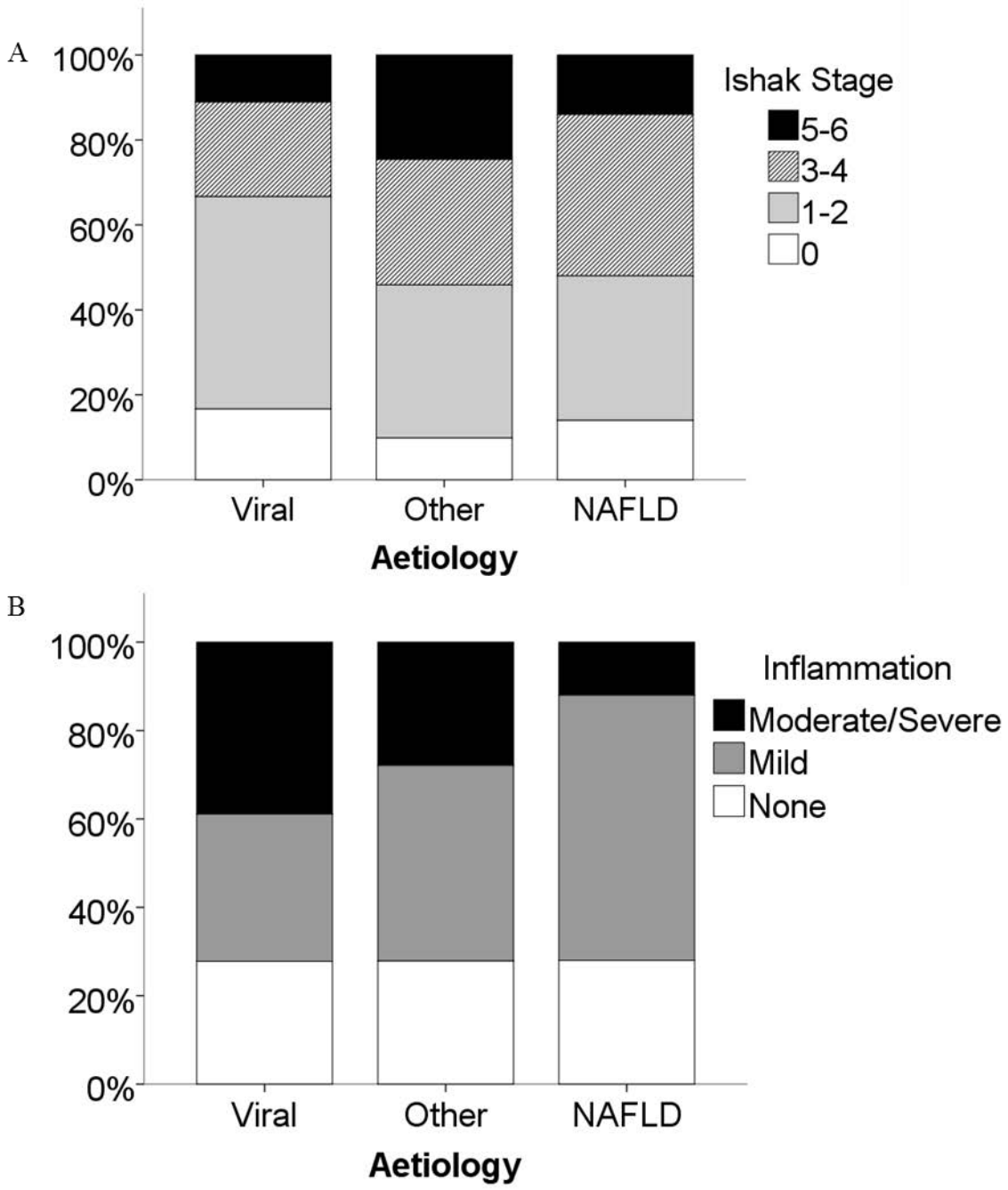


Figure 3.3-10: Distribution of fibrosis and inflammation between different aetiologies. A) The distribution of fibrosis stages is not significantly different between aetiologies (p=0.548 by Fisher's exact test). B) The distribution of histological inflammation is not significantly different between aetiologies (p=0.098 by Fisher's exact test).

3.3.3 Discussion

Given the known limitations of currently available methods of fibrosis assessment there is clear need to develop effective and accurate non-invasive biomarkers of hepatic fibrosis. Such a method could have utility in clinical practice and as a useful surrogate endpoint in research. This study is the first independent validation of cT1 as a biomarker of hepatic fibrosis. Presented are data evaluating the performance of cT1 in a large and well characterised cohort of patients and HVs.

Demographic data show that patients are older, have higher BMI and are less likely to consume alcohol than HV. The lower BMI in HV is expected due to exclusion of potential HV with BMI above the normal range and the lower age in HV reflects the fact that the majority of HV were students who tend to be younger. In this study the HV are treated as a separate group from patients without fibrosis on biopsy. Thus the demographic differences do not introduce confounding.

The biopsies included in this study are of a good size, which has been shown to improve the reliability and repeatability of the histological assessment of fibrosis.[37] This suggests that, notwithstanding the intrinsic limitations of liver biopsy histology, it is likely to have provided a reliable reference standard against which to compare multiparametric MRI. Further confidence in the reliability of biopsy as a reference standard in this study is provided by the good agreement between the pathologists involved in this study. Weighted kappa statistics show good agreement and these kappa values are in line with published data.[183-185] The relationship between Ishak stage and CPA seen in this study is comparable to previous

studies.[26, 186] CPA technique is not well standardised and this agreement with published data provides reassurance that our CPA technique is reliable.

Within the patient cohort all stages of fibrosis are represented. This increases the applicability of the results to a wide range of patients. Even though all fibrosis stages are represented, there are fewer patients with cirrhosis than earlier stage fibrosis. This is likely to be due to the fact that cirrhosis is more readily detectable by non-invasive tests and therefore standard of care biopsy is less likely to be indicated in those with cirrhosis. Within the study cohort there is a relative excess of autoimmune liver disease and a paucity of alcohol related liver disease. This demonstrates that the patients referred for liver biopsy are not representative of liver patients as a whole. This should be born in mind when applying these results to other populations. It is possible that the findings of this study would not be applicable to the wider population.

Data show a stepwise increase in cT1 with increasing Ishak stage and a fairly weak but highly statistically significant correlation between cT1 and CPA. Post hoc tests show statistically significant differences in cT1 between HV and mild fibrosis, moderate fibrosis and cirrhosis. Patients without fibrosis on biopsy had statistically significant differences in cT1 when compared to those with advanced fibrosis and cirrhosis. There were no statistically significant differences between any two adjacent groups. Of note there is no statistically significant difference between advanced fibrosis and cirrhosis or between no fibrosis and mild fibrosis.

Other evaluated non-invasive tests have performed broadly in keeping with published data however the correlations between CPA and non-invasive tests were generally less strong than published studies. No statistically significant correlation between CPA and APRI was found.

When comparing non-invasive tests for the detection of any fibrosis (Ishak >0), advanced fibrosis (Ishak ≥4) or cirrhosis (Ishak ≥5) cT1 had AUROC comparable to all other evaluated tests. Only VCTE outperformed multiparametric MRI across all stages of fibrosis.

Although these data show that the performance of multiparametric MRI is comparable to other non-invasive markers of liver fibrosis the strength of the relationship between cT1 and hepatic fibrosis is less strong in this cohort than in previous published data from Banerjee et al.[163] The Banerjee study reports Rho=0.68 ($p<0.0001$) for the correlation between cT1 and Ishak stage and Rho=0.54 ($p<0.0001$) for the correlation between cT1 and CPA. This is in contrast to Rho=0.43 ($p<0.001$) for the comparison with Ishak stage and Rho=0.33 ($p<0.001$) for the comparison with CPA in this study. Banerjee et al also showed statistically significant differences between all fibrosis groups except between Ishak 1-2 and Ishak 3-4 as shown in Figure 2.6-1. This has not been replicated in this work where there was no statistically significant difference between no fibrosis vs early stage fibrosis nor moderate fibrosis vs cirrhosis.

This weaker relationship is reflected in reduced diagnostic performance. For the detection of any fibrosis AUROC (95% CI) for multiparametric MRI was 0.71 (0.59-0.82) in this study, which is notably lower than 0.94 (0.90–0.99) in the Banerjee study. When making this comparison it should be noted that, in the Banerjee paper, HV and patients without fibrosis on biopsy are analysed together whereas here they are analysed as two separate groups. In these data there is a non-statistically significant trend towards lower cT1 in HV than those with Ishak 0 fibrosis on biopsy. It is possible that this difference in the analysis has contributed to the lower AUROC in this work.

The difference in analysis cannot completely explain the weaker correlation between cT1 and fibrosis seen in this work and it remains unclear why multiparametric MRI should have performed less well in this study than in previous work. Other non-invasive markers of liver fibrosis assessed in this study have performed broadly in line with published data. This would suggest that if patient factors are confounding the ability of cT1 to measure fibrosis that it is an issue specific to multiparametric MRI.

The reduced diagnostic performance of multiparametric MRI in this study compared to previous work may be driven partly by inflammation. Figure 3.3-6 shows that, in line with published work by Hoad et al cT1 is significantly different between those with and without inflammation on biopsy.[164] As seen in Figure 3.3-7, this effect of inflammation on cT1 is most marked in patients without fibrosis suggesting that a raised cT1 may be due to either fibrosis or inflammation or both. This is a clear confounding factor when using cT1 to measure hepatic fibrosis.

It may be more clinically useful to consider cT1 a measure of 'fibro-inflammation' rather than any single process. In clinical practice a tool that could reliably identify patients without significant fibrosis or inflammation could be of value as it may allow these patients to be quickly and safely reassured and discharged from follow-up. AUROC for the identification of participants with fibro-inflammation was higher than for the detection of fibrosis alone. However, to be a useful test to exclude fibro-inflammation and avoid further testing, multiparametric MRI would require a high NPV. Using the cT1 cut off value of 822 msec[163] the NPV for the exclusion of fibro-inflammatory liver disease is very low at 0.36. The PPV is high at 0.90 however, the accuracy of both these assessments are limited by the recruitment

strategy. Participants were recruited from patients having a standard of care biopsy and therefore there was already a suspicion of either fibrosis or inflammation. This influences the pre-test probability and therefore influences NPV and PPV.

The performance of multiparametric MRI is not consistent across aetiologies. When analysing only patients with NAFLD (the single biggest group in this cohort) cT1 shows no significant relationship with either Ishak stage or CPA. Excluding liver transplant recipients from the analysis improves the correlation of cT1 with Ishak stage but the performance remains less good than in the Banerjee study. The mix of aetiologies is another possible reason for the poorer performance of multiparametric MRI in this study. Although the cohort used for this study has broadly similar demographics to that used previously, the mix of diagnoses is notably different. Autoimmune liver disease forms the second largest group in the current study and it is an uncommon diagnosis in the Banerjee study's cohort. There are also no liver transplant recipients in the Banerjee cohort and 23.9% of patients in the current study were liver transplant recipients.

As well as the diagnostic performace of multiparametric MRI varying with aetiology there is a significant increase in cT1 in patients with NAFLD compared to other aetiologies. As seen in Figure 3.3-10 this difference cannot be explained by either the severity of fibrosis or the presence of inflammation. This suggests that factors other than fibrosis and inflammation also contribute to the measured cT1.

One obvious difference between patients in the NAFLD group and those in the other groups is the hepatic fat content. By histological assessment of steatosis there is significantly more fat

in the liver of patients with NAFLD than without ($p<0.001$ by Fisher's exact test). Published evidence using MRI simulation and phantom experiments suggests that T1 is overestimated in the presence of steatosis when, as in this study, it is measured by the ShMoLLI sequence.[187] Further work by Perspectum Diagnostics suggests that this effect is small and not clinically relevant (Personal communication, Dr R. Banerjee, CEO Perspectum Diagnostics). The difference in performance and the difference in absolute cT1 values between aetiologies remain unexplained and warrants further study.

Another factor that may be important when assessing cT1's ability to assess hepatic fibrosis is the use of liver biopsy as the reference standard and in particular the sampling error inherent in assessment with liver biopsy.[31] It is unlikely that the ROI used to make the cT1 measurement would be placed in the same place as the site of biopsy and so variations in fibrosis across the liver may confound assessment with multiparametric MRI. This possible source of confounding is considered further in section 3.4.

3.4 Evaluation of cT1 for the staging of fibrosis using multiple colocalised histology specimens

3.4.1 Aims

As discussed above, a factor that may reduce the correlation between cT1 and liver biopsy histology is the limitations of liver biopsy histology as a reference standard. One key aspect of liver biopsy histology that limits its accuracy is sampling error. In the study described in section 3.3 it is likely that the ROI for cT1 measurement will have been placed in a different part of the liver to the site of the biopsy. With the variation in fibrosis that can be seen in chronic liver disease, the severity of fibrosis within the cT1 measurement ROI could be different to the severity of fibrosis at the site of the biopsy, thus confounding the comparison of cT1 and histology.

The goal of this study was threefold:

- To devise a method to allow cT1 measurement and histological assessment to be colocalised to eliminate the effects of sampling error and fibrosis variation.
- To determine the variation of fibrosis across the liver and evaluate the potential for T1 mapping to demonstrate this variation.
- To assess if multiparametric MRI assessment of the whole liver is superior to a single ROI

3.4.2 Methods

This study conforms to the research ethics guidelines of the 1975 declaration of Helsinki. It has received approval from the national research ethics service (14/WM /0010) and local research and development office.

Patient recruitment

The only practical way to get access to multiple samples from the same liver was from hepatectomy specimens. The Centre for Liver Research has a close relationship with the liver transplant programme at University Hospitals Birmingham NHS Foundation Trust and hepatectomy specimens taken at the time of liver transplantation (the explant) are already used within the Centre for Liver Research for tissue collection.

Patients awaiting liver transplant at University Hospitals Birmingham NHS Foundation Trust were identified from the liver transplant waiting list. To minimise the time between the scan and transplant, patients who could be expected to have a short waiting time were selected. This included patients with blood group A, B or AB, all patients (of any blood group) with a UKELD ≥ 60 and patients (of any blood group) being transplanted for hepatocellular carcinoma (HCC). Specific exclusion criteria were the presence of polycystic liver disease and any contraindication to MRI. The flow chart for recruitment is shown in Figure 3.4-1. For this pilot work a target of 10 patients was set.

Patients meeting inclusion and without exclusion criteria were invited for a single research visit where they underwent research MRI, FibroScan, blood sampling and the collection of

demographic and clinical information. When recruited patients underwent transplantation, their explant was collected from the operating theatre and processed in the Centre for Liver Research. Once the target of 10 patients with paired MRI and histology data was reached no further patients were enrolled and no further explants were processed for those who had already been enrolled.

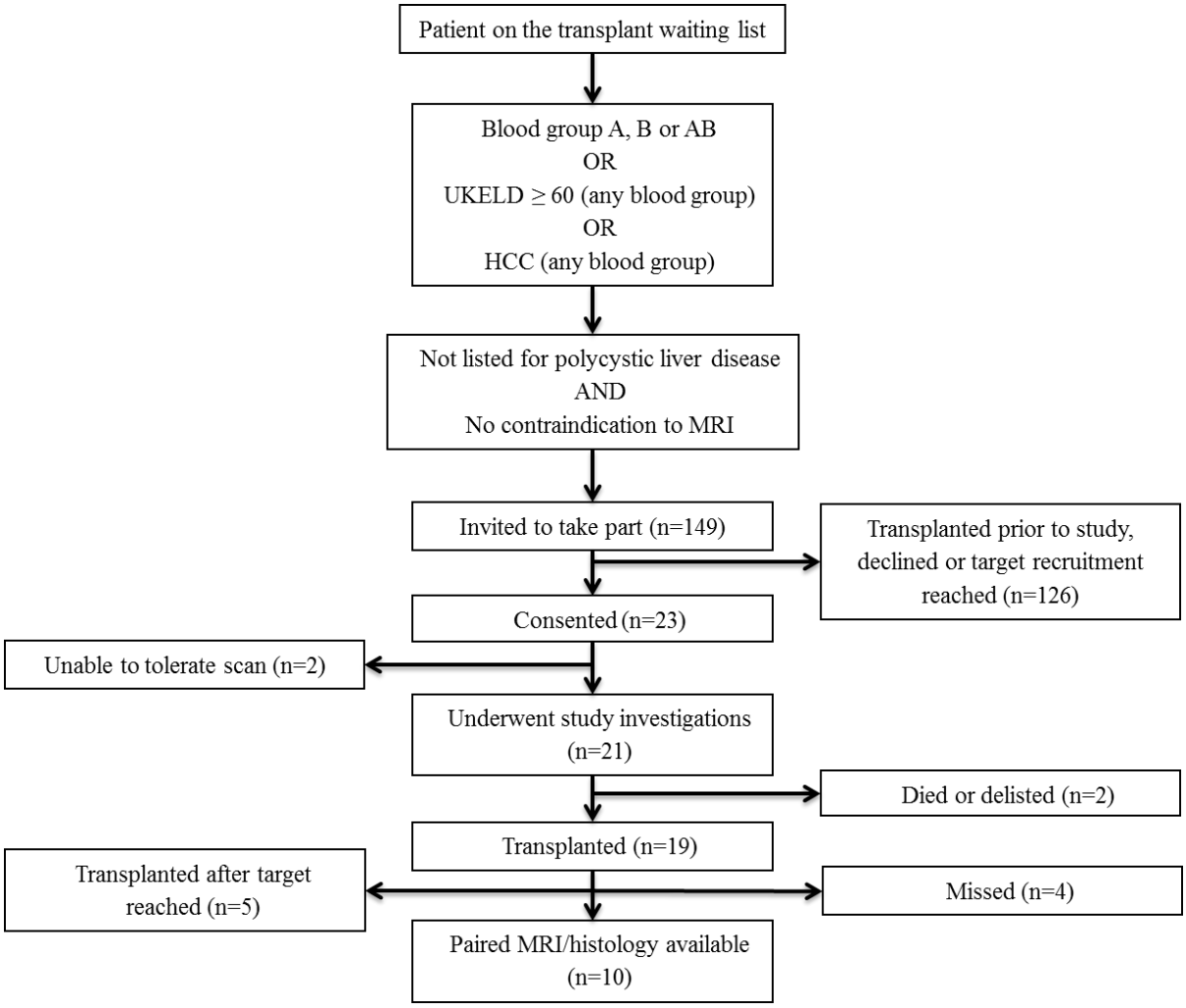


Figure 3.4-1: Flow chart for recruitment to this study indicating reasons for drop out from the study

Study Interventions

MRI scans were conducted as described in Chapter 2 with the exception that the T1 maps for this study were acquired in the sagittal plane to facilitate colocalisation of histology and MRI data. FibroScan examinations were performed by trained operators in accordance with manufacture's guidelines. Examinations were regarded as 'possible' if at least 10 valid readings could be recorded and 'reliable' if they contained at least 10 valid readings and had interquartile range (IQR) to median ratio ≤30%.[182] The decision on using the M probe or XL probe was made on the recommendation from the FibroScan machine based on the automated skin to liver capsule distance measurement.

Blood samples were analysed for routine markers of liver disease including aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), bilirubin and albumin. Simple blood biomarker panels including AST:ALT ratio, AST/platelet ratio index (APRI) and Fib-4 were calculated according to published formulae.[46, 48, 49] Serum samples were also sent to iQur ltd. (London, UK) for analysis to determine the enhanced liver fibrosis (ELF) score.

Histological specimen preparation

Explanted livers were laid on their posterior surface and slices cut in the sagittal plane. 2 slices were taken from the right lobe and 2 from the left. Each slice was 1cm thick and the position of each of these slices within the liver was measured from the right lateral edge of the liver. This is shown schematically in Figure 3.4-2. Slices were labelled A to D from right to left.

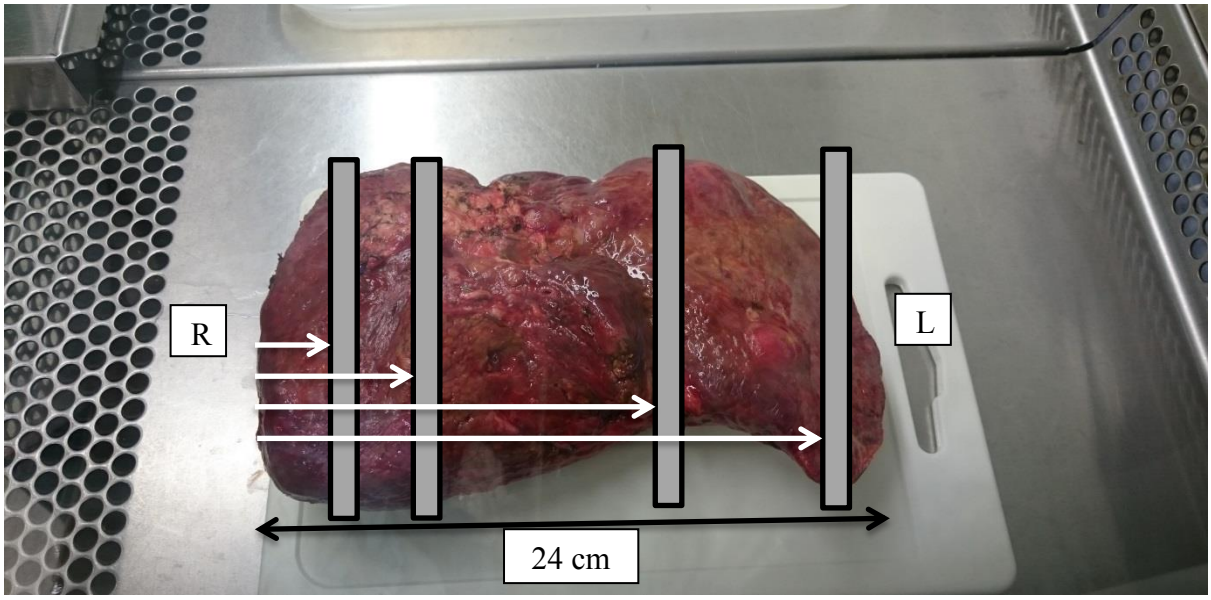


Figure 3.4-2: Anterior view of explant from MURAL-B-006 demonstrating positioning of slices. Grey bars: position of the slices, black arrow: medio-lateral size of explant, white arrows: measurement of slice position.

Four samples of approximately 2x2cm were then taken from each slice as demonstrated in Figure 3.4-3. These were labelled 1 to 4 from cranial to caudal. Measurements were made from the superior and anterior surface of the liver so that the position of the sample could be matched to the corresponding MRI slice. Samples were placed in histology cassettes, formalin fixed and paraffin embedded.

Figure 3.4-3: Photograph of slice B from MURAL-B-006 showing the position of samples.

Cut sections were stained with Haematoxin & Eosin, Perl's stain and van Gieson stain by Dr Gary Reynolds (Senior Research Fellow, Centre for Liver Research, University of Birmingham). Semi-quantitative assessment of steatosis, siderosis and fibrosis of each sample was performed by Dr Owen Cain (Pathology academic clinical fellow, University Hospitals Birmingham NHSFT). Steatosis, siderosis and fibrosis were scored using the systems described by Brunt et al,[188] Scheuer et al[189] and Ishak et al[8] respectively. Sections were also stained with Picro Sirius Red stain and photographed with a Leica Aperio slide scanner (Leica, Nussloch, Germany). Digital imagaes were analysed with the Definiens Tissue Studio (Definiens, Munich, Germany) to calculate the collagen proportionate area (CPA). Representative images of liver tissue stained with Picro Sirius Red at varying magnification are shown in Figure 3.4-4 demonstrating the staining of collagen fibres within pathological hepatic fibrosis.

Due to the large area of the histology specimens in this study (approximately 400mm^2) an unedited technique was used for CPA calculation. An unedited technique is where structural collagen present in normal liver, such as supporting portal tracts and blood vessel walls is not edited from the digital image prior to CPA calculation. It has been shown by Standish et al in 2006 that an unedited technique is appropriate for large samples, such as the ones used in this study.[190]

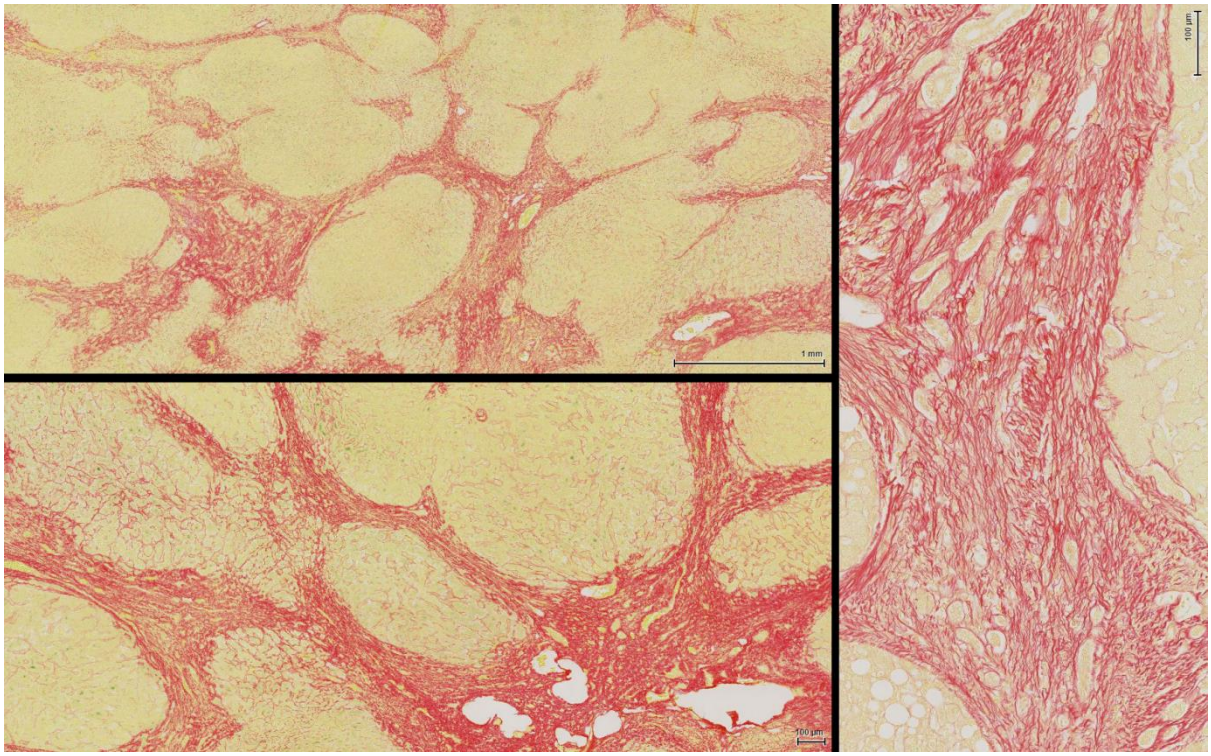


Figure 3.4-4. Liver tissue stained with Picro Sirius Red at varying magnifications.

MRI data analysis

MRI data were analysed using Liver*Multiscan*^{TM} software by a single operator at Perspectum Diagnostics ltd blinded to the participant characteristics and histology results. A single ROI was placed on the T1 map with the position of the ROI chosen to match the location of the corresponding histology sample. The method of colocalisation of MRI ROI and histology

sample is described below. cT1 was calculated from the measured T1 value by the Liver*Multiscan*™ software using a T2* measurement taken from the T2* map taken through the hilum of the liver in a transverse plane.

In addition to single ROI analysis cT1 values for the whole liver were calculated. To perform this analysis the Liver*Multiscan*™ software calculates the cT1 value for each pixel in the liver and the mean, median, mode and coefficient of variation around the mean of these values is calculated. Liver*Multiscan*™ was used to identify blood vessels within the liver with a threshold technique based on T1 values. Once identified, vessels are digitally subtracted from the image and the mode cT1 of the liver parenchyma without the vessels is calculated. An example of vessel subtracted liver slice is shown in Figure 3.4-5.

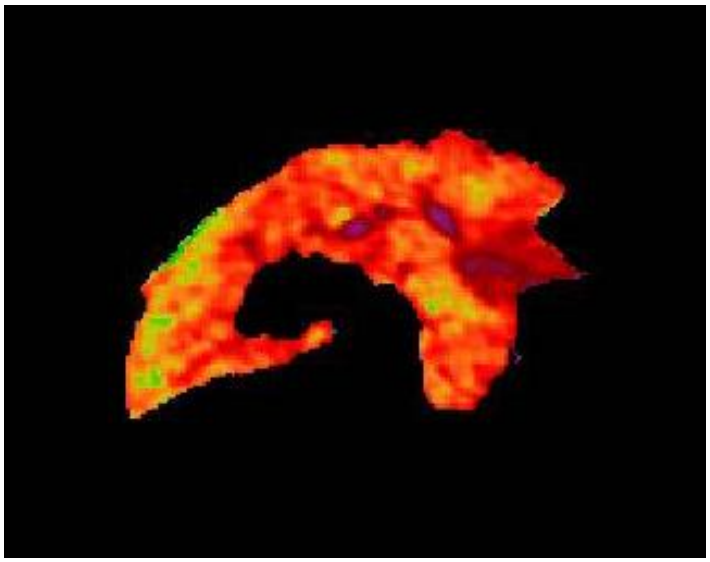


Figure 3.4-5: cT1 map of a single liver slice with the blood vessels removed with a threshold technique. Image courtesy of Perspectum Diagnostics Ltd.

Colocalisation of histology sample and MRI ROI

The colocalisation of ROIs for cT1 calculation and the location of the histology samples has proven to be very challenging. Compared to an explant lying flat on its posterior surface (as shown Figure 3.4-2) a liver in situ is rotated around both its cranio-caudal and medio-lateral axes. Rotation around the medio-lateral axis does not inhibit colocalisation as it does not alter the plane of the slice taken through the liver. However, rotation around the cranio-caudal axis causes the plane of the slice to differ between the MRI and the histology. This rotation is shown in

Figure 3.4-6.

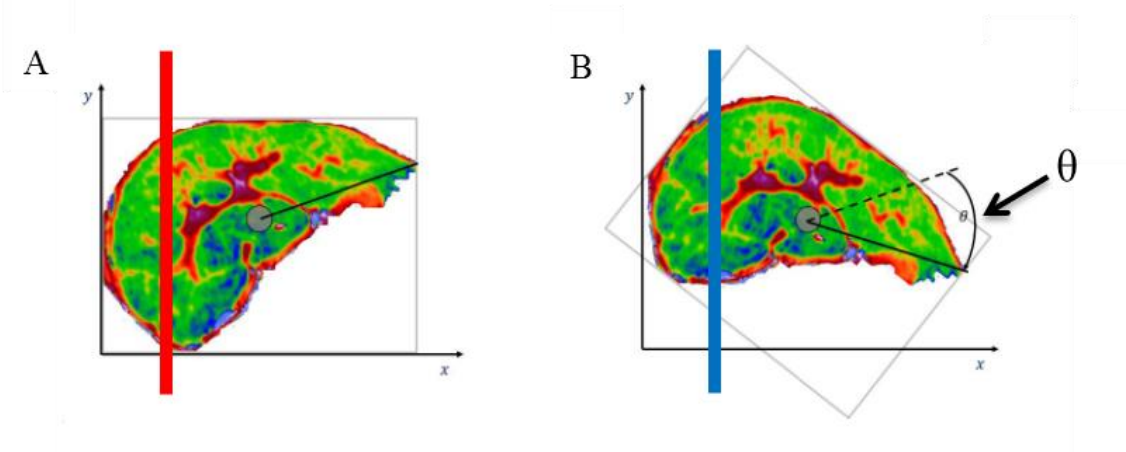


Figure 3.4-6: cT1 maps of transverse sections through the liver demonstrating rotation around the craino-caudal axis. A: Orientation of liver in situ with the red line demonstrating orientation of MRI slices. B: Orientation of the liver during preparation of histological specimens with the blue line demonstrating the orientation of histological slices. Image courtesy of Perspectum Diagnostics Ltd.

To overcome the issues relating to the orientation of the liver slice, the angle through which the liver had rotated was measured for all patients from transverse images of the liver. This angle of rotation is denoted θ in

Figure 3.4-6. A rotation matrix calculation was performed by Stella Kin at Perspectum Diagnostics Ltd. To calculate the position of the MRI slices that corresponds to the histological specimens. This leads to MRI slices being positioned as shown in Figure 3.4-7.

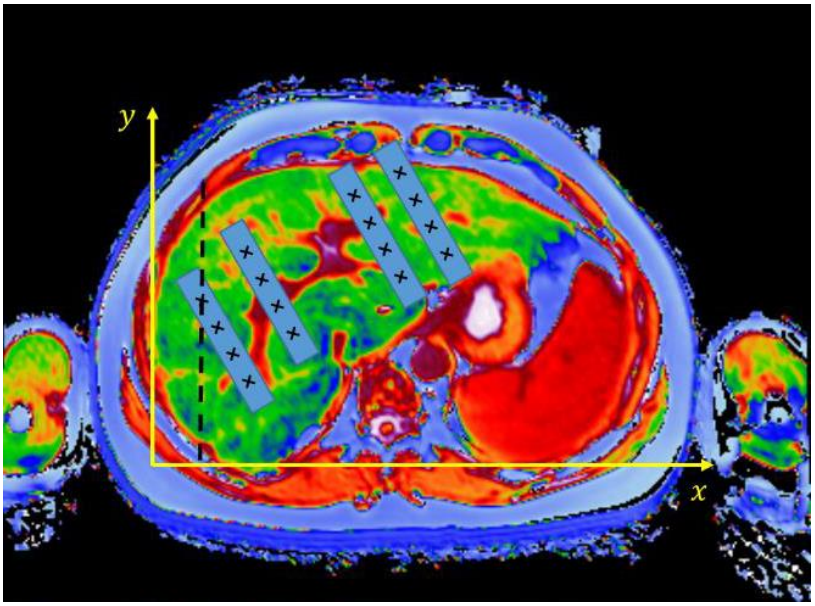


Figure 3.4-7: Demonstration of the plane of MRI slices following rotation through θ. Image courtesy of Perspectum Diagnostics Ltd.

The second challenge in the colocalisation of histology sample and MRI ROI was the deformation of the liver that occurs after removal from the body. The liver is not a rigid structure and removal of arterial and portal perfusion makes the liver flaccid and susceptible to changes in shape. Mobilisation of the liver and removal of the usual supporting structures caused a significant change in the shape of the liver. Some degree of shape change on removal from the body was anticipated but this was found to be much more pronounced than expected.

The two images in Figure 3.4-8 are the MRI slice and corresponding histology specimen showing the profound shape change that can occur.

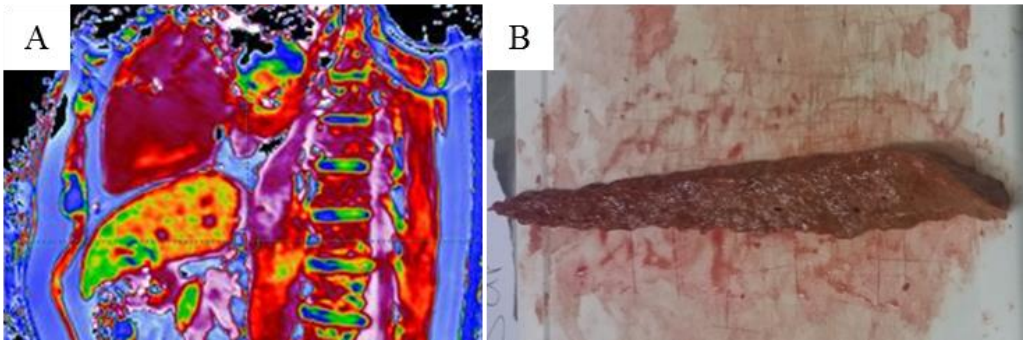


Figure 3.4-8: MRI image (A) and histology specimen (B) for the same slice from MURAL-B-007 showing the marked shape change that occurs between the liver in situ and ex-vivo.

The effect of this shape change is that the position of the histology sample measured from the explant does not always fall within the liver parenchyma on the MRI image. This is demonstrated in Figure 3.4-9.

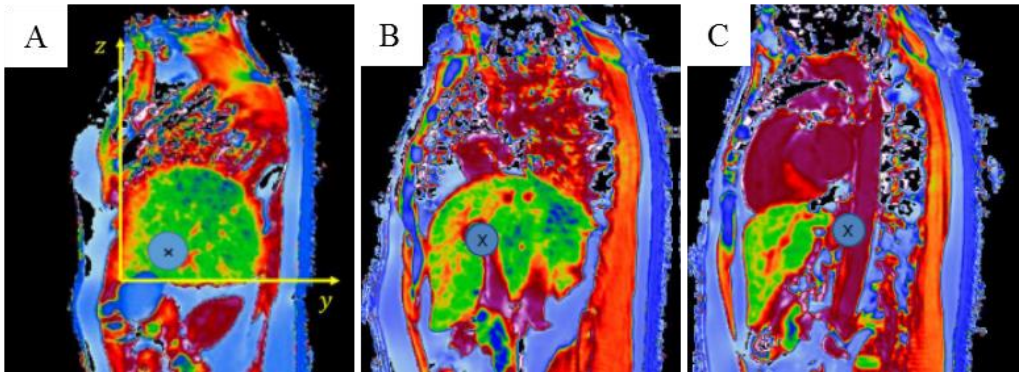


Figure 3.4-9: MRI slices from the same patient where the calculated position of the histology specimen (blue circle) falls A: on a major vessel, B: within the liver parenchyma and C: outside of the liver.

3.4.3 Results

Patient demographics

Demographic information, baseline investigations, underlying diagnoses and indication for transplant are outlined in Table 3.4-1.

Demographics

Total participants	10
Male	6 (60%)
Age (years)	60 (24-70)
BMI (kg/m^2)	25.3 (±4.0)
Interval between MRI and transplant (weeks)	11.3 (±6.3) (range 3.8 to 24.9)

Underlying diagnosis

Primary Biliary Cholangitis	3 (30%)
Hepatitis C	2 (20%)
Alcohol related liver disease	2 (20%)
Haemochromatosis	1 (10%)
Hepatitis B	1 (10%)
Primary Sclerosing Cholangitis	1 (10%)

Indication for transplantation

Hepatocellular carcinoma	6 (60%)
Hepatic failure	3 (30%)
Intractable puritis	1 (10%)

Fibrosis biomarkers

APRI	1.22 (0.35-2.61)
AST:ALT Ratio	0.96 (0.82-2.35)
FIB-4	3.12 (1.31-10.69)
ELF test	10.8 (7.7-12.1)
Median liver stiffness (kPa)	22.2 (7.9-45.0)

CONTINUED…

CONTINUED…

Prognostic scores

United Kingdom End-stage Liver Disease (UKELD) score	49.5 (±3.2)
Child-Pugh score	
5	5 (50%)
6	2 (20%)
7	2 (20%)
9	1 (10%)

Data presented as n (%), mean (±SD) or median (range) as appropriate.

Table 3.4-1: Demographic information and baseline investigations

Histological assessment

Semi-quantitative histological staging was possible for all 152 samples collected. MURAL-B-002 had a severely atrophic left lobe and it was not possible to take histological samples from this lobe. CPA analysis was available from 141 samples with technical failure of the image analysis process accounting for this reduction in available samples. Within individual patients there was little variation in semi-quantitative histology scores between the different samples from the liver. In particular there was very little variation in Ishak stage with no patient having variation of more than 1 stage. There was much greater variation in CPA between the different samples taken from the same liver. A summary of the CPA analysis is shown in Table 3.4-2.

	Number of samples	Mean (±SD) CPA (%)	Coefficient of variation
MURAL-B-002	8	4.09 (±0.75)	0.56
MURAL-B-006	15	19.70 (±7.55)	0.38
MURAL-B-007	16	25.78 (±2.74)	0.11
MURAL-B-011	8	21.08 (±3.03)	0.40
MURAL-B-013	16	10.20 (±2.59)	0.25
MURAL-B-017	16	16.98 (±3.98)	0.23
MURAL-B-018	16	18.71 (±5.07)	0.27
MURAL-B-020	15	14.66 (±2.02)	0.30
MURAL-B-021	16	23.98 (±6.36)	0.27
MURAL-B-022	15	18.39 (±10.51)	0.65
	Sum 141	Mean (±SD) of coefficient of variation	0.34 (±0.16)

Table 3.4-2: Summary of collagen proportionate area analysis.

For the 141 individual samples, CPA and Ishak stage had a strong relationship as shown in Figure 3.4-10.

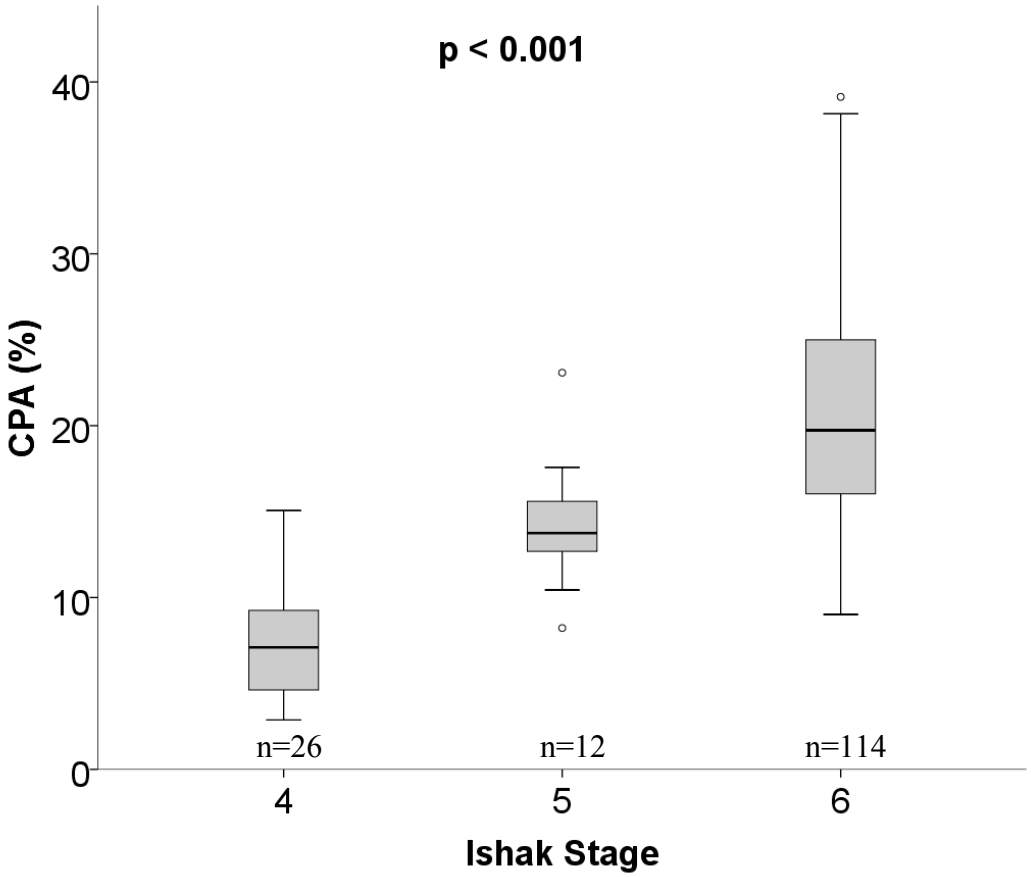


Figure 3.4-10: Box plot demonstrating the relationship between Ishak stage and CPA. P-value from the Jonckheere-Terpstra test.

The absence of significant variation in semi-quantitive scores across the different samples from each liver allows an overall histological opinion to be given for each explanted liver. Semi-quantitative scores for fibrosis, steatosis, siderosis and inflammation were as shown in Table 3.4-3.

Ishak stage

0-3	0 (0%)
4	2 (20%)
5	1 (10%)
6	7 (70%)

Brunt grade

0	6 (60%)
1	3 (30%)
2	1 (40%)
3	0 (0%)

Scheuer grade

0	3 (30%)
1	3 (30%)
2	1 (10%)
3	3 (30%)
4	0 (0%)

Portal inflammation

Minimal	0 (0%)
Mild	6 (60%)
Moderate	3 (30%)
Severe	1 (10%)

Lobular inflammation

Minimal	3 (30%)
Mild	6 (60%)
Moderate	1 (10%)
Severe	0 (0%)

Table 3.4-3: Semi-quantitative histological assessment of liver tissue samples.

Non-invasive markers of fibrosis do not correlate with histological assessment of fibrosis in this cohort with the exception of a strong relationship between APRI and CPA. Correlation coefficients are shown in Table 3.4-4. There was no statistically significant correlation between non-invasive markers.

Histological parameter		AST:ALT Ratio	APRI	ELF	FIB-4	Liver stiffness
Overall Ishak Stage	Rho	0.261	0.242	0.634	0.149	0.283
	p	0.498	0.530	0.067	0.702	0.496
	n	9	9	9	9	8
Mean CPA	Rho	0.400	**0.700**	0.383	0.583	0.587
	p	0.286	**0.036**	0.308	0.099	0.126
	n	9	9	9	9	8

Table 3.4-4: Correlation between established non-invasive biomarkers of fibrosis and histological assessment of fibrosis.

In this cohort, markers of liver disease severity (UKELD and Child-Pugh score) correlated strongly with each other (Spearman's Rho = 0.71, p=0.02) but, as shown in Table 3.4-5, did not correlate with markers of fibrosis.

	UKELD			Child-Pugh Score		
	Rho	p	n	Rho	p	n
AST:ALT Ratio	0.183	0.637	9	0.158	0.685	9
APRI	0.367	0.332	9	0.474	0.197	9
ELF	-0.217	0.576	9	-0.061	0.875	9
FIB-4	-0.133	0.732	9	-0.018	0.964	9
Median liver stiffness	-0.132	0.756	8	-0.051	0.904	8
Ishak stage	-0.166	0.646	10	-0.175	0.629	10
CPA	0.188	0.603	10	0.052	0.886	10

Table 3.4-5: No correlation between severity scores and markers of fibrosis.

Multiparametric MRI assessment of whole liver

cT1 values for the whole liver analysis are shown in Table 3.4-6. cT1 varies little between the different summary statistic used (mean, median or mode) with a trend towards higher values when using the mean. Exclusion of vascular structures before calculating the mode has little impact on the cT1 value.

ID number	Mean cT1 (msec)	Coefficient of variation	Median cT1 (msec)	Mode cT1 (msec)	Mode cT1 - excluding vessels (msec)
MURAL-B-002	784.3	0.13	760.3	740	743
MURAL-B-006	943.2	0.10	917.3	900	897
MURAL-B-007	1063.1	0.10	1015.0	980	982
MURAL-B-011	820.2	0.11	808.3	780	793
MURAL-B-013	811.0	0.11	782.3	740	748
MURAL-B-017	838.6	0.11	822.3	820	808
MURAL-B-018	865.9	0.11	843.3	820	807
MURAL-B-020	1002.1	0.09	1001.0	1020	1002
MURAL-B-021	1059.4	0.09	1050.0	1020	1035
MURAL-B-022	763.7	0.16	738.9	700	704

Table 3.4-6: cT1 values for whole liver analysis of each participant. Summarised by various summary statistics.

For whole liver cT1 analysis the mode cT1 without vessel subtraction had the strongest correlation with overall Ishak stage (Rho 0.798, p=0.006). Mean, Median and Mode with vessel subtraction also had strong correlation with overall Ishak stage (Rho 0.791, p=0.006 for all). Only mean cT1 correlated with mean CPA in a statistically significant manner (Rho=0.636, p=0.048). None of the whole liver cT1 measures correlated with prognostic markers. These correlations are shown in Table 3.4-7 and the relationships between mean cT1 and histology parameters can be seen graphically in Figure 3.4-11.

		Overall Ishak Stage	Mean CPA	UKELD	C-P score
mean cT1 whole liver	Rho	**0.791**	**0.636**	-0.018	-0.169
	p	**0.006**	**0.048**	0.960	0.640
	n	**10**	**10**	10	10
mode cT1 whole liver	Rho	**0.798**	0.489	-0.153	-0.125
	p	**0.006**	0.151	0.673	0.731
	n	**10**	10	10	10
mode (excluding vessels) cT1 whole liver	Rho	**0.791**	0.515	-0.030	-0.007
	p	**0.006**	0.128	0.934	0.986
	n	**10**	10	10	10
median cT1 whole liver	Rho	**0.791**	0.624	-0.030	-0.078
	p	**0.006**	0.054	0.934	0.830
	n	**10**	10	10	10

Table 3.4-7: Correlation of whole liver cT1 with overall Ishak stage, mean CPA, UKELD and Child-Pugh score.

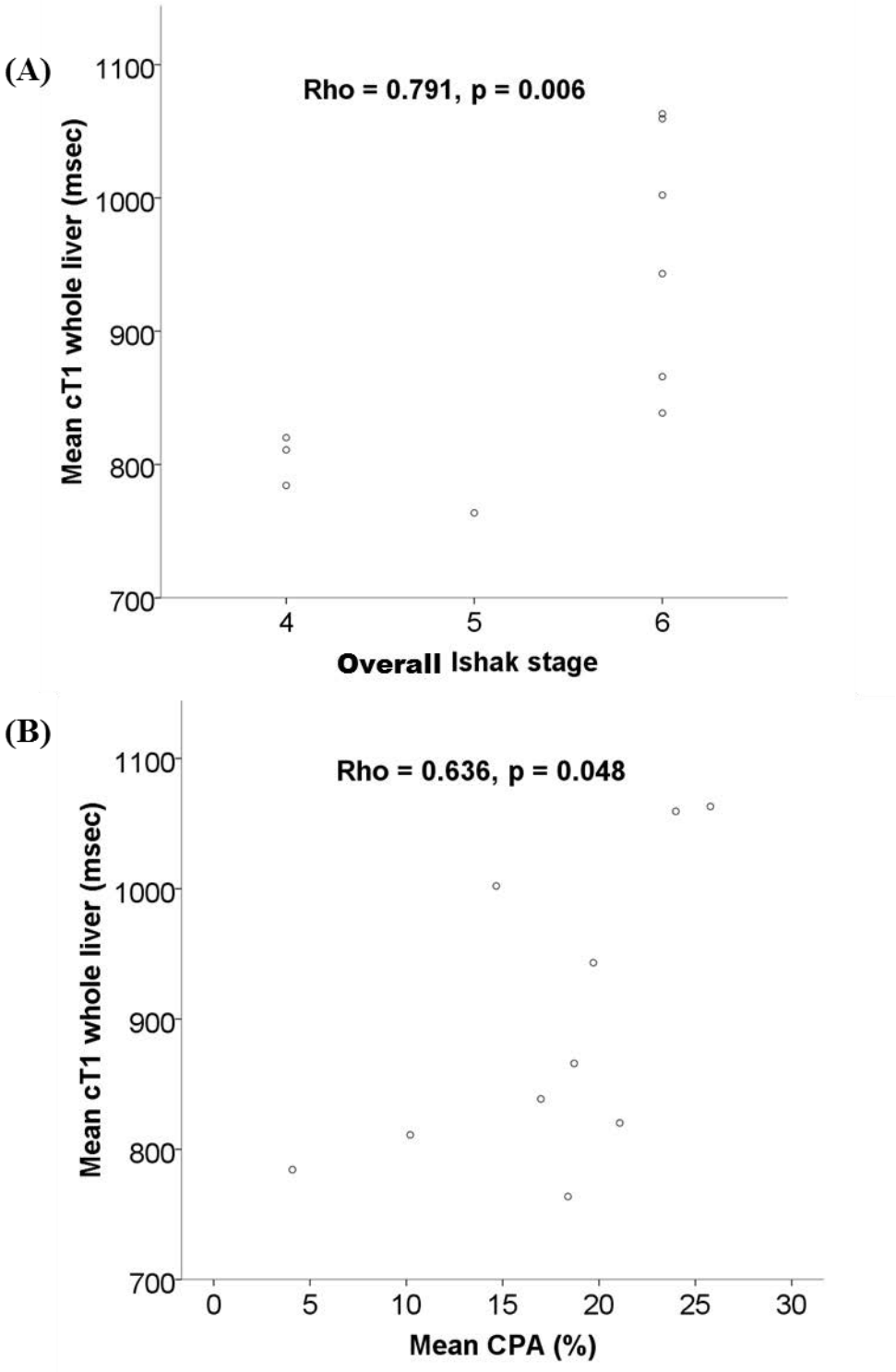


Figure 3.4-11Scatter plots showing the relationship between (A) whole liver mean cT1 and Ishak stage and (B) whole liver mean cT1 and CPA. Correlation assessed with Spearman's Rho.

Multiparametric MRI assessment of individual colocalised samples

Due to the aforementioned difficulties in colocalising MRI and histology samples there are a limited number of paired samples for each participant. Overall, 54/160 (34%) samples have both histology and MRI data. Success in colocalising histology and MRI data was not uniform either between patients or between slices within individuals. This is summarised in Table 3.4-8. It can be seen that, in general, the problem with colocalising MRI and histology samples is worse for slices C and D than for A and B. The measurement used to position the slices was taken from the right lateral edge of the liver. This means that slices C and D are further from the origin of the measurement and it would be expected that they are more severely affected by the deformation of the liver. For MURAL-B-002 the left lobe of the liver was severely atrophied and therefor it was not possible to take histological samples from the left lobe. For MURAL-B-006 there were no sagittal plane images acquired and so it was not possible to match histology and MRI data for this participant.

ID number	Number of paired samples available				
	Whole liver (max 16)	Individual slices			
		Slice A (max 4)	Slice B (max 4)	Slice C (max 4)	Slice D (max 4)
MURAL-B-002	6	4	2	Atrophic left lobe	
MURAL-B-006	No sagittal plane MRI images				
MURAL-B-007	0	0	0	0	0
MURAL-B-011	6	3	3	0	0
MURAL-B-013	4	2	2	0	0
MURAL-B-017	6	3	3	0	0
MURAL-B-018	10	4	4	2	0
MURAL-B-020	7	2	3	2	0
MURAL-B-021	10	3	4	2	1
MURAL-B-022	5	2	3	0	0

Table 3.4-8: Number of paired samples per participant

For individual colocalised samples there was moderate correlation between cT1 and Ishak stage (Rho=0.525, p<0.001). The Jonckheere-Terpstra test shows significant difference in cT1 across the three Ishak stages seen (p < 0.001) with post hoc tests giving significant differences between Ishak stages 4 and 6 only (p < 0.001). This relationship is shown in Figure 3.4-12.

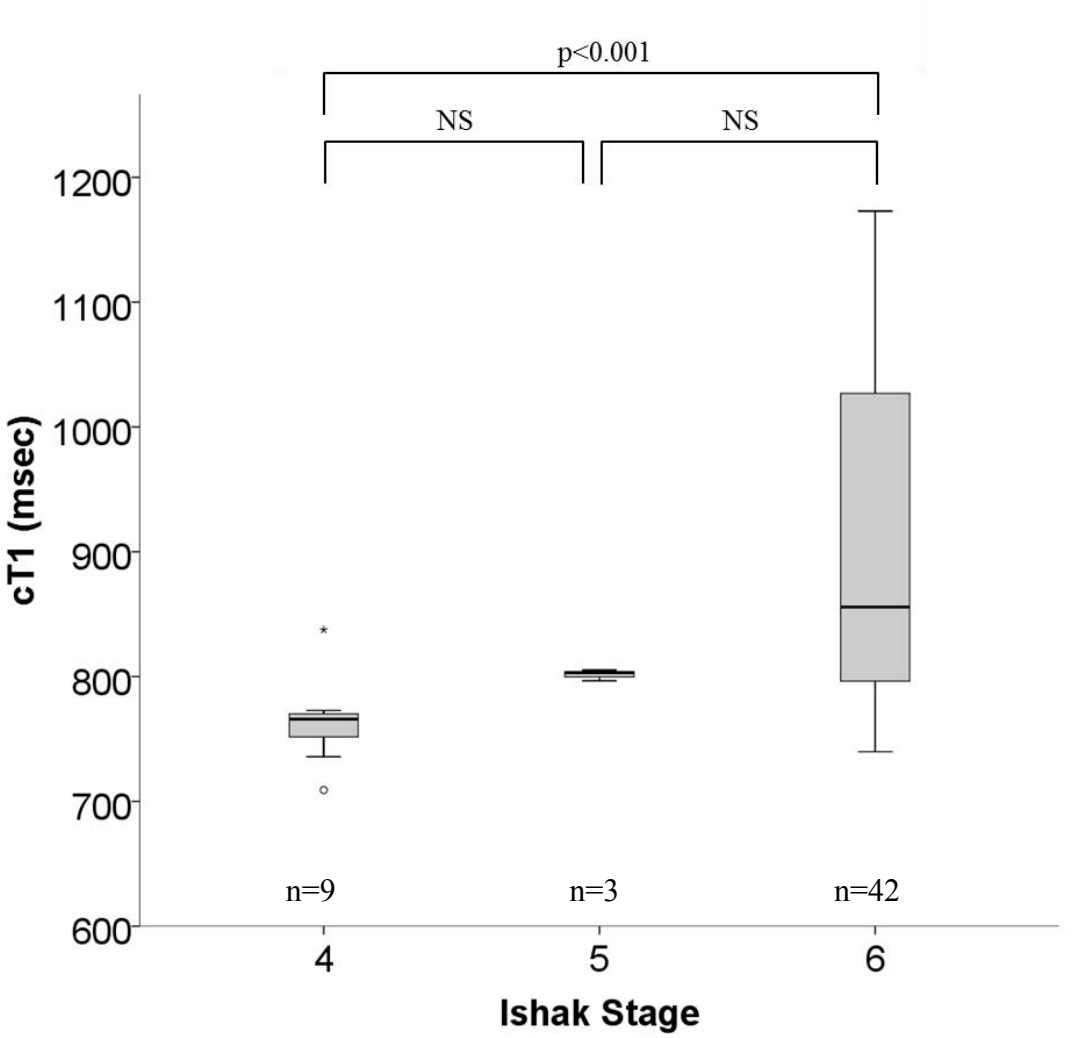


Figure 3.4-12: Box plot showing the relationship between cT1 and Ishak stage for paired samples (n=54). Overall significance p<0.001 by the Jonckheere-Terpstra test. Inter-group differences from Dunn's test.

As noted above, the variation in Ishak stage across the 54 samples for which it was possible to place an MRI ROI was very low and, as shown in Table 3.4-9, the variation in cT1 is also low. Coefficient of variation was from 2-6%. The variation in CPA was much greater with coefficient of variation from 11-65%. The coefficient of variation for each participant is shown in Table 3.4-9.

	Coefficient of Variation	
	CPA (n=141)	**cT1 (n=54)**
MURAL-B-002	56%	6%
MURAL-B-006	38%	-
MURAL-B-007	11%	4%
MURAL-B-011	40%	6%
MURAL-B-013	25%	2%
MURAL-B-017	23%	4%
MURAL-B-018	27%	5%
MURAL-B-020	30%	6%
MURAL-B-021	27%	4%
MURAL-B-022	65%	4%
Mean (±SD) of coefficient of variation	34 (±16)%	4.5 (±1.3)%

Table 3.4-9: Coefficient of variation for CPA and cT1 for colocalised samples. This shows that there was much greater variation in CPA between different areas of the liver than was evident from cT1 measurement.

For individual colocalised samples, CPA has no significant correlation with cT1 (Rho = 0.252, p = 0.077). The relationship can be seen in Figure 3.4-13. When cT1 is plotted against CPA, there appears to be two distinct groups. These are marked on Figure 3.4-13.

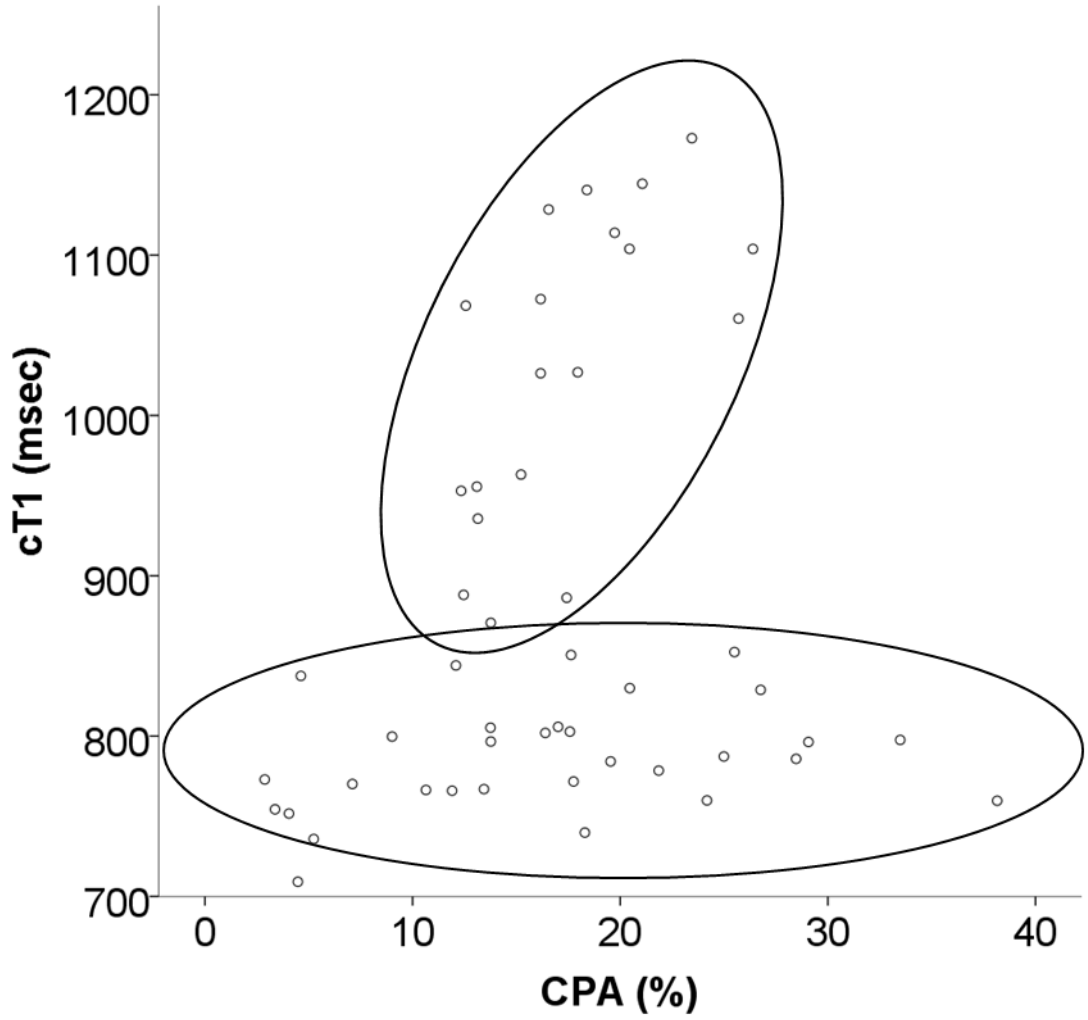


Figure 3.4-13: Scatter plot showing the relationship between CPA and cT1 for paired samples. The two ovals highlight what appear to be two distinct groups.

In Figure 3.4-14 it can be seen that the presence of inflammation has an impact on cT1 that appears to be independent of CPA. The samples without inflammation (circles ●) show very little change in cT1 over the whole range of CPA seen in this cohort. Inflammation in either portal areas or lobules (crosses +) leads to a marked elevation in cT1 and extensive inflammation throughout the liver (triangles ▲) shows marked elevation in cT1 that is independent of the degree of fibrosis.

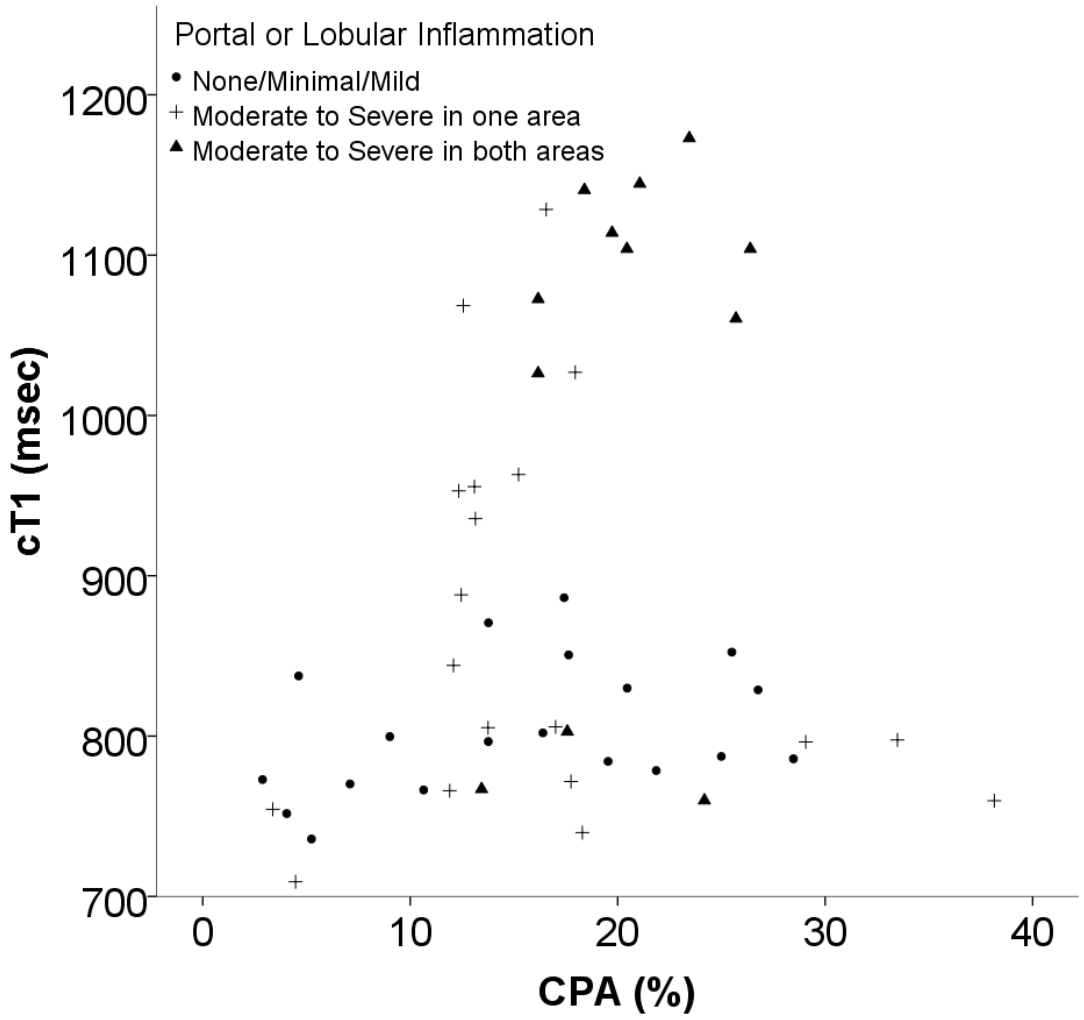


Figure 3.4-14: Scatter plot showing the relationship between CPA and cT1 stratified by the degree of inflammation. Samples with more significant inflammation show markedly higher cT1.

3.4.4 Discussion

This experiment aims to further investigate the ability of cT1 to stage hepatic fibrosis using multiple histological samples from explanted livers as the reference standard. The explanted livers were taken from a cohort of patients on the liver transplant waiting list at University Hospitals Birmingham. The demographics of this cohort are typical of patients awaiting liver transplant in the UK except that UKELD and Child's-Pugh score are lower due to the relative excess of patients in this cohort being transplanted for HCC. These patients tend to have less severe liver failure and so a lower UKELD than the majority of patients on the list. However, despite this driver towards less severe liver disease, it is clear from histology and non-invasive biomarkers that all patients have advanced hepatic fibrosis with all but 2 patients having cirrhosis. In this cohort the established non-invasive biomarkers used do not differentiate between different histological stages of fibrosis and do not correlate with CPA. The blood biomarkers used in this experiment are designed to detect the presence of advanced fibrosis and are not designed to grade the 'severity' of fibrosis once advanced fibrosis or cirrhosis is established. In particular the indirect markers used (AST:ALT, APRI and Fib-4) are calculated from blood tests that reflect the consequences of portal hypertension rather than the volume of fibrosis per se. This is likely to explain why these tests do not correlate with CPA in this cohort. Non-invasive fibrosis markers also do not correlate with either of the prognostic markers used in this experiment. Many of the factors taken into account when calculating both UKELD and Child's-Pugh score are the result of portal hypertension and thus are related to the architectural changes in the liver found in cirrhosis. These changes are not dependent on the volume of collagen in the liver and this may be why there is no correlation between markers of fibrosis and prognostic markers in this cohort.

When looking at whole liver assessment the summary statistic used has little effect on the recorded cT1. There is a trend towards higher values when the mean is used. The digital subtraction of vessels before calculation of the mode cT1 value has little impact. It should also be noted that cT1 values calculated from the whole liver analysis are markedly lower than was expected from previous data. In the experiment using biopsy as a reference standard (section 3.3) the mean cT1 for patients with Ishak stage 0 fibrosis was 851.7 msec. In this experiment, 5 of the patients (who all have advanced fibrosis or cirrhosis) have a mean cT1 lower than this. This is particularly unexpected given that the cT1 of vascular and biliary structures is higher than that of liver parenchyma regardless of fibrosis stage. Unless specifically excluded, whole liver analysis leads to the inclusion of vascular structures and so it would be expected that whole liver analysis would give a higher cT1 value than a single ROI. The reason for this difference in the range of cT1 values seen is not clear. To date, comparison of single ROI measurement and whole liver analysis in a large cohort has not been conducted. In order to find if whole liver analysis improves the accuracy of multiparametric MRI assessment of liver fibrosis, further work comparing the cT1 value from single ROI and whole liver analysis should be performed in a larger cohort.

Measurement of cT1 from whole liver assessment shows strong correlation between cT1 and overall Ishak stage regardless of the summary statistic used. Mean cT1 also shows a strong correlation with mean CPA and median cT1 has a moderate correlation with mean CPA that just fails to reach statistical significance. The correlation between mean CPA and mean cT1 is far stronger than the correlation between mean CPA and cT1 measured from a single ROI ($Rho=0.333$, $p<0.001$). As discussed in section 3.3, the use of human judgement to place the ROI and the inaccuracy this introduces is a potential source of error. The improved correlation

with whole liver analysis could support the use of mean cT1 measured on whole liver analysis rather than single ROI for fibrosis assessment. This is a potential avenue for further work in this field.

cT1 from whole liver assessment has not been shown in this cohort to correlate with either UKELD or Child's-Pugh score. This is however a small sample and the inability of cT1 to give prognostic information should be viewed in this context. Formal power calculations have not been performed but the study is almost certainly underpowered to detect this association.

There is a clear theoretical advantage in matching the location of histology sample and cT1 measurement. This should reduce sampling error in the reference standard and therefore improve the confidence in the reliability of our evaluation of multiparametric MRI. The variation in CPA demonstrated in Table 3.4-9 clearly demonstrates the variations in fibrosis severity across the liver and supports the need for colocalised samples. However, the assessment of colocalised samples in this cohort is limited by the difficulties with the technique as described in section 3.4.2. Although some shape change in the liver at the time of hepatectomy was anticipated it was far more marked than expected. It is difficult to be fully confident that supposedly colocalised samples are indeed assessing the same volume of liver tissue.

MRI data were not assessed until after all patients had been transplanted and so the problem with rotation and deformation of the liver were not appreciated until it was too late to modify the methods of tissue collection. To avoid the need to re-format and rotate the MRI slices, the MRI slices could be acquired perpendicular to the posterior surface of the liver. To reduce

deformation of the liver it may be possible to reduce the time between hepatectomy and tissue sampling, however, I would expect that the two most significant factors in the shape change in the liver are the mobilisation of the liver from its supporting tissue and removal of vascular inflow causing the liver to become flaccid. Both of these events occur at the time of hepatectomy and therefore they are not possible to avoid. It therefore may be that precise colocalisation of a MRI ROI and a histology sample is not technically possible. A better method may be to take cT1 measurements from each of the 8 anatomical segments of the liver. This anatomy is maintained after the liver is removed from the body and a histological sample could be taken from each segment. This would result in fewer samples per liver but seems more likely to give confidence that colocalised samples are accurately matched.

When assessing the accuracy of fibrosis assessment from individual colocalised samples there is a moderate correlation between cT1 and Ishak stage. There is no significant correlation between CPA and cT1 in these colocalised samples. The very small numbers of samples with less than Ishak stage 6 fibrosis (Ishak 5: 3, Ishak 4: 9) should be kept in mind when looking at the correlation between cT1 and Ishak stage.

In the CPA data presented here we have, as expected,[191] demonstrated that the severity of fibrosis varies across the liver. The mean (±SD) coefficient of variation across a single liver was 34.0 (±16.0)%. This variation is not reflected in variations in Ishak stage. The descriptive, nature of Ishak staging is not influenced directly by the volume of collagen and, once cirrhosis is established, there is no further capacity for Ishak stage to grade the severity of fibrosis. CPA however has been shown to add further relevant clinical information on top of the Ishak stage.[186] In people with cirrhosis, CPA correlates with the severity of portal hyper

tension and predict the likelihood of clinical decompensation.[26, 42] This would suggest that CPA is a more useful measure of fibrosis when assessing the performance of cT1. In contrast to CPA, cT1 shows very little variation across the liver. When performing whole liver analysis, the mean (±SD) coefficient of variation is 11.1 (±2.1)% and when looking at multiple ROIs, the mean (±SD) coefficient of variation is lower at 4.5 (±1.3)%. This implies that cT1 is not sensitive to small changes in fibrosis severity clearly demonstrated by variations in CPA.

As noted previously, inflammation seems to have a major impact on cT1. Plotting cT1 against CPA for the colocalised samples shows 2 distinct groups as shown in Figure 3.4-13. It seems that the driver that differentiates these two distinct groups is inflammation. Samples are stratified by the presence of inflammation in Figure 3.4-14. No sample with anything less than moderate inflammation had a cT1 over 900 msec. There is a trend for samples with more extensive inflammation (▲) to have a higher cT1.

3.5 Conclusion

These data show that cT1 calculated from Liver*Multiscan*™ analysis of multiparametric MRI data has the ability to stage hepatic fibrosis with accuracy comparable to most of the established non-invasive biomarkers of hepatic fibrosis included in this work. The initial pilot data[163] for Liver*Multiscan*™ demonstrated a clear advantage for this technology over existing biomarkers, particularly in the identification of early stage fibrosis however this has not been replicated in this work. The reason for this is not clear but in this larger and more diverse cohort the confounding effects of inflammation and steatosis seem relevant. In particular the significant confounding effect of inflammation would make cT1 very difficult to interpret in clinical practice. There is a suggestion that the use of 'whole liver analysis' may improve the accuracy of Liver*Multiscan*™ for the identification of fibrosis however, this remains to be proven and is an interesting avenue for further work in this field.

In this work cT1 is outperformed by liver stiffness measured with VCTE across all stages of fibrosis. Although the multiparametric MRI acquisition protocol can be completed in approximately 20 minutes, MRI remains an expensive and time consuming activity compared to all other evaluated non-invasive biomarkers. A formal assessment of cost effectiveness is beyond the scope of this thesis but this will need careful consideration if multiparametric MRI is to find a role in clinical practice.

The value of matching multiparametric MRI ROI and histology samples is supported by the variation in fibrosis demonstrated with CPA analysis of these samples. Refinement of the

method to use the segmental anatomy of the liver should improve this technique and should allow the collection of more accurate data.

cT1, clearly is not simply a marker of liver fibrosis and the multiple significant confounding factors make it a difficult parameter to interpret. It is possible that cT1 should be regarded as a marker of 'fibro-inflammatory disease'. Whether fibro-inflammatory disease is clinically relevant remains to be seen however it is plausible that a technology that can identify, and perhaps more importantly exclude, fibrosis and/or inflammation may have a role in clinical practice. NAFLD is a highly prevalent liver disease where inflammation (non-alcoholic steatohepatitis) and fibrosis are relevant to the disease process. Chapter 4 examines the potential application of multiparametric MRI in NAFLD.

CHAPTER 4: THE ASSESSMENT OF NON ALCOHOLIC FATTY LIVER DISEASE WITH MULTIPARAMETRIC MRI

4.1 Non-Alcoholic Fatty Liver Disease

Hepatic steatosis is the accumulation of fat within the liver and has been recognised since Thomas Addison's description of "*fatty degeneration of the liver*" secondary to alcohol excess in 1836.[192] Hepatic steatosis occurs most commonly as a result of alcohol excess and for most of the 19th and 20th centuries hepatic steatosis in the absence of alcohol excess was regarded as being of no clinical consequence. Through the 1960s and 1970s a pattern of liver disease, histologically identical to alcoholic liver disease but without associated alcohol excess, was increasingly recognised as a disease entity in its own right[193-195] culminating in the seminal description by Ludwig et al in 1980 of a series of 20 patients with "*liver disease that histologically mimics alcoholic hepatitis and that also may progress to cirrhosis*".[196] This paper coined the term non-alcoholic steatohepatitis (NASH) and noted the strong association with obesity and type two diabetes.[196]

Later work highlighted that the phenotype of non-alcoholic fatty liver disease (NAFLD) is a spectrum ranging from a relatively benign and non-progressive steatosis to a progressive, inflammatory and fibrotic liver disease that can lead to cirrhosis and hepatocellular carcinoma (HCC).[197]Today, NAFLD is recognised as the hepatic manifestation of the metabolic syndrome and due to the enormous prevalence of obesity and insulin resistance it is the most common chronic liver disease worldwide.[198] Estimates of the prevalence of NAFLD vary widely due to differences in the populations studied and the different diagnostic tools used[198] but is thought to affect 20-30% of adults in western societies.[199] The incidence of NAFLD rises dramatically in the obese and, in patients undergoing bariatric surgery, reaches 90%.[200]

A recent meta-analysis estimated the prevalence of NALFD to be 25.24% of the global popluation.[201]

Although it is clear that the prevalence of NAFLD in the community is enormous, the progression to advanced liver disease is less well understood. Not all patients with NAFLD will develop the complications of advanced liver disease and identifying patients at risk of progressive disease and complications is vital to allow the focusing of limited resources.

4.1.1 Non-Alcoholic Steatohepatitis and Simple Steatosis

The spectrum of NAFLD is divided into two main diagnostic groups. Simple steatosis (SS) is hepatic steatosis without evidence of liver injury. It is widely accepted that SS leads to little or no progression of fibrosis and no increase in liver related mortality.[197, 202-204] In stark contrast to SS is NASH, which is characterised histologically by hepatocyte ballooning and lobular inflammation. NASH is a source of ongoing liver injury, causes progressive fibrosis and cirrhosis in up to 20%[205] and carries a HCC risk of 5.29 cases per 1,000 person-years.[201] A 2009 systematic review shows that the presence of NASH is an independent risk factors for progression of fibrosis in NAFLD.[206]

Current clinical practice is focused on the identification of fibrosis in NAFLD[80] however identification of NASH may identify patients at an earlier point in their natural history and allow intervention to prevent fibrosis.[207] Identifying those with NASH may identify patients with more serious liver disease, guide the intensity of clinical follow-up and indicate prognosis.[208] Another benefit in identifying patients with NASH is to enrich clinical studies.

In the current absence of effective and evidence based therapies for NASH,[209] interventional trials are of great importance and the identification of NASH highlights patients eligible for clinical studies.[210]

4.2 Currently available techniques for the identification of NASH

The distinction between SS and NASH is currently dependent on liver biopsy histology although insights into the mechanisms underpinning NASH have led to investigation into other, non-invasive, markers.

4.2.1 Adipokines

Adipokines are cell signalling hormones secreted by adipose tissue and include leptin, adiponectin, tissue necrosis factor alpha (TNF-α) and interlukin 6 (IL-6). These cytokines control diverse metabolic processes including satiety, fatty acid metabolism, glycaemic control and regulation of fat stores. Adipokines have been implicated in the mechanisms underlying insulin resistance and thus are potentially relevant in the pathogenesis of NAFLD.[211] Adiponectin has been shown in small, single centre studies to be lower in patients with NASH when compared to those with SS[212] and healthy controls.[213] A combination panel of adipokines has also been shown to discriminate between healthy volunteers and NAFLD.[214] These studies, although mechanistically interesting are small and do not support the routine use of adipokine measurements for the staging of NAFLD.

4.2.2 Cytokeratin-18

Cytokeratin (CK)-18 is a filament protein abundant in hepatocytes. During cellular damage or apoptosis CK-18 is released from cells. This can be as complete protein (M65 antigen) or as enzymatically cleaved fragments (M30 antigen). Several studies have investigated the

potential of CK-18 (both M65 and M30) to detect the presence of NASH. To date however studies are small and diverse in their methodology. A 2012 study including 146 patients with NAFLD found that CK-18 (M30) was significantly elevated in NASH compared to SS but sensitivity and specificity were modest at 66%.[215] A 2013 meta-analysis found slightly improved performance with pooled specificity for both M65 and M30 of 71% while pooled sensitivity was higher for M30 (83%) than M65 (77%).[216] However, this meta-analysis found significant heterogeneity and publication bias within the included studies.[216] To date, the largest single study assessing the utility of CK-18 fragments has shown a modest AUROC (95% CI) of 0.65 (0.59–0.71) and the optimal cut off value gave sensitivity and specificity of 58% and 68% respectively.[217]

4.2.3 Liver biopsy

The above mentioned disappointing results have cemented the position of liver biopsy histology as the gold standard test for the differentiation of NASH from SS. The diagnosis of NASH on biopsy is dependent on the identification of ballooned hepatocytes and lobular inflammation. These histological features can also be graded to give an overall diseases activity. One such scoring system in the NAFLD activity score (NAS).[25, 188] NAS grades several histological features of NAFLD on a categorical scale. The sum of these scores gives the NAS. This score was designed to be used in research to monitor changes in disease activity and it is not a score by which NASH can be differentiated from SS. Indeed the diagnosis of NASH by expert pathologists has been shown to be independent of NAS.[218]

4.3 Assessment of fibrosis in NAFLD

It is clear that the staging of fibrosis in NAFLD is a key indicator of clinical outcome.[14, 21, 219] Angulo and colleagues looked retrospectively at 619 patients with NAFLD who had undergone liver biopsy between 1975 and 2005. Their data show that the presence of NASH did not predict transplant free survival whereas the presence of fibrosis did.[14] This is shown in Figure 4.3-1. Similar results were demonstrated by Ekstedt et al who showed that advanced fibrosis (Kleiner stage 3-4) was the only factor associated with increased risk of death.[21] Recent work from the same group using Swedish registry data shows that fibrosis stage predicts survival and the rate of progression of liver disease whereas the presence or absence of NASH does not.[219] These studies are limited by their retrospective design and, in earlier work, the relatively small numbers of patients in some groups. However, the importance of fibrosis assessment in determining the prognosis for patients with NAFLD is now widely accepted.[220]

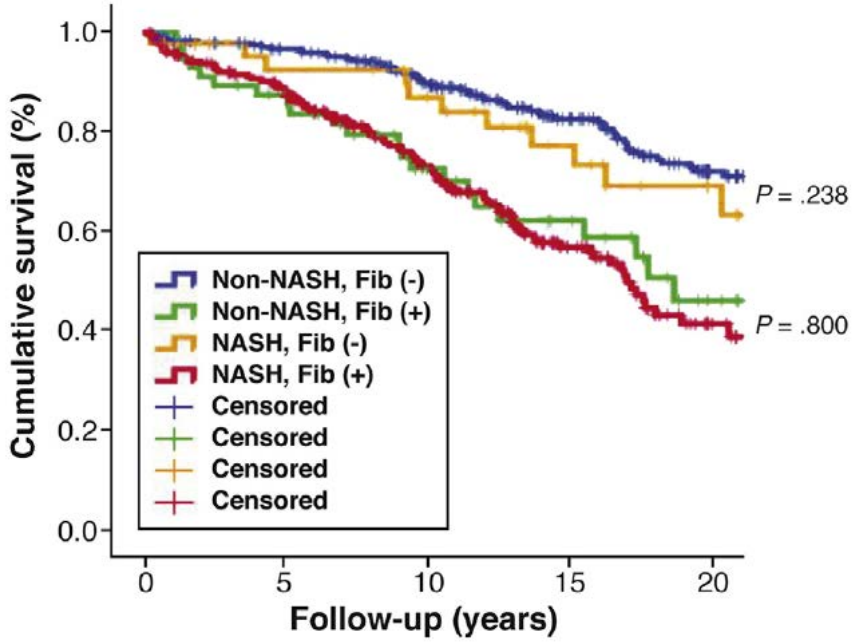


Figure 4.3-1: Kaplan-Myer plot demonstrating that transplant free survival is predicted by the presence of fibrosis at baseline (red and green lines) and not the presence of NASH (yellow and blue lines). Taken from Angulo P et al, Gastroenterology, 2015.

4.3.1 Currently available techniques for fibrosis assessment in NAFLD

Methods for invasively and non-invasively assessing hepatic fibrosis are discussed in detail in Chapter 1 and the performance of non-invasive biomarkers in patients with NAFLD is comparable to performance in other aetiologies. Overall, non-invasive tests of hepatic fibrosis have high accuracy for the diagnosis of advanced fibrosis and cirrhosis but do not have the sensitivity to reliably identify lower stage fibrosis whereas liver biopsy histology is limited by patient acceptance and sampling error.

One notable addition to the previously discussed biomarkers is the NAFLD fibrosis score (NFS). The NFS is an indirect fibrosis marker derived from, and validated in, a population of patients with NAFLD.[221] It is comprised of readily available clinical and laboratory parameters (age, presence of insulin resistance, BMI, platelet count, albumin, AST and ALT) and is designed to predict advanced fibrosis (Kleiner stage ≥3). More recent evidence has emerged that high NFS also predicts mortality in patients with NAFLD.[222] The use of a dual cut-off values allows the low cut off to achieve a very high negative predictive value of 88-92%.[103, 221, 223] The use of dual cut-off values is however a limitation of NFS in practice. Approximately 25% of cases fall between the cut-off values, are therefore indeterminate, and require alternative testing.[221]

4.3.2 The need for novel biomarkers in NAFLD

There is a clear need for the development of a novel biomarker for use in NAFLD that can identify patients with NAFLD and stratify patients in terms of fibrosis and the presence of NASH. Such a biomarker would be of value in clinical practice and as an endpoint in clinical trials.

4.3.3 Liver*Multiscan*™

Liver*Multiscan*™ (Perspectum Diagnostics Ltd., Oxford, UK) is a proprietary multiparametric MRI technology used to quantify liver fat, iron and fibro-inflammatory liver injury by proton density fat fraction (PDFF), T2* mapping and corrected T1 (cT1)[224] mapping respectively.

cT1 of the liver has previously been reported to stage hepatic fibrosis in NAFLD[225] and an unselected population of patients undergoing liver biopsy.[163] PDFF is measured using a modified Dixon sequence is a well-established and accurate technique for the assessment of hepatic fat content.[226-228] Iron concentration was estimated from T2* according to a previously determined model.[229]

4.4 Aims

Liver*Multiscan*™ has potential to be a useful biomarker in NAFLD for use in both clinical practice and as an endpoint in trials. We sought to determine how Liver*Multiscan*™ performed in terms of utility and comparative effectiveness, in the assessment of a prospective cohort of patients with NAFLD having routine liver biopsy as standard of care.

4.5 Methods

4.5.1 Study Participants

Our prospective study was undertaken at the Queen Elizabeth Hospital Birmingham and Royal Infirmary of Edinburgh between February 2014 and September 2015. The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki, and was approved by the National Research Ethics Service (West Midlands – The Black Country; REC Ref: 14/WM/0010). The study was registered with the ISRCTN registry (ISRCTN39463479) and the National Institute of Health Research (NIHR) portfolio (15912). The study sponsor was the University of Birmingham. Male and female adult (≥18 years of age) patients booked for non-targeted liver biopsy for any indication were prospectively recruited to a validation study of Liver*MultiScan*™ (reported in Chapter 3). Those patients with a histological diagnosis of NAFLD were included in this sub-group analysis. Exclusion criteria were: biopsy of a distinct focal lesion, inability to give fully informed consent and any contraindication to MRI. Patients with a histologically confirmed diagnosis of NAFLD without secondary cause and without history of alcohol excess (men >21 UK units/week, women >14 UK units/week) were included in this sub-group analysis.

Healthy volunteers were recruited from staff and students at the University of Birmingham. Exclusion criteria were obesity (Body mass index (BMI) >30kg/m^2), current or previous history of liver disease, significant medical co-morbidity, family history of liver disease, excess alcohol intake or any contraindication to MRI. Participants gave written, informed consent and attended for a single study visit during which they underwent multiparametric

MRI, FibroScan™ examination, blood sampling and collection of clinical and demographic data. All study investigations were performed after a 4 hour fast. Patients undertook their study visit in the 2 weeks prior to liver biopsy. Healthy volunteers did not undergo liver biopsy.

4.5.2 Study Investigations

MRI scans were performed as described in Chapter 2 and analysed with Liver*Multiscan*™ to generate values for cT1, T2* and fat fraction (PDFF-Dixon). Patients also underwent MR spectroscopy as described in Chapter 2 to measure fat fraction (PDFF-MRS).

FibroScan™ examinations were performed by trained operators (PJE and NM) in accordance with manufacture's guidelines and validated local clinical practice.[91] The decision on using the M probe or XL probe was made on the skin to liver capsule distance measured by the FibroScan machine. Examinations were regarded as 'possible' if at least 10 valid readings could be recorded and 'reliable' if they contained at least 10 valid readings and had interquartile range (IQR) to median ratio ≤30% (Boursier's criteria).[182] At the start of the study the Controlled Attenuation Parameter (CAP) was not available on the FibroScan™ XL probe. CAP on the XL probe was enabled during study recruitment so was recorded if available in addition to median liver stiffness (LS).

Blood samples were analysed routinely for markers of liver disease. Simple blood biomarker panels including AST/ALT ratio, FIB-4 and NAFLD fibrosis score (NFS) were calculated

according to published formulae.[49, 221] Sera were also analysed to determine the ELF score (iQur Limited, London, UK).

4.5.3 Histological assessment

Liver biopsy samples were taken with 16 gauge biopsy needles. Histology was assessed by experienced academic liver histopathologists blinded to the MRI, ELF and elastography findings. Biopsies that were less than 15mm in length or that contained fewer than 11 portal tracts were regarded as inadequate for histological assessment and were therefore excluded.[30, 41] Fibrosis and steatosis were staged according to the system described by Kleiner et al in 2005 [25, 188] and siderosis according the Scheuer grading system.[189]

Biopsies were categorised as NASH based on the presence of lobular inflammation and hepatocyte ballooning.[230] Overall disease activity was graded according to the NAFLD activity score (NAS).[25] Biopsy sections were also stained with Picro Sirius Red and morphometry used to determine the collagen proportionate area (CPA, %) as previously described.[26]

4.5.4 Statistical analysis

Statistical analysis was carried out using IBM SPSS Statistics for Windows version 22 (IBM Corp, Armonk, NY). Variables are summarised with mean ± standard deviation (SD) if normally distributed and with median and range if not normally distributed. Comparisons between patients and healthy volunteers were performed using independent samples t-tests,

Mann-Whitney tests, or Fisher's exact tests, as applicable. Correlation between continuous variables was determined with Spearman's correlation coefficient (Rho). Comparisons across variables with multiple groups were performed using Kruskal-Wallis tests for nominal variables, or Jonckheere–Terpstra tests for ordinal variables. Post-hoc pairwise comparisons between groups were performed using Dunn's test. Diagnostic performance was compared by calculation of the receiver operating characteristic and determination of the area under the curve (AUROC) with 95% confidence intervals (CI). For all tests, a p-value <0.05 was taken to indicate statistical significance.

4.6 Results

4.6.1 Patient demographics

Of 54 patients with NAFLD recruited into the study 50 had sufficient data for analysis. 3 MRI data sets were unusable and 1 biopsy was judged too small for reliable fibrosis assessment. 7 healthy volunteers were recruited. One volunteer was subsequently excluded from analysis due to the discovery of abnormal liver biochemistry. The study flow chart is shown in Figure 4.6-1.

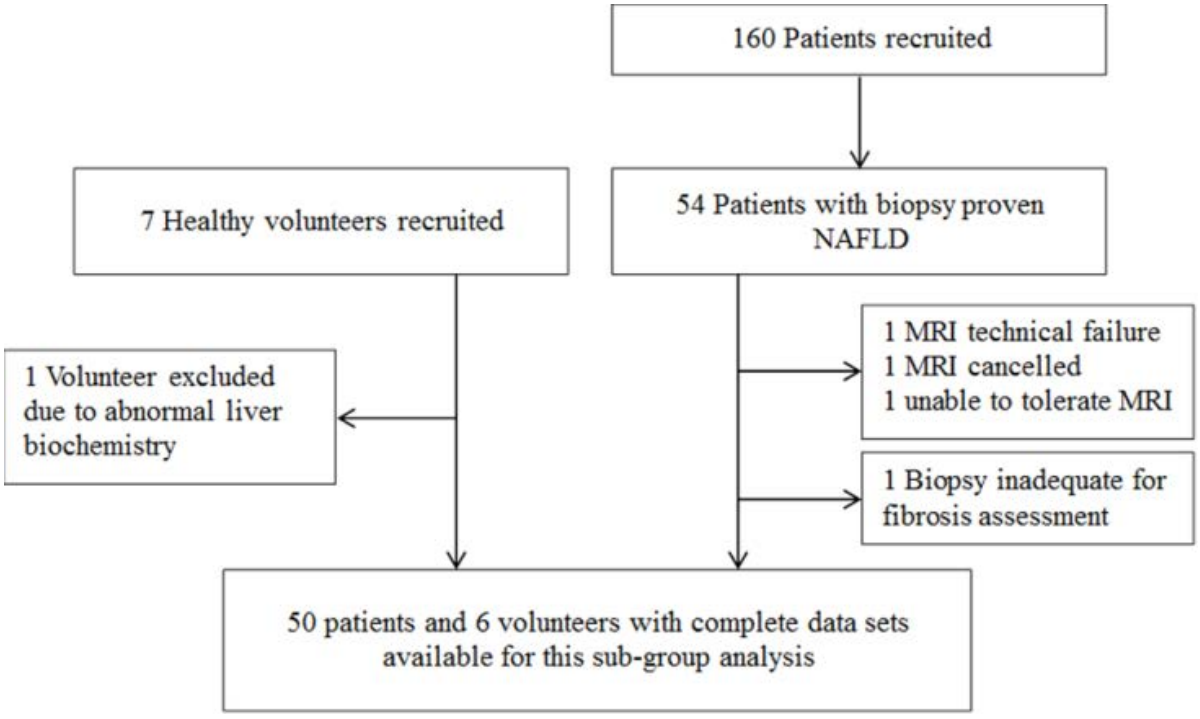


Figure 4.6-1: Study flow chart.

The characteristics of the 50 patients and 6 volunteers are outlined in Table 4.6-1. Comparisons between these groups found that the healthy volunteers were significantly

younger (median 32 vs. 54 years, p=0.011), had significantly lower BMI and lower waist to hip (W:H) ratio. Volunteers were more likely to consume alcohol than patients but there was no difference in the median consumption of drinkers and no patient or volunteer drank alcohol to excess.

49/50 (98%) of patient FibroScan™ examinations were possible (≥10 valid readings) and 47/50 (94%) were reliable by Boursier's criteria.[182] Non-reliable FibroScan™ examinations were excluded from further analysis. 16/47 (34%) of patient FibroScan™ examinations were completed with the M probe and the remainder with the XL probe. All volunteer FibroScan™ examinations were performed with the M probe and were reliable by Boursier's criteria.

	Patients n=50	Healthy Volunteers n=6	p-Value
Age (years)	54 (18-73)	32 (23-55)	**0.011**
Male	28 (56%)	3 (50%)	1.000
Caucasian	43 (86%)	6 (100%)	1.000
BMI (Kg/m^2)	33.6 ±5.1	24.0 ±2.5	**0.001**
W:H ratio			
Male	0.98 ±0.07	0.81 ±0.05	**0.001**
Female	0.90 ±0.06	0.72 ±0.03	**0.001**
Post-transplant	5 (10%)	n/a	-
Type 2 diabetes	26 (52%)	n/a	-
Hypertension	25 (50%)	n/a	-
Dyslipidaemia	26 (52%)	n/a	-
Smoking Status			1.000
Non-smoker	26 (58%)	4 (67%)	-
Ex-smoker	15 (30%)	2 (33%)	-
Current smoker	6 (12%)	0 (0%)	-
Consume alcohol	13 (26%)	6 (100%)	**0.001**
UK units/week*	8 (1-20)	13 (1-15)	0.701

Data reported as mean ±SD, with p-values from t-tests; median (range), with p-values from Mann-Whitney tests; or n (%), with p-values from Fisher's exact tests, as applicable.

Bold p-values are significant at $p<0.05$

**In patients that consume alcohol*

Table 4.6-1: Baseline characteristics of patients with NAFLD and healthy volunteers.

4.6.2 Histology results

The histological characteristics of the patients in the study are shown in Table 4.6-2.

Characteristic	n	%
Kleiner Fibrosis Stage		
0	6	12%
1	10	20%
2	9	18%
3	20	40%
4	5	10%
Diagnosis		
SS	12	24%
NASH	38	76%
Brunt Steatosis Grade		
0	0	0%
1	23	46%
2	17	34%
3	10	20%
Lobular Inflammation (NAS)		
0	11	22%
1	23	46%
2	15	30%
3	1	2%
Hepatocyte Ballooning (NAS)		
0	10	20%
1	15	30%
2	25	50%
Total NAS		
0	0	0%
1-2	9	18%
3-4	16	32%
5-6	22	44%
7-8	3	6%
Scheuer Siderosis Grade		
0	42	84%
1	7	14%
2	1	2%
3	0	0%

Table 4.6-2: Liver histology characteristics of study participants

Median (range) length of liver biopsy samples was 25 (15-50) mm. Median (range) CPA was 5.3 (0.6-34.2) %. CPA correlated strongly with Kleiner fibrosis stage ($p<0.001$) (Figure 4.6-2). The characteristics of the histology and distribution of fibrosis stages in the cohort are shown in Table 4.6-2. 12 (24%) of the patients had SS and 38 (76%) had NASH.

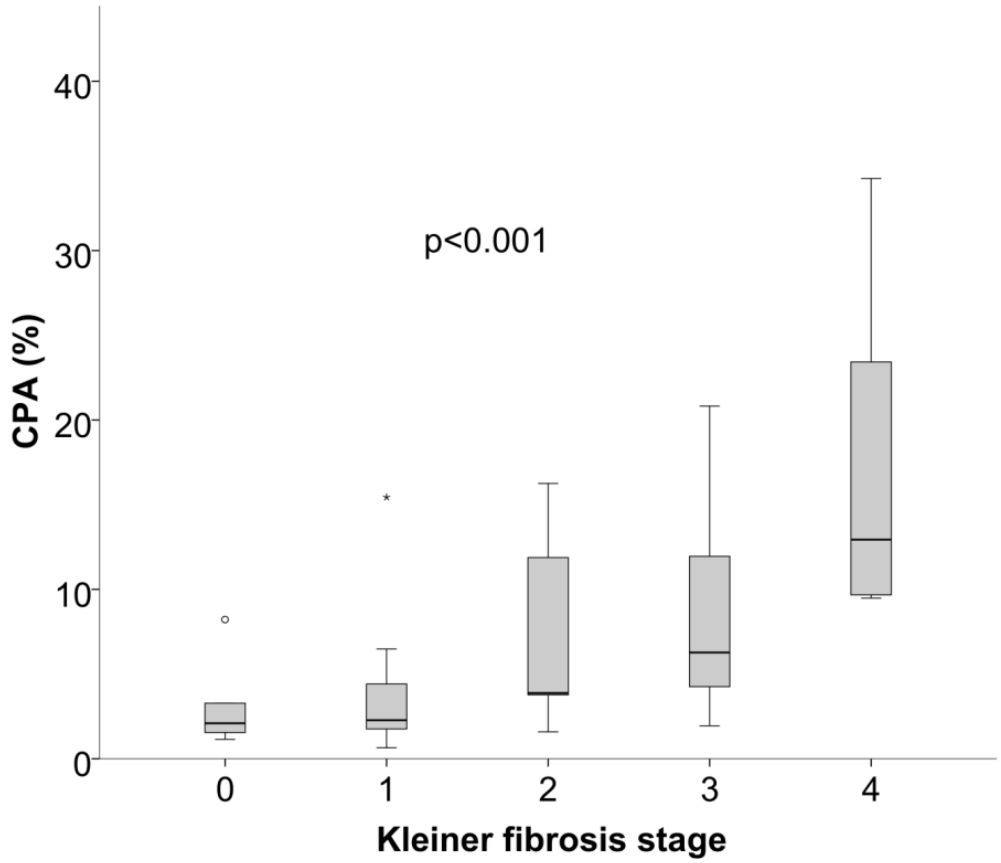


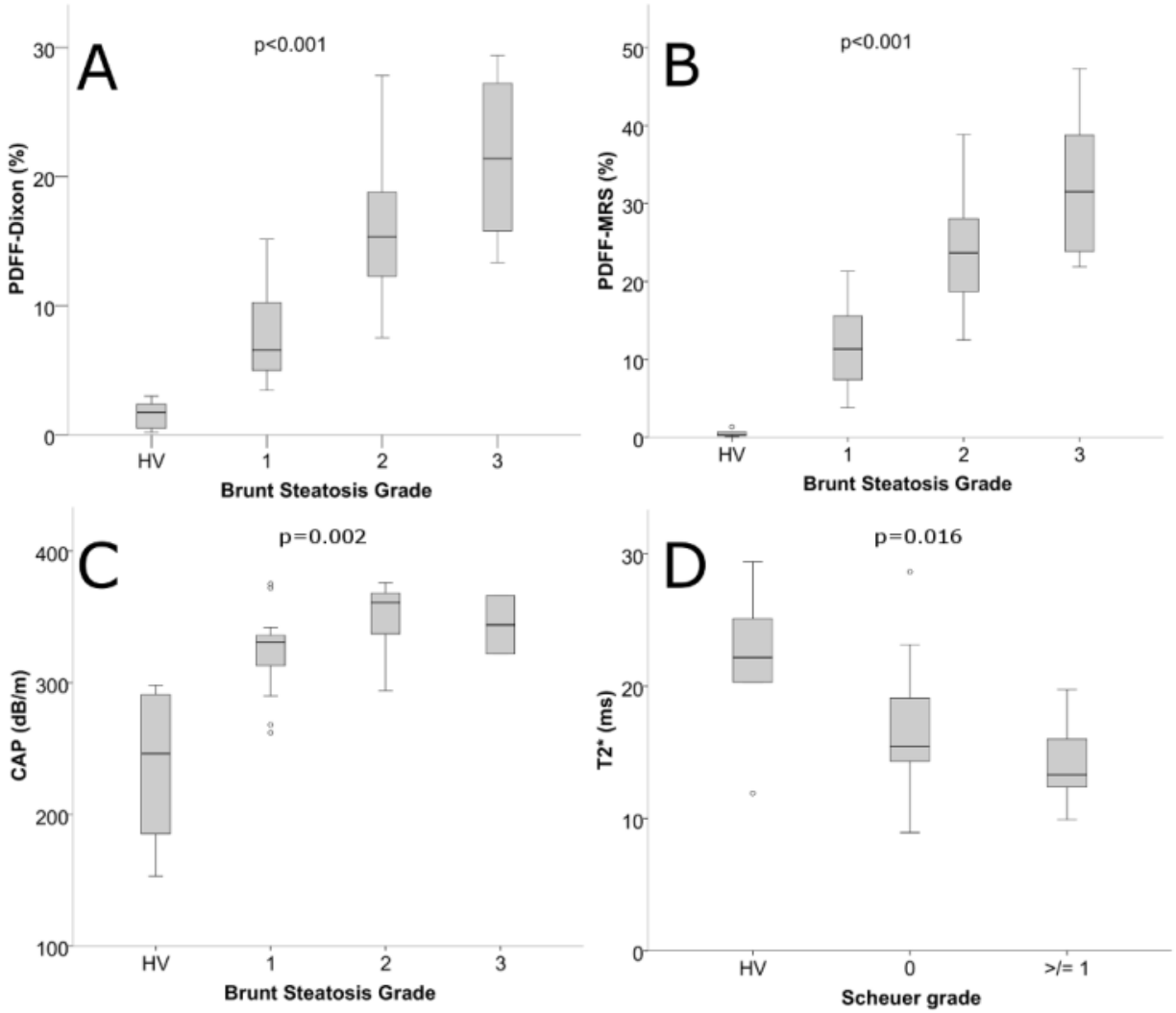
Figure 4.6-2: Box plot showing the association of CPA and Kleiner stage in the study cohort. $P<0.001$ by the Jonckheere-Terpstra test.

4.6.3 Grading of steatosis

PDFF-Dixon data was available for 38/50 (76%) patients and all volunteers. Median PDFF-Dixon for volunteers, grade 1, grade 2 and grade 3 steatosis were 1.8, 6.6, 15.3 and 21.4% respectively ($p<0.001$) (Figure 4.6-3 A). PDFF-MRS was available for 43/50 (86%) patients and 5/6 (83%) volunteers. Median PDFF-MRS for volunteers, grade 1, grade 2 and grade 3 steatosis were 0.3, 11.3, 23.7, and 31.5% respectively ($p<0.001$) (Figure 4.6-3 B). AUROC (95%CI) for the identification of steatosis (Brunt grade ≥1) for both PDFF-Dixon and PDFF-MRS was 1.00 (1.00-1.00).

CAP was available in 24/50 (48%) patients and all volunteers. Median CAP for volunteers, grade 1, grade 2 and grade 3 steatosis were 246, 331, 361, 344dB/m respectively ($p=0.002$) (Figure 4.6-3 C). AUROC (95% CI) for CAP for the identification of steatosis (Brunt grade ≥1) was 0.95 (0.87-1.00).

There was a strong correlation between PDFF-Dixon and PDFF-MRS ($n=38$, $Rho=0.975$, $p<0.001$), between CAP and PDFF-MRS ($n=27$, $Rho=0.682$, $p<0.001$) and between CAP and PDFF-Dixon ($n=26$, $Rho=0.712$, $p<0.001$,).



Comparison	Brunt vs. PDFF-Dixon	Brunt vs. PDFF-MRS	Brunt vs. CAP
Overall	**<0.001**	**<0.001**	**0.002**
HV vs. 1	0.211	0.312	0.090
HV vs. 2	**<0.001**	**<0.001**	**<0.001**
HV vs. 3	**<0.001**	**<0.001**	0.277
1 vs. 2	**0.011**	**0.004**	0.376
1 vs. 3	**0.002**	**<0.001**	1.000
2 vs. 3	1.000	1.000	1.000

Comparison	Scheuer vs. T2
Overall	**0.016**
1 vs. 0	0.267
1 vs. HV	**0.012**
0 vs. HV	0.116

Figure 4.6-3: Box plots demonstrating the relationships between A) PDFF-Dixon and Brunt steatosis grade, B) PDFF-MRS and Brunt steatosis grade, C) CAP and Brunt steatosis grade and D) T2* and Scheuer siderosis grade. Overall significance calculated with the Kruskal-Wallis test and inter-group differences assessed with Dunn's tests.

4.6.4 Grading of siderosis

7/50 (14%) patients had grade 1 siderosis on biopsy and only 1/50 (2%) patient had grade 2 siderosis. Mean T2* in healthy volunteers, patients without siderosis on biopsy and patients with siderosis on biopsy (Scheuer grade ≥1) had mean T2* of 21.8 (±5.8), 16.7 (±3.7) and 14.1 (±3.1) milliseconds (ms) respectively (p=0.016) (Figure 4.6-3 D). AUROC for the differentiation of patients with and patients without siderosis on biopsy was 0.705 (0.498-0.912). Median serum ferritin for healthy volunteers, patients without siderosis and patients with siderosis was 52, 119 and 238 μg/L respectively (p=0.003). Serum ferritin significantly correlated with T2* (Rho=-0.374, p=0.005). Serum iron did not show a statistically significant relationship with histological assessment of siderosis (p=0.059) or correlation with T2* (Rho=0.029, p=0.834).

4.6.5 Differentiation between NASH and simple steatosis

Demographic characteristics and blood results presented in Table 4.6-3 showed no significant difference between patients with NASH and SS. cT1 showed a significant difference between SS and NASH and, although not validated for this purpose, LS and ELF also showed significant differences between patients with SS and those with NASH (Table 4.6-4). Whilst cT1 did differentiate between NASH and SS, the AUROC (95% CI) for cT1 0.69 (0.50-0.88) was inferior to ELF 0.87 (0.77-0.79) and LS 0.82 (0.70-0.94) (Figure 4.6-4).

	NASH (n=38)	Simple Steatosis (n=12)	p-Value
Age (years)	54 (18-73)	46 (23-69)	0.216
Male	19 (50%)	9 (75%)	0.186
Caucasian	34 (90%)	9 (75%)	0.337
BMI (Kg/m^2)	34.2 ±4.8	31.6 ±5.7	0.125
Type 2 diabetes	22 (58%)	4 (33%)	0.190
Hypertension	20 (53%)	5 (42%)	0.742
Hyperlipidaemia	21 (55%)	5 (42%)	0.514
Consume alcohol	8 (21%)	5 (42%)	0.256
*Alcohol intake (UK units/week)**	*7 (1-20)*	*12 (2-16)*	*0.831*
Bilirubin (µmol/L)	11 (4-45)	15 (5-50)	0.318
Aspartate transaminase (AST) (U/L)	38 (19-119)	41 (16-112)	0.526
Alanine transaminase (ALT) (U/L)	53 (18-153)	74 (15-176)	0.707
Alkaline phosphatase (ALP) (U/L)	94 (45-251)	76 (50-149)	0.114
Gamma-glutamyl transferase (gGT) (U/L)	78 (22-381)	58 (21-547)	0.071
Albumin (g/L)	45 ±4	46 ±4	0.477
Fasting glucose (mmol/L)	6.3 (2.8-17.3)	5.6 (4.6-11.4)	0.111
Cholesterol (mmol/L)	4.8 ±1.5	5.2 ±1.2	0.470
Triglycerides (mmol/L)	2.0 (0.7-5.8)	1.4 (0.9-2.7)	0.099
Ferritin (µg/L)	115 (10-689)	177 (45-346)	0.159
Transferrin saturation (%)	24.3 (7.5-49.6)	30.5 (14.7-43.7)	0.080
Creatinine (µmol/L)	73 (46-143)	77 (52-97)	0.225
Platelet count ($x10^9/L$)	219 ±71	197 ±37	0.172

Data reported as mean ±SD, with p-values from t-tests; median (range), with p-values from Mann-Whitney tests; or n (%), with p-values from Fisher's exact tests, as applicable.
**In patients that consume alcohol*

Table 4.6-3: Demographic, clinical and laboratory parameters in patients with simple steatosis and NASH.

	NASH (n=38)	SS (n=12)	p-Value
cT1 (ms)	1007 ±94	907 ±120	**0.004**
Liver Stiffness (kPa)*	10.2 (4.9-27.7)	6.1 (3.6-9.1)	**<0.001**
ELF	9.3 ±1.0	7.8 ±0.8	**<0.001**
AST:ALT ratio	0.76 (0.27-1.56)	0.62 (0.34-1.07)	0.077
NFS	-0.95 ±1.64	-1.78 ±1.99	0.150
FIB-4	1.20 (0.40-5.80)	1.12 (0.38-4.61)	0.351

Data reported as mean ±SD, with p-values from t-tests or median (range), with p-values from Mann-Whitney tests, as applicable.
Bold p-values are significant at p<0.05
**Based on reliable scans only (n=47)*

Table 4.6-4: cT1, LS and ELF showed significant differences between those with simple steatosis and those with NASH.

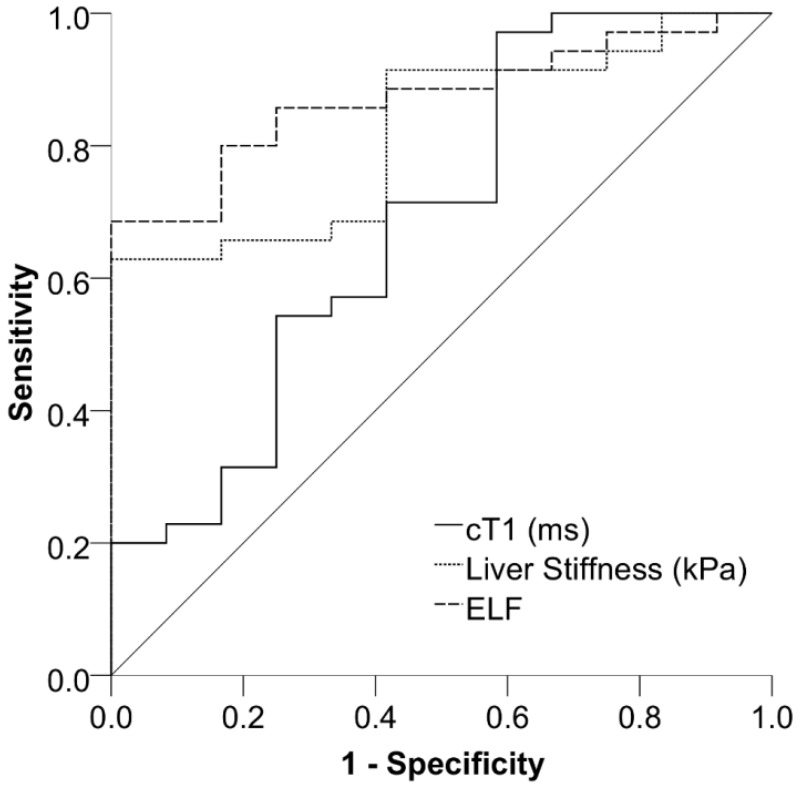


Figure 4.6-4: ROC curve for the differentiation of patients with simple steatosis from those with NASH (n=47).

4.6.6 Grading of NAFLD disease activity

In patients with NAFLD, semi-quantitative assessment of hepatocyte ballooning showed a statistically significant correlation with cT1 ($p=0.045$), LS ($p=0.002$), and ELF ($p=0.011$). Lobular inflammation was significantly associated with LS ($p=0.005$) and ELF ($p=0.001$) but not cT1 ($p=0.588$). Overall assessment of disease activity as defined by the total NAS showed significant correlation with cT1, LS and ELF score with $p<0.001$, $p=0.005$ and $p=0.001$ respectively. These associations can be seen in Figure 4.6-5. AUROC (95% CI) to differentiate those with NAS <5 and NAS ≥5 was statistically significant for cT1, LS, ELF and FIB-4, 0.74 (0.59-0.88), 0.74 (0.59-0.89), 0.74 (0.59-0.89) and 0.73 (0.58-0.88) respectively. Statistical significance was not reached by AST:ALT ratio and NFS 0.60 (0.43-0.77) and 0.63 (0.47-0.77) respectively.

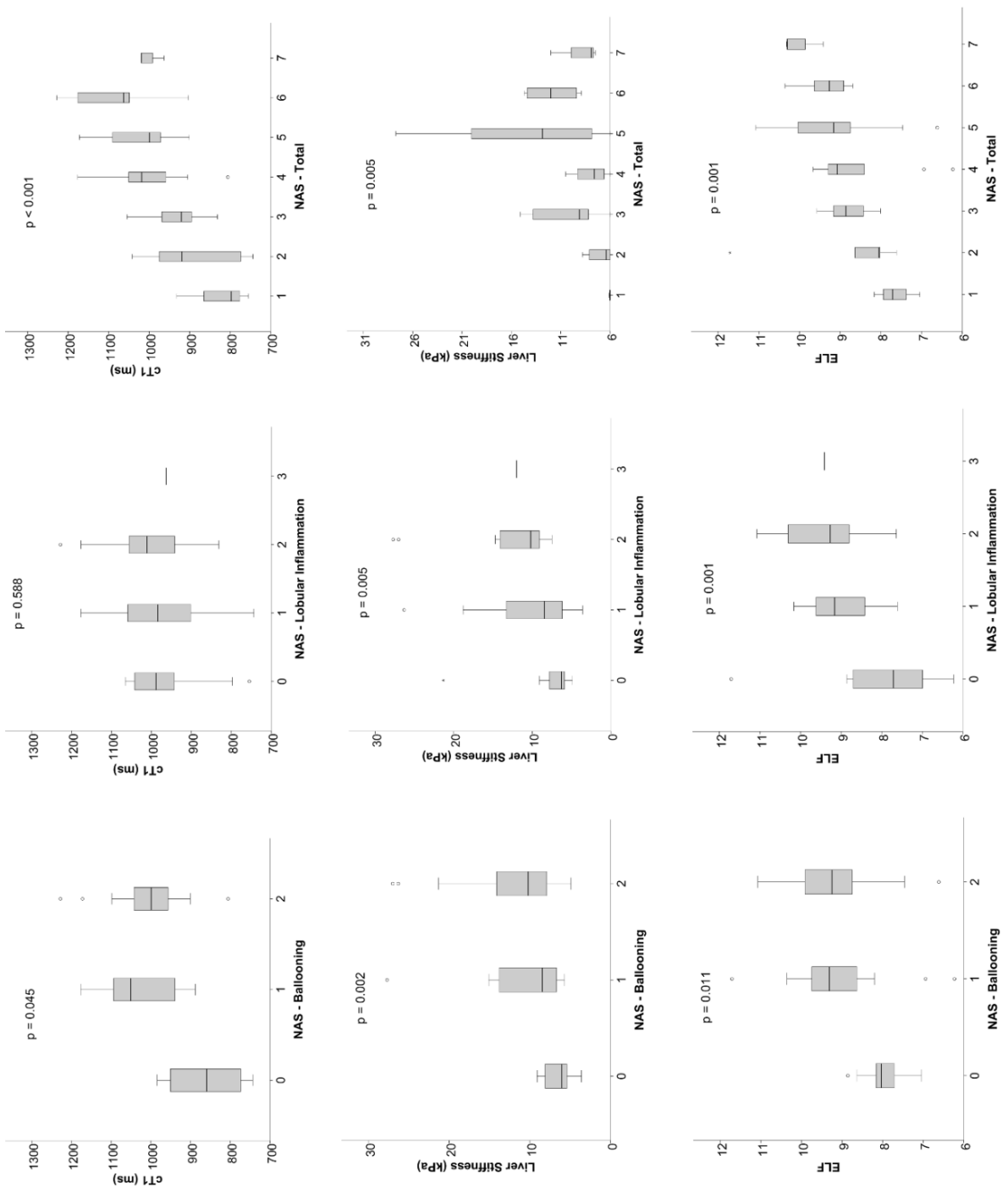


Figure 4.6-5: Box plots showing the relationship between the individual components of NAS and non-invasive markers of liver disease. P-values calculated with the Jonckheere–Terpstra test.

4.6.7 Staging of liver fibrosis using multiparametric MRI

Mean (±SD) cT1 for healthy volunteers was 791 (±42)ms. For patients with NAFLD with F0, F1, F2, F3 and F4 fibrosis mean (±SD) cT1 was 882 (±141), 969 (±115), 985(±93), 1016 (±97) and 997 (±86)ms respectively as can be seen in Figure 4.6-6. Statistically significant differences were demonstrated between healthy volunteers and F2 fibrosis (p=0.048) and F3 fibrosis (p=0.003). However, cT1 showed no significant trend across the fibrosis stages (p=0.068), with pairwise comparisons finding no evidence of significant differences between individual fibrosis stages in patients with NAFLD. As shown in Figure 4.6-7, there was no evidence of significant correlation between cT1 and CPA in patients with NAFLD (Rho=0.142, p=0.324).

In NAFLD patients there was a significant association between Kleiner fibrosis stage and ELF (p<0.001), Liver stiffness (LS) (n=47) (p<0.001), NFS (p=0.003), AST/ALT ratio (p=0.002) and FIB-4 (p=0.013) (Figure 4.6-6). CPA showed significant correlation with: ELF (Rho=0.404, p=0.004), LS (n=47) (Rho=0.511, p<0.001), NFS (Rho=0.306, p=0.030), AST/ALT ratio (Rho=0.453, p=0.001) and FIB-4 (Rho=0.292, p=0.039) (Figure 4.6-7).

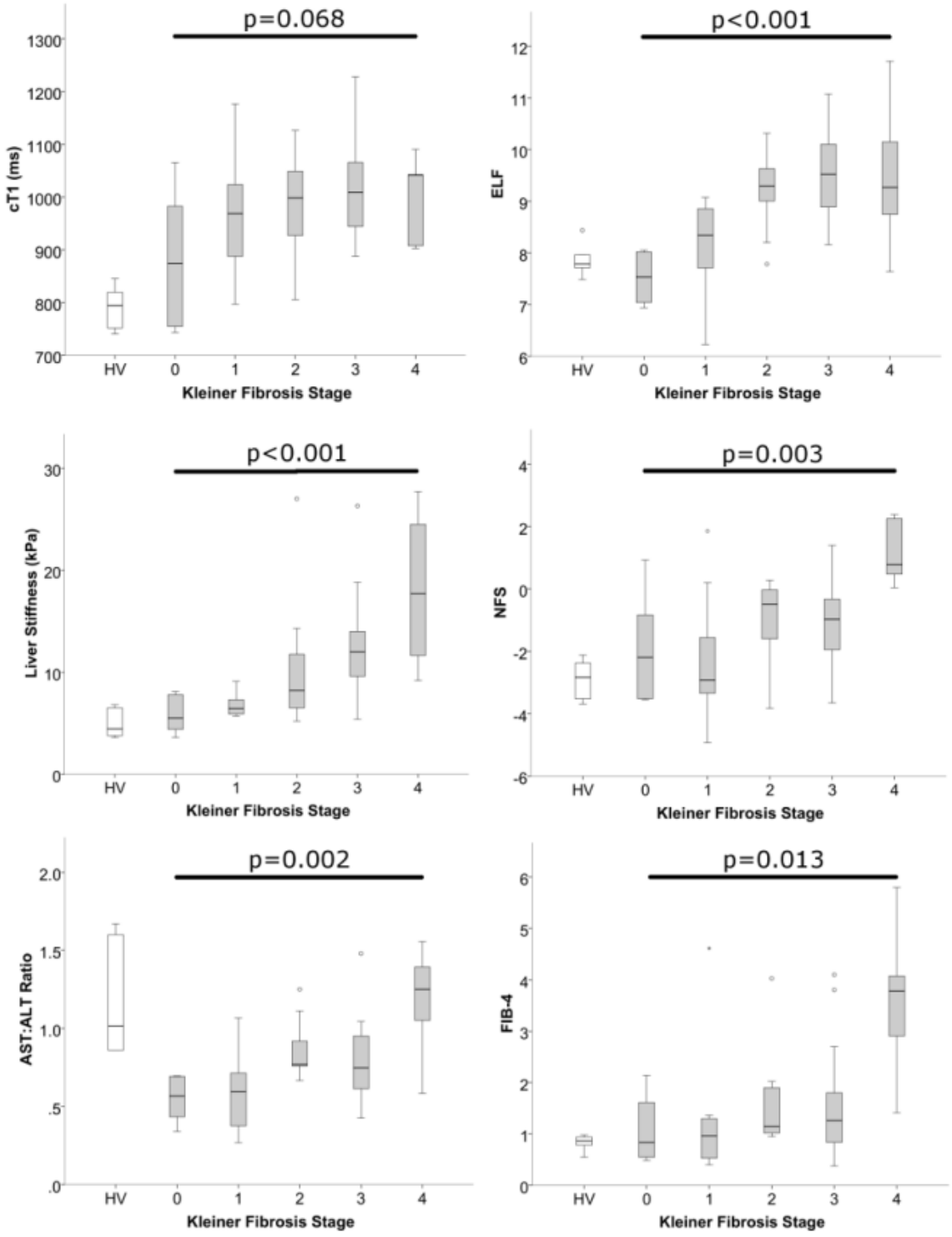


Figure 4.6-6: Box plots showing the relationship between non-invasive markers of liver disease and Kleiner fibrosis stage in the study cohort.

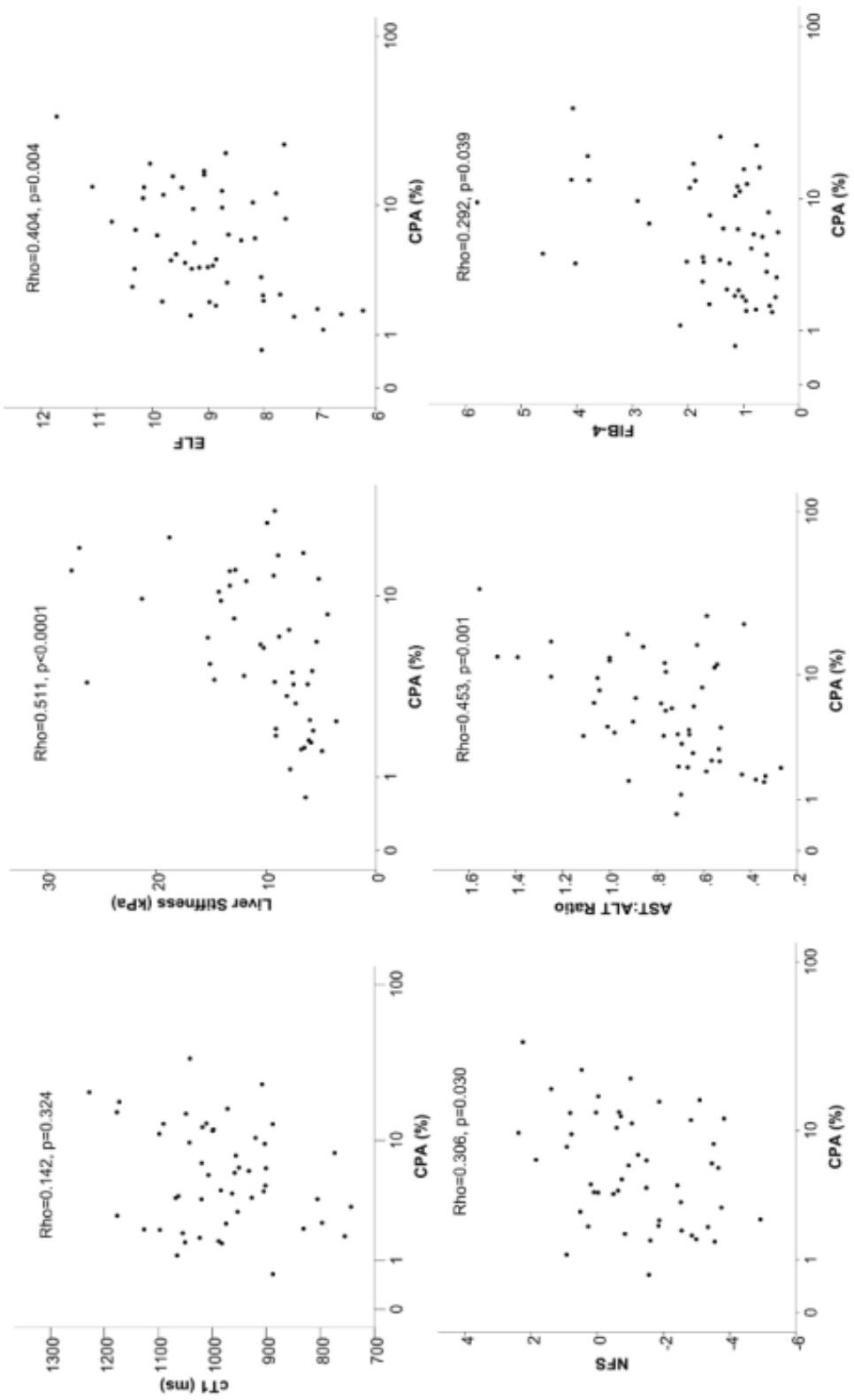


Figure 4.6-7: Scatter plots showing the relationship between non-invasive markers of liver disease and collagen proportionate area (CPA) in the study cohort.

To diagnose clinically significant (defined as ≥F2) fibrosis in patients with NAFLD, the AUROC (95% CI) for ELF, LS, AST:ALT ratio, NFS and FIB-4 were statistically significant; 0.90 (0.82-0.99), 0.90 (0.81-0.99), 0.78 (0.64-0.93), 0.72 (0.54-0.89) and 0.69 (0.52-0.86) respectively. cT1 did not reach statistical significance with AUROC (95%CI) of 0.63 (0.45-0.81).

To diagnose advanced (defined as ≥F3) fibrosis in patients with NAFLD, AUROC (95% CI) for LS, ELF, NFS were statistical significance; 0.88 (0.76-0.99), 0.80 (0.68-0.93) and 0.66 (0.50-0.82) respectively. AST:ALT ratio, cT1 and FIB-4 did not reach statistical significance with AUROC (95%CI) of 0.63 (0.47-0.79), 0.62 (0.46-0.78) and 0.61 (0.45-0.78) respectively.

It is proposed that the influence of inflammation on cT1 leads to the lack of correlation with fibrosis stage. Figure 4.6-8 shows the influence of disease activity on assessment of fibrosis. In patients with early stage fibrosis there was significant difference in cT1 between those with NAS <5 compared to those with a NAS ≥5. This distinction had a trend towards significance in higher stage fibrosis.

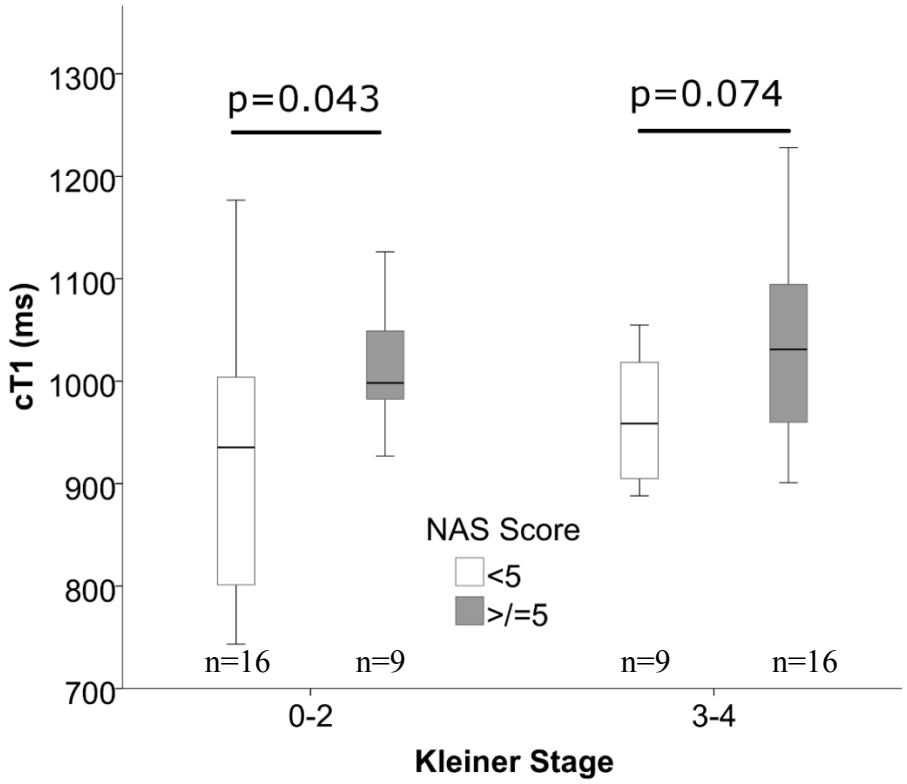


Figure 4.6-8: Box plot demonstrating that cT1 is elevated in patients with a high NAS implying that cT1 reflects NAFLD disease activity as well as fibrosis. This was statistically significant by the Mann-Whitney U test in those with early stage fibrosis only.

4.6.8 Comparative utility of cT1 to exclude clinically significant liver disease

10/50 (20%) of patients in the cohort were classified as being at low risk for progressive liver disease. This was defined as SS without clinically significant (>F1) fibrosis. cT1, LS, ELF, AST:ALT ratio and NFS showed statistically significant differences between healthy volunteers, low risk patients and high risk patients as shown in Table 4.6-5.

	High risk patients* (n=40)	Low risk Patients** (n=10)	Healthy volunteers (n=6)	p-Value
cT1 (ms)	1007 ±93	890 ±122	790 ±42	**<0.001**
Liver Stiffness (kPa)†	9.9 (4.9-27.7)	6.1 (3.6-9.1)	4.5 (3.6-6.8)	**<0.001**
ELF	9.2 ±1.0	7.7 ±0.8	7.9 ±0.3	**<0.001**
AST:ALT ratio	0.76 (0.27-1.56)	0.65 (0.34-1.07)	1.02 (0.86-1.67)	**0.018**
NFS	-1.05 ±1.66	-1.54 ±2.09	-2.90 ±0.62	**0.047**
FIB-4	1.15 (0.38-5.80)	1.23 (0.53-4.61)	0.87 (0.55-0.98)	0.158

Data reported as mean ±SD, with p-values from one-way ANOVA or median (range), with p-values from Kruskal-Wallis tests, as applicable. Bold p-values are significant at p<0.05
** Patients with either NASH or >F1 fibrosis ** Patients with SS and ≤F1 fibrosis*
† Based on reliable scans only (n=53)

Table 4.6-5: cT1, LS, ELF, AST:ALT ratio and NFS show significant differences between high risk patients, low risk patients and healthy volunteers.

The AUROC (95% CI) to differentiate the different groups is shown inTable 4.6-6, and confirmed effective utility of cT1, LS and ELF to exclude significant liver disease.

	Low risk patients* (n=10) *vs* High risk patients** (n=37)	Healthy volunteers (n=6) *vs* All patients (n=47)	Healthy volunteers and low risk patients (n=16) *vs* High risk patients (n=37)
cT1	**0.73 (0.53-0.93)**	**0.93 (0.86-1.00)**	**0.83 (0.69-0.96)**
LS†	**0.82 (0.69-0.94)**	**0.89 (0.77-1.00)**	**0.86 (0.76-0.96)**
ELF	**0.89 (0.80-0.99)**	**0.81 (0.69-0.92)**	**0.89 (0.81-0.98)**
AST:ALT	0.64 (0.45-0.84)	**0.82 (0.67-0.97)**††	0.52 (0.34-0.70)††
NFS	0.55 (0.32-0.77)	**0.79 (0.66-0.91)**	0.64 (0.47-0.81)
FIB-4	0.51 (0.31-0.71)	0.72 (0.59-0.85)	0.59 (0.43-0.75)

Bold values were significant at p<0.05
** Patients with SS and ≤F1 fibrosis ** Patients with either NASH or >F1 fibrosis*
† Based on reliable scans only (n=53)
††Inverse relationship, i.e. AST:ALT ratio was higher in the healthy volunteers group

Table 4.6-6: AUROC (95% CI) for stratification of low and high risk patients. cT1, LS and ELF showed statistically significant results across all comparisons.

Taking common cut off values for the three best performing tests, sensitivity, specificity, negative predictive value and positive predictive value for the diagnosis of high risk patients were calculated and are shown in Table 4.6-7.Negative predictive values, suggesting those patients for whom biopsy could potentially be avoided, were substantially higher for cT1 (80.0-83.3%) compared to LS (39.1-42.9%) and ELF (26.3-57.1%). However, attention must be paid to the disconnection between NPV and sensitivity in this cohort due to the high prevalence of high risk patients in the study cohort. This is discussed further in section 4.7.

		AUROC (95% CI)	**Cut off**	**Sensitivity**	**Specificity**	**PPV**	**NPV**
Low* vs high risk** patients	cT1	0.73	822 ms[163]	97.5%	40.0%	86.7%	80.0%
		(0.53-0.93)	875 ms[165]	97.5%	50.0%	88.6%	83.3%
	LS	0.82	5.8 kPa[103]	89.2%	30.0%	82.5%	42.9%
		(0.69-0.94)	7.0 kPa[103]	75.7%	60.0%	87.5%	40.0%
			7.9kPa[103]	64.9%	70.0%	88.9%	35.0%
			9.0 kPa[103]	62.2%	90.0%	95.8%	39.1%
	ELF	0.89	7.7[62]	92.5%	40.0%	86.0%	57.1%
		(0.80-0.99)	9.8[62]	30.0%	100%	100%	26.3%
HV vs patients	cT1	0.93	822 ms	90.0%	83.3%	97.8%	50.0%
		(0.86-1.00)	875 ms	88.0%	100.0%	100%	50.0%
	LS	0.89	5.8 kPa	85.1%	66.7%	95.2%	36.4%
		(0.77-1.00)	7.0 kPa	68.1%	100%	100%	28.6%
	ELF	0.81	7.7	86.0%	16.7%	89.6%	12.5%
		(0.69-0.92)	9.8	24.0%	100%	100%	13.6%

** Patients with SS and ≤F1 fibrosis*

*** Patients with either NASH or >F1 fibrosis*

HV: healthy volunteers

Table 4.6-7: Sensitivity, specificity, PPV and NPV at commonly accepted cut off values for the differentiation of low and high risk patients.

4.7 Discussion

The global burden of NAFLD is increasing inexorably and validated non-invasive diagnostic tests are important for patients, clinicians and industry. This is not only the first independent validation study to assess the diagnostic accuracy of multiparametric MRI with Liver*Multiscan*^{TM} in NAFLD, but also the first study to compare the performance and potential cost-effectiveness of this emerging methodology against more established non-invasive biomarkers of liver disease. In our prospectively recruited population we demonstrated the ability of multiparametric MRI to grade hepatic steatosis with a high degree of accuracy. Moreover, multiparametric MRI demonstrated accurate differentiation of patients with simple steatosis from those with NASH and also correlated in a highly significant manner with overall disease activity as defined by NAFLD activity score. However, in this cohort, multiparametric MRI did not predict the severity of histological liver fibrosis. Identifying those patients with NAFLD requires accurate detection of steatosis. In clinical practice, steatosis is typically assessed by visual grading of standard liver ultrasound images.[32] Although the sensitivity of ultrasound in detecting moderate and severe steatosis is good, there is wide interobserver and intraobserver variability.[231] Other non-invasive techniques for steatosis assessment such as the Fatty Liver Index have moderate to good accuracy but are significantly confounded by fibrosis stage and are unable to monitor changes in steatosis.[232, 233] PDFF-Dixon has been shown in this study to have excellent accuracy in differentiating patients with steatosis on liver biopsy from healthy volunteers with AUROC of 1.0. PDFF-Dixon also correlated strongly with PDFF-MRS (Rho=0.975, $p<0.001$), which is widely regarded at the most accurate method for non-invasive quantification of liver fat.[228, 234] Comparison of the accuracy of PDFF-Dixon and Controlled Attenuation Parameter for the

detection of steatosis must be made with caution due to the small numbers of patients in this study with both a PDFF-Dixon and Controlled Attenuation Parameter reading. Both techniques had very high accuracy for the detection of any steatosis.

PDFF-Dixon demonstrated a clear, stepwise increase with advancing Brunt steatosis grade in patients with NAFLD suggesting that multiparametric MRI could be used as an accurate method of monitoring steatosis progression and regression and assessing the therapeutic response to lifestyle or drug interventions in the context of clinical trials. In this study, Controlled Attenuation Parameter did not demonstrate the same stepwise increase with Brunt grade observed with PDFF-Dixon measurement.

A single test to reliably exclude NAFLD would be of considerable value in clinical practice. In this study multiparametric MRI showed a high degree of accuracy for differentiating between healthy volunteers and those with NAFLD with AUROC (95% CI) of 0.93 (0.86-1.00). It should be recognised however, that the healthy volunteers and patients enrolled in this study were not well matched in terms of age, waist to hip ratio or BMI. Accepting this limitation, using a cT1 cut-off value of 875ms gave multiparametric MRI a sensitivity of 88.0% with specificity of 100% for the detection of any liver disease. This was superior to all other non-invasive tests. In addition, the negative predictive value for excluding any liver disease was substantially higher than those for the other non-invasive techniques.

Although the NPV is notably higher for cT1 than other tests, there are two key limitations that must be recognised. Firstly the relatively small number of patients included in this study means that small variations in how patients are classified by the various tests has a large

impact on the calculated sensitivity, specificity, NPV and PPV. For example cT1 and liver stiffness at the low cut-off of 5.8kPa have broadly similar sensitivities but for liver stiffness the NPV is half that of cT1. For cT1 the calculation (true negative / all negatives) is 5/6 = 83.3% and for LS 3/7 = 42.9%. These changes in the classification of just one or two patients leads to the large variation in NPV and the disconnect between sensitivity and NPV seen in these data. Secondly, it must be recognised that PPV and NPV are influenced not only by the performance of the test but also the prevalence of the condition within the population being studied. Our study cohort is drawn from patients selected to undergo a liver biopsy as a routine part of their care within tertiary liver units. To reach the point where liver biopsy was requested it is very likely that the clinicians in charge of their care would be suspicious of advanced liver disease and the majority of these patients would already have failed attempts to stratify disease severity with non-invasive tests. For example most patients will have had a high or indeterminate value from the NAFLD fibrosis score and then either an unreliable or elevated reading from FibroScan. This leads to the low prevalence of low risk patients within our study cohort. This highly selected group is very unlikely to be representative of patients with NAFLD in the community, nor is it necessarily representative of the majority of patients with NAFLD seen in secondary care. This will have influenced the evaluation of the test and the applicability of the results; particularly in respect to NPV and PPV.

An interesting opportunity to consider further work to would be to establish the ability of multiparametric MRI to be used as a single one-stop comprehensive MRI examination to screen for significant liver disease in patients with NAFLD. However, to be useful in this regard, multiparametric MRI would need to demonstrate its ability to reliably differentiate low and high risk patients in populations with a much lower prevalence of advanced liver

disease. The data presented in this thesis is unable to support this conclusion due, in part, to the strategy for patient recruitment.

The differentiation of those with NASH from those with simple steatosis is an important distinction in clinical practice as our current understanding of NAFLD recognises NASH as the harbinger of progressive fibrosis and hepatocellular carcinoma.[202, 203, 235] Identifying individuals with NASH stratifies patients at risk of significant disease and may do so at an earlier stage than tests that reflect fibrosis alone.[207] Detection of NASH guides decision making about clinical management and follow-up intensity and identifies patients who may be eligible for recruitment to clinical studies.[210] To date, the differentiation of simple steatosis and NASH has been reliant on liver biopsy. Liver biopsy has low patient acceptability due to its invasiveness and associated risk. Liver biopsy is also prone to sampling error and interobserver variation of histological assessment. These factors reduce the suitability and reliability of liver biopsy for disease stratification in NAFLD. Currently available methods to non-invasively differentiate NASH and simple steatosis are suboptimal. Conventional blood tests and imaging techniques have low accuracy for the differentiation of simple steatosis and NASH.[210]

To determine the severity of disease, cT1 showed a highly significant, positive correlation with NAFLD activity score. The correlation between NAFLD activity score and cT1 was stronger than between NAFLD activity score and any other evaluated test, although the authors acknowledge that these tests are not designed to assess disease activity. This indicates the potential utility of multiparametric MRI as a sensitive diagnostic tool to monitor changes in disease activity. A NAFLD activity score of ≥5 is frequently used as a criterion to enrich

clinical trials with patients with more significant liver disease. The performance of cT1 to make this distinction was comparable to the other non-invasive markers evaluated in this study.

Staging of fibrosis in NAFLD has been clearly shown to predict clinical outcomes [14, 21, 219] and thus is an important part of the assessment of patients with NAFLD in clinical practice and for inclusion in current late stage clinical trials. Additionally both cT1 and ELF have also been reported to have utility in predicting clinical outcomes.[165, 236] In this cohort, cT1 did not predict fibrosis defined by either Kleiner stage or collagen proportionate area whereas the other non-invasive markers assessed in this study performed in line with published work.[237, 238]

The lack of correlation between cT1 and fibrosis in this study was unexpected as previous work by Banerjee *et al*. in unselected patients[163] and by Pavlides *et al*. in patients with NAFLD[225] has shown a clear correlation between cT1 and fibrosis stage. It may be that our study is underpowered to detect this correlation, however, in our study it appears that the influence of disease activity on cT1 has hampered the ability of cT1 to detect stage differences in fibrosis; larger studies are needed to explore this further. Figure 4.6-8 shows the heavy confounding of disease activity on fibrosis assessment in this cohort. When grouped by fibrosis stage, the only statistically significant difference in cT1 is between low and high NAFLD activity score in patients with early stage fibrosis. This identification of a group of patient with early stage fibrosis and less severe disease as graded by NAFLD activity score characterises a group of patients at low risk of progressive liver disease. As shown in Table 4.6-7, cT1 had comparable sensitivity, specificity and positive predictive value to ELF and

liver stiffness for the exclusion of significant disease (NASH or fibrosis ≥F1). As discussed above, the greater negative predictive value of multiparametric MRI is likely to represent the nature of the study cohort rather than the qualities of the test.

A limitation of this study is that the statistical power is limited by recruitment volume, which may have resulted in some of the more subtle associations between variables being missed. Although liver biopsy remains the gold standard, as a comparator it has known limitations of sampling error and interobserver variation.[30, 239] However, in our study we demonstrated that semi-quantitative histology scores correlated strongly with collagen proportionate area and other biomarkers performed as per previous publications. Despite these limitations, our independently collected data confirms the opportunities for new non-invasive biomarkers in liver disease severity assessment.

4.8 Conclusion

In conclusion, multiparametric MRI with Liver*Multiscan*^{TM} has the ability to identify patients with NAFLD and to quantify steatosis and overall disease activity (by NAFLD activity score). Additionally, it stratified between healthy volunteers and patients with NAFLD as well as between patients at low risk and those at high risk of progressive liver disease. When evaluated alongside existing biomarkers, multiparametric MRI was superior for the grading of steatosis, grading of NASH severity, and for excluding disease, but in this cohort was inferior for the staging of fibrosis. The potential application of multiparametric MRI in clinical and research settings thus remains of great interest and further studies in appropriate clinical populations will facilitate improved understanding of the opportunities for applying MR imaging in evaluation of liver disease.

CHAPTER 5: In Vivo Proton Magnetic Resonance Spectroscopy for the Diagnosis and Assessment of Non-Alcoholic Fatty Liver Disease

5.1 Preamble

Insulin resistance and the associated alterations in lipid metabolism are mechanistically important in the development of non-alcoholic fatty liver disease (NAFLD).[240] These metabolic differences lead to an altered lipidome between health, simple steatosis (SS) and non-alcoholic steatohepatitis (NASH).[241] Proton magnetic resonance spectroscopy (^{1}H-MRS) is a technique used widely in biochemistry and medical imaging to analyse the chemical composition of substances or tissues. In vivo, ^{1}H-MRS can identify metabolites present within tissues and thus characterise different tissues. The most common clinical application of ^{1}H-MRS is the characterisation of brain tumours based on their metabolite content.[242] In addition to this established clinical application, ^{1}H-MRS has been shown to characterise lipids within adipose tissue and the liver. Thus, in vivo ^{1}H-MRS could potentially be used as a non-invasive biomarker to identify the lipidomic difference within hepatic tissue that characterise health, SS and NASH.[180, 243-245]

In the work presented here we investigate the feasibility of using in vivo ^{1}H-MRS to characterise the hepatic lipids profile of patients with NAFLD with a view to establishing the potential as a biomarker to detect NAFLD and assess disease severity.

5.2 Rationale for the investigation of ^{1}H-MRS as a biomarker of NAFLD

5.2.1 Lipotoxicity in NAFLD

The term lipotoxicity refers to inflammation and cellular damage within the liver of patients with NAFLD due to the toxic effects of fatty acids and their metabolites; the central driver for which is insulin resistance and a high carbohydrate diet.[246] Lipid metabolites implicated in this process include phosphatidic acid, choline, ceremides and diacylglycerols.[246] Insulin resistance leads to increased lipolysis in adipose tissue and the release of free fatty acids into the circulation. Circulating free fatty acids are then taken up by the liver [247] and de novo generation of lipid within the liver due to dietary carbohydrate excess further contributes to increased flux of free fatty acids into hepatocytes.[240] The increased movement of fatty acids into hepatocytes overwhelms the normal mechanism of fatty acid excretion as very low density lipoprotein. This fatty acid excess generates reactive oxygen species, which contribute to cellular injury, cytokine production and tissue inflammation. In addition free fatty acids and particularly saturated fatty acids can directly activate Toll like receptors leading to activation of the innate immune system, release of pro-inflammatory cytokines and activation of apoptotic pathways.

The lipotoxicity theory of the pathogenesis of NASH outlined in Figure 5.2-1 indicates that the accumulation of triglyceride droplets within hepatocytes is a by-product of the pathological process and is not hepatotoxic per se.[240, 246]

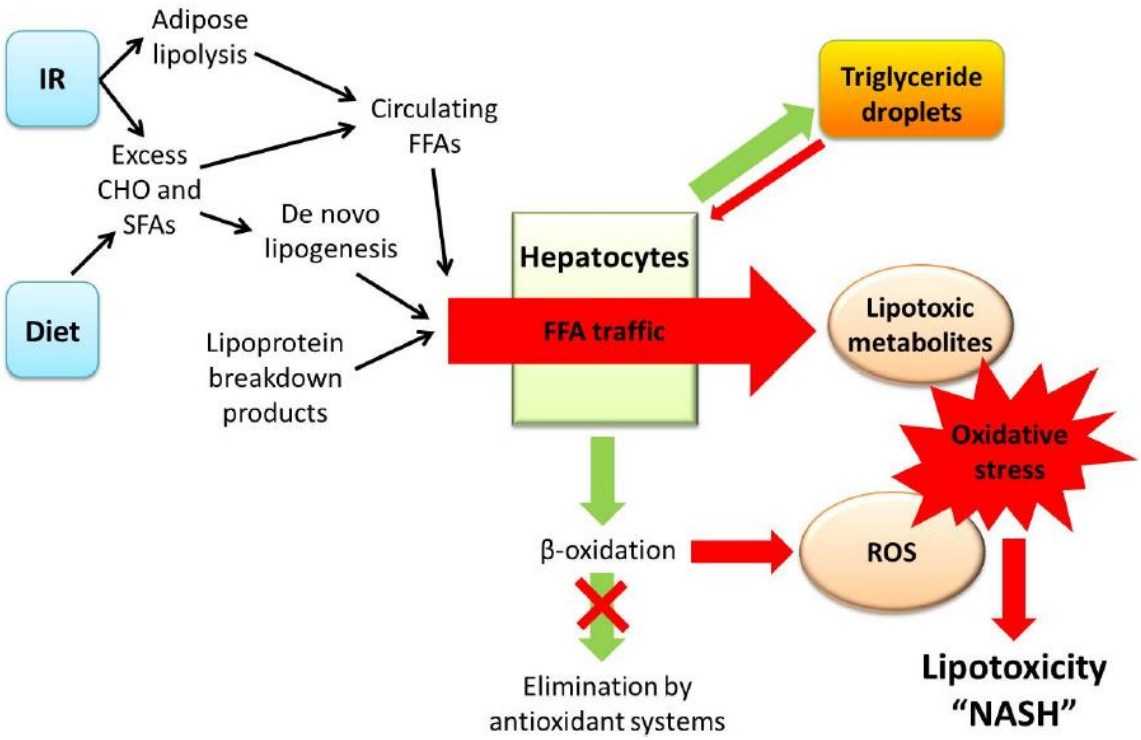


Figure 5.2-1: Schematic representation of the increase in fatty acid uptake by the liver and the associated lipotoxicity in the pathogenesis of NASH. Taken from Peverill W et al, *Int. J. Mol. Sci*, 2014.

5.2.2 Defining the lipidomic differences between health, SS and NASH

With the understanding of the metabolic pathways involved in NAFLD and NASH come attempts to define the metabolic profile of a patient with NAFLD with a view to the identification of diagnostic and prognostic markers. Studies using analytical biochemistry techniques such as mass spectrometry and gas chromatography to analyse liver tissue samples have identified some clear differences between health, SS and NASH.

In 2007 Puri P et al showed that, as might be expected, the total amount of triacylglycerol (TAG) in liver tissues increases from controls to those with NAFLD. Although cholesterol esters (CE) remain fairly constant, free cholesterol shows a stepwise increase from healthy controls to SS to NASH with statistically significantly lower levels of choline in those with NAFLD compared to controls.[241] This work also showed that, in patients with NAFLD, fatty acid chains showed a trend towards being more saturated with a statistically significant decrease in amount of TAG containing poly-unsaturated fatty acid (PUFA) chains. In line with previous work[248] there was also a trend for progressive decrease in PUFAs within the hepatic free fatty acid (FA) pool of patients with NAFLD. However, in this small study (9 patients per group) this change in PUFA concentration did not reach significance between SS and NASH.

The finding that, in NASH, liver tissue contains fatty acids that are more saturated than the liver tissue of controls is supported by Yamada K et al who found a reduction in the C16:1/C16:0 ratio (indicating an increase in the saturation of FAs) that was specifically associated with lobular inflammation, a key feature of NASH.[249] In addition, FA chain length was shown to be shorter in NASH than controls (reduced C18:0/C16:0 ratio). [249]

In 2015 Gorden D and colleagues have presented a very carefully conducted analysis of tissue and plasma metabolomics in a much larger cohort than previously studied and looked specifically at metabolomic biomarkers that could differentiate SS and NASH. This study found that clear differences were evident in both lipid and aqueous metabolites between patients with SS and those with NASH.[250] When looking at the metabolomic analysis of liver tissue the main differences were that the FA chains making up TAG were shorter and more

saturated in NASH than SS. There were also fewer TAGs containing PUFA chains and a reduction in some CE species. 20 lipid species were shown in this paper to differentiate SS and NASH including several phosphatidylcholines (PC), phospholipids of various classes, CEs, ceremides and aqueous metabolites involved in the Krebs cycle. [250]

5.2.3 The identification of lipids with ^{1}H-MRS

^{1}H-MRS is able to non-invasively characterise fatty acids (FAs) in lipid phantom experiments and in vivo adipose tissue[243] and liver.[180, 251, 252] These measurements have been shown to be highly repeatable.[253] Therefore it can be concluded that ^{1}H-MRS can accurately measure several parameters relating to the fatty acid composition of the liver in vivo.

5.3 Aims

As outlined above it is plausible that ^{1}H-MRS will be able to characterise hepatic lipids and that this information can be used to identify the differences that exist in the hepatic lipidome in health, SS and NASH.

To my knowledge, no previous work looks specifically at the differentiation of SS and NASH with ^{1}H-MRS. Here, we investigate the ability of ^{1}H-MRS to characterise hepatic FAs with a view to identifying patients with NAFLD and differentiating those with SS from those with NASH.

To this end the aim is to calculate the following parameters from the ^{1}H-MRS data using the method described in Chapter 2:

- Fat fraction (FF)
- Mean FA chain length (mCL)
- Number of double bonds per FA chain (nDB)
- Number of diacyl groups per chain – indicating polyunsaturated fatty acids (PUFAs)
- Concentration of choline

5.4 Method

Study participants described in Chapter 3 all had ^{1}H-MRS sequences acquired during their research MRI. These ^{1}H-MRS spectra were acquired and analysed as described in Chapter 2 to characterise the hepatic lipidome. Here presented are the data from the patients with biopsy proven NAFLD as well as 6 healthy volunteers and 7 patients with chronic viral hepatitis as controls.

5.5 Results

5.5.1 Participants

54 patients with NAFLD, 7 patients with chronic viral hepatitis and 6 healthy volunteers (HV) underwent ^{1}H-MRS. As shown in Figure 5.5-1, 1 patient with NAFLD was excluded due to an inadequate liver biopsy and 9 were excluded due to the MR spectrum failing visual quality control. One patient with chronic viral hepatitis and one healthy volunteer had an MR spectrum unsuitable for analysis due to very low fat fraction and one patient in the viral hepatitis group had an inadequate biopsy.

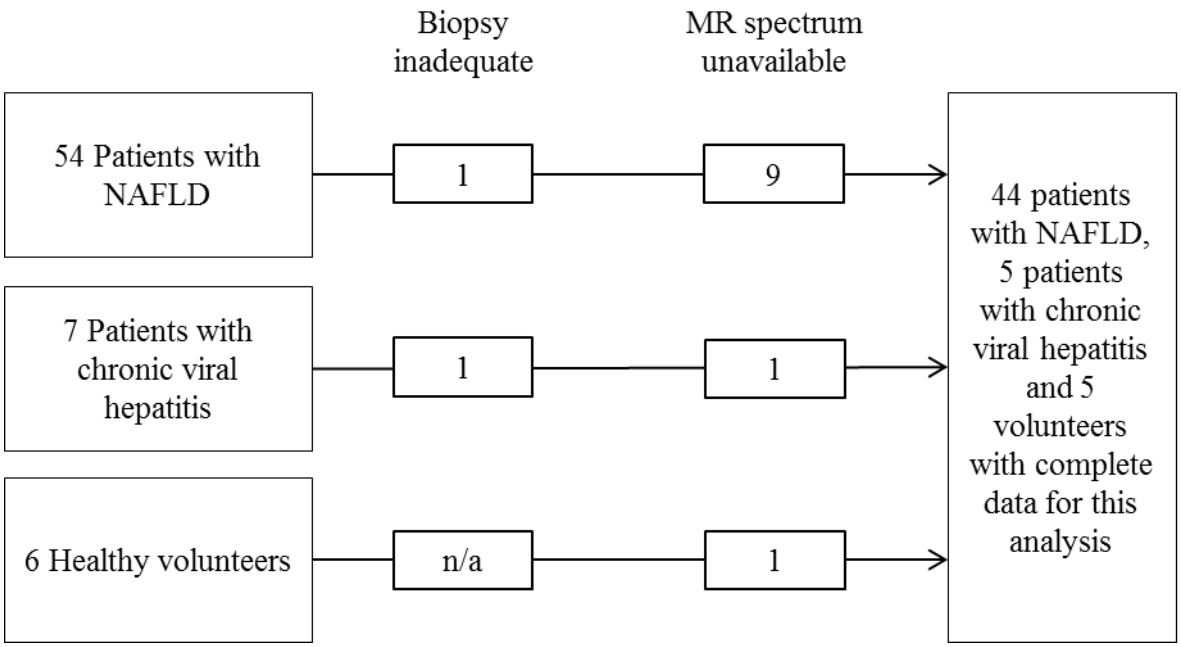


Figure 5.5-1: Study flow chart for the ^{1}H-MRS study

5.5.2 ^{1}H-MRS correlates strongly with histological assessment of steatosis

^{1}H-MRS derived fat fraction correlated strongly with semi-quantitative histological assessment of hepatic steatosis (Brunt grade) with Spearman's Rho of 0.837 ($p<0.001$). Mean (±SD) fat fraction for HVs and Brunt stages in patients are shown in Table 5.5-1 and this relationship is shown in Figure 5.5-2.

	n	**Fat fraction (%)**
HV	5	0.53 (±0.50)
Brunt grade		
0	2 (all viral)	1.88 (±0.25)
1	24 (21 NAFLD, 3 viral)	10.97 (±5.45)
2	14 (all NAFLD)	23.53 (±6.97)
3	9 (all NAFLD)	32.38 (±9.01)

Data reported as mean ±SD.

Table 5.5-1: Fat fraction as measured by ^{1}H-MRS for HVs and each Brunt steatosis grade.

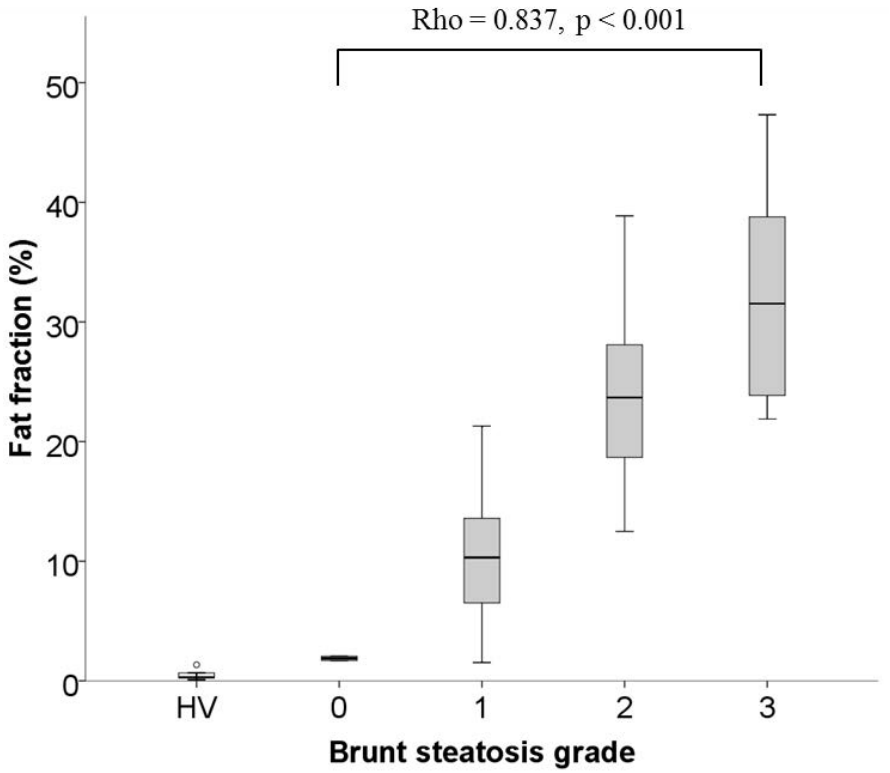


Figure 5.5-2: Box plot demonstrating the strong relationship between Brunt grade and fat fraction.

5.5.3 Differentiation between aetiologies with ^{1}H-MRS

The characteristics of the participants in each group are outlined in Table 5.5-2. There is a statistically significant difference in mean BMI between the groups. However, post-hoc tests show that there is no significant difference between healthy volunteers and patients with viral hepatitis. Healthy volunteers were more likely to drink alcohol than other participants. No participant drank alcohol to excess.

	Patients with NAFLD n=44	Patients with chronic viral hepatitis n=5	Healthy Volunteers n=5	p-value
Age (years)	49.3 ±13.4	48.6 ±15.1	37.6 ±10.3	0.186
Male	26 (59%)	5 (100%)	2 (40%)	0.128
Caucasian	37 (84%)	5 (100%)	5 (100%)	1.000
BMI (Kg/m^2)	33.3 ±5.2	26.3 ±2.5	24.1 ±2.8	**<0.001**
Post-transplant	5 (11%)	2 (40%)	n/a	0.143†
Type 2 diabetes	22 (50%)	1 (20%)	n/a	0.215†
Smoking Status				0.063
Non-smoker	24 (55%)	3 (60%)	4 (80%)	
Ex-smoker	14 (32%)	2 (40%)	1 (20%)	
Current smoker	6 (14%)	0 (0%)	0 (0%)	
Consume alcohol	10 (23%)	0 (0%)	5 (100%)	**0.007**
UK units/week*	6 (1-20)	n/a	12 (1-15)	0.679†
ALT (U/L)	57 (15-488)	40 (19-550)	20 (9-35)	**0.003**
Bilirubin (µmol/L)	12 (4-50)	9 (4-19)	7 (3-21)	0.234
Albumin (g/L)	45 (34-53)	47 (40-50)	49 (44-55)	0.072
Sodium (mmol/L)	140 (134-146)	141 (136-146)	140 (138-141)	0.958
INR	1.0 (0.9-1.4)	1.00 (0.9-1.2)	1.0 (0.9-1.0)	0.261

**In participants that consume alcohol*
† Comparing valid groups only
Data reported as mean ±SD, with p-values from one way ANOVA; median (range), with p-values from Mann-Whitney U or Kruskal-Wallis tests; or n (%), with p-values from Fisher's exact tests, as applicable. Bold p-values are significant at p<0.05

Table 5.5-2: Demographic and baseline characteristics of the three aetiology groups

Fat fraction correlated with BMI (Rho=0.440, p=0.001), waist circumference (Rho=0.494, p<0.001) and waist:hip ratio (Rho=0.504, p<0.001). There was no statistically significant correlation between fat fraction and serum total cholesterol (Rho=0.155, p=0.269) or serum triglycerides (Rho=0.255, p=0.065).

As shown in Figure 5.5-3 there were clear differences in median (range) fat fraction between patients with NAFLD (18.75 (3.79-47.31)% and other groups (viral: 2.06 (1.52-9.84)% and HV 0.31 (0.08-1.34)%). The difference between the fat fraction of healthy volunteers and patients with viral hepatitis was not statistically significant.

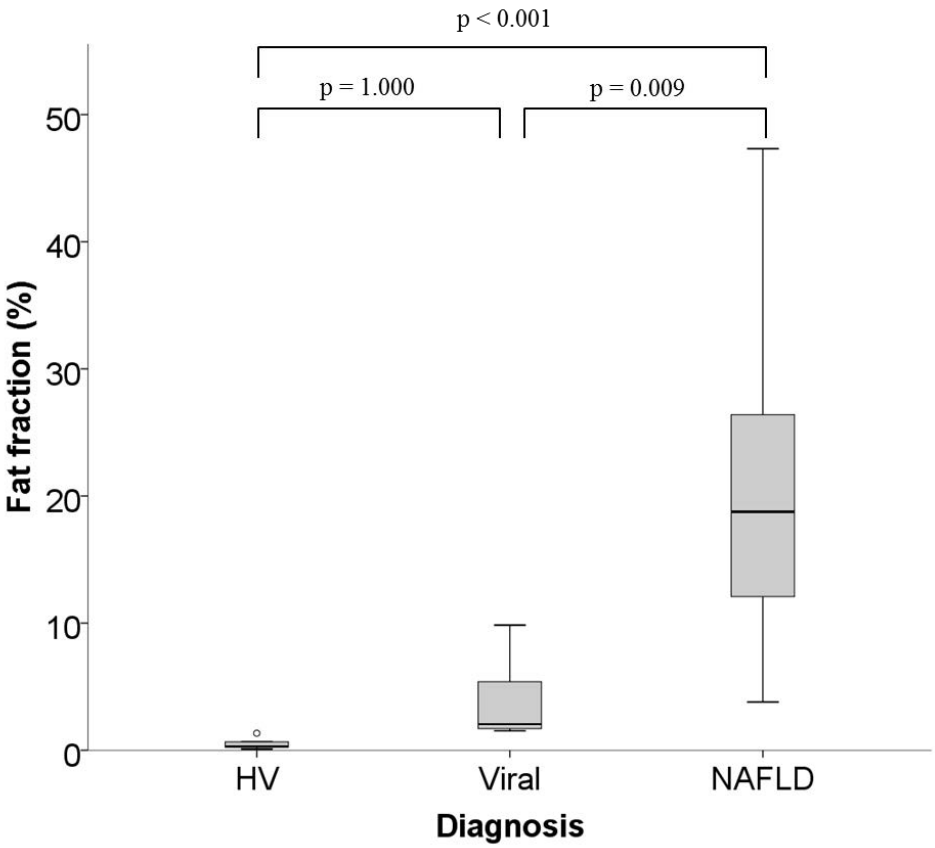


Figure 5.5-3: Box plot demonstrating the relationship between diagnosis and fat fraction. Overall p<0.001 by the Kruskal-Wallis test. Post hoc tests show significant differences between patients with NAFLD and other groups as indicated above.

Although fat fraction can be calculated for HVs the very low fat content makes further analysis of the ^{1}H-MRS spectrum unreliable and it was not possible to further characterise the FAs present. HVs are therefore excluded from analysis of FA characteristics.

In patients with NAFLD compared to those with chronic viral hepatitis there fewer double bonds per chain (1.57 vs 3.94 arbitrary units) and shorter fatty acid chains (25.52 vs 19.89 arbitrary units) but this did not reach statistical significance. In patients with NAFLD there was a statistically significant reduction in the concentration of PUFA (0.00 vs 0.28 arbitrary units) and choline (0.00076 vs 0.018%). This is shown in Figure 5.5-4. None of the calculated fatty acid parameters showed a statistically significant difference between patients with and without type 2 diabetes.

To differentiate chronic viral hepatitis and NAFLD, AUROC (95% CI) for PUFA and choline were 0.88 (0.74-1.00) and 0.88 (0.78-0.98) respectively as shown in
Figure 5.5-5.

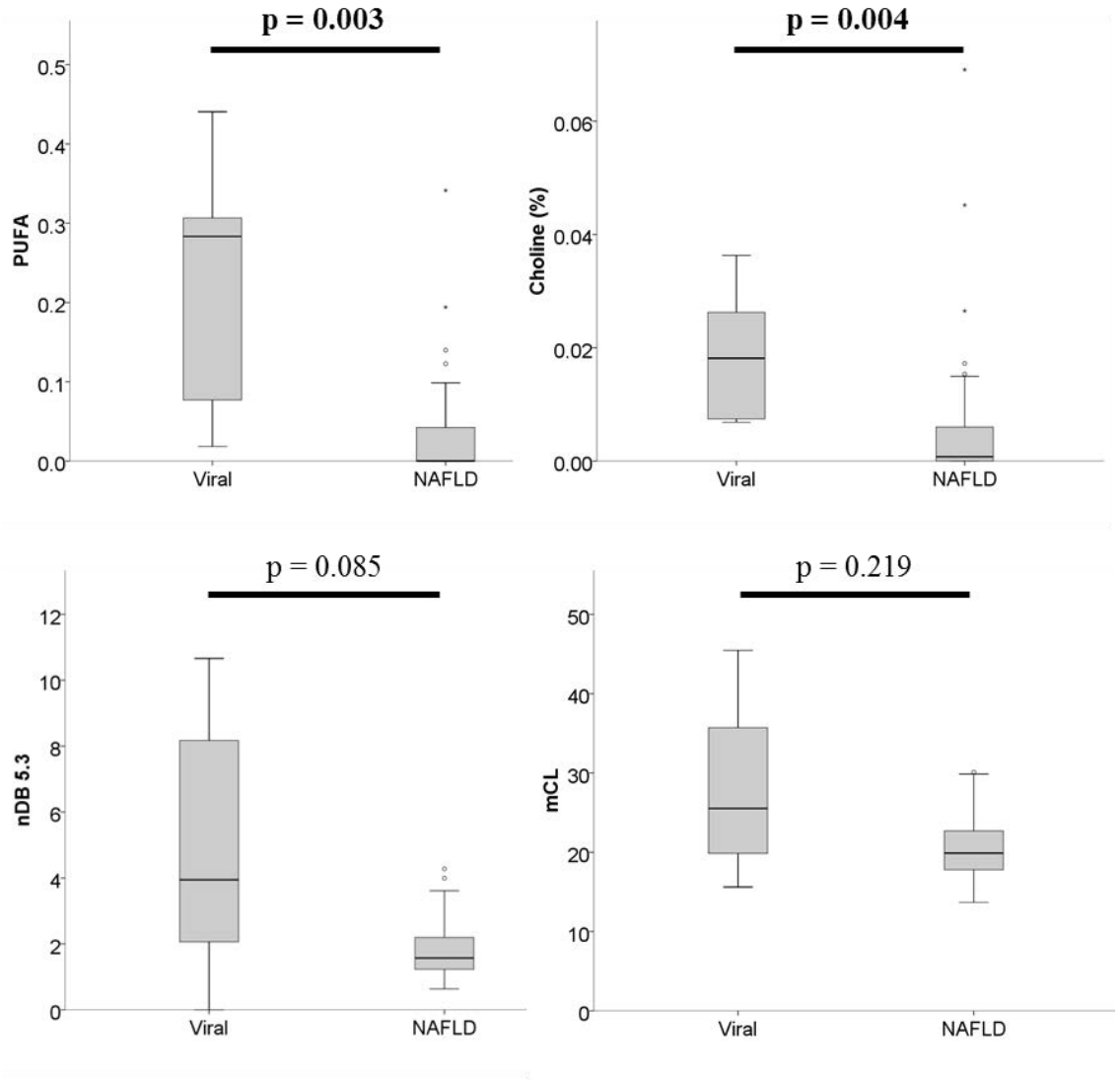


Figure 5.5-4: Box plots demonstrating the relationships between fatty acid characteristics of patients with chronic viral hepatitis and those with NAFLD. Significant differences are seen in the amount of PUFA and the amount of choline. nDB and mCL were non-significantly lower in patients with NAFLD. p-values calculated with the Mann-Whitney U test.

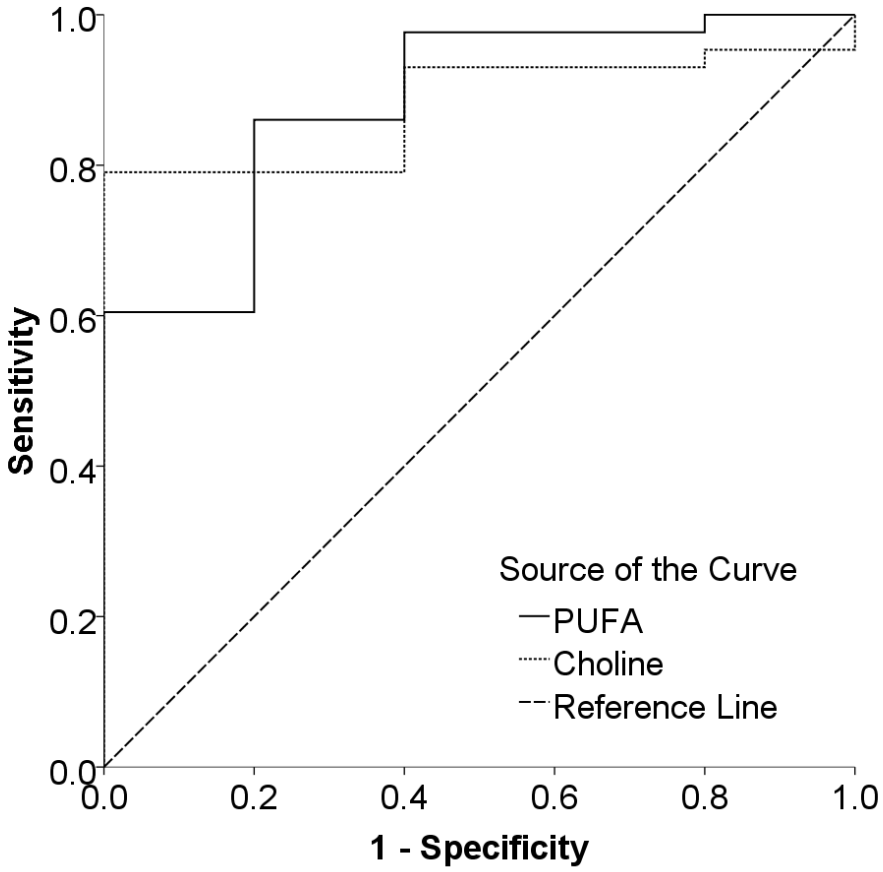


Figure 5.5-5: ROC curve for the differentiation of chronic viral hepatitis and NAFLD with the two ^{1}H-MRS parameters to show a statistically significant difference.

5.5.4 Differentiation of SS and NASH

The characteristics of the patients with NAFLD divided into those with SS and those with NASH are shown in Table 5.5-3. Other than a higher NAS score in the patients with NASH there were no statistically significant differences between the two groups.

	Patients with SS n=11	**Patients with NASH n=33**	**p-value**
Age (years)	46.2 (±13.5)	47.6 (±17.5)	0.248
Male	9 (82%)	17 (52%)	0.076
Caucasian	3 (27.3%)	4 (12.1%)	0.341
BMI (Kg/m^2)	30.3 (±5.2)	34.4 (±5.8)	0.217
Waist:hip ratio	0.97 (±0.06)	0.89 (±0.07)	0.993
Post-transplant	2 (18%)	3 (9%)	0.276
Type 2 diabetes	4 (36%)	18 (54.5%)	0.163
Smoking Status			0.154
Non-smoker	7 (64%)	17 (52%)	
Ex-smoker	3 (27%)	11 (33%)	
Current smoker	1 (9%)	5 (15%)	
Consume alcohol	4 (36.4%)	6 (18.2%)	0.237
UK units/week*	7 (2-16)	6 (1-20)	0.914
Bilirubin (units)	21.6 (±16.0)	15.7 (±3.9)	0.251
Albumin (units)	42.2 (±2.6)	40.1 (±3.7)	0.537
Fasting glucose (units)	7.0 (±2.9)	6.5 (±1.7)	0.135
Brunt grade			0.491
0	0 (0%)	0 (0%)	
1	7 (63.6%)	14 (42.4%)	
2	2 (18.2%)	12 (36.4%)	
3	2 (18.2%)	7 (21.2%)	
NAS			**0.001**
1	3 (27.3%)	0 (0%)	
2	3 (27.3%)	2 (6.1%)	
3	3 (27.3%)	4 (12.1%)	
4	2 (18.2%)	7 (21.2%)	
5	0 (0%)	13 (39.4%)	
6	0 (0%)	4 (12.1%)	
7	0 (0%)	3 (9.1%)	

Data reported as mean ±SD, with p-values from t tests; median (range), with p-values from Mann-Whitney U tests; or n (%), with p-values from Fisher's exact tests, as applicable. Bold p-values are significant at $p<0.05$

Table 5.5-3: Characteristics of patients with NAFLD divided into SS and NASH

There is no statistically significant difference between the two groups in any of the ^{1}H-MRS derived FA parameters. In patients with NAFLD, none of the calculated fatty acid parameters showed a statistically significant difference between patients with and without type 2 diabetes.

	Patients with SS n=11	**Patients with NASH n=33**	**p-value**
FF	0.11 (0.06-0.42)	0.20 (0.04-0.47)	0.226
nDB	1.72 (1.02-3.99)	1.57 (0.63-4.28)	0.487
PUFA	0.01 (0.00-0.34)	0.00 (0.00-0.19)	0.702
mCL	22.29 (16.07-29.87)	19.84 (13.69-30.10)	0.470
Choline	2.6×10^{-5} ($0.0-4.5 \times 10^{-4}$)	2.2×10^{-6} ($0.0-6.9 \times 10^{-4}$)	0.600

Data reported as median (range), with p-values from Mann-Whitney U tests.
Bold p-values are significant at $p<0.05$

Table 5.5-4: ^{1}H-MRS parameters show no statistically significant differences between SS and NASH.

5.5.5 Assessing disease activity in NAFLD

Within patients with NAFLD, fatty acid characteristics correlate with hepatic steatosis but do not correlate with histological markers of disease activity as shown in Table 5.5-5.

		FF	**nDB**	**PUFA**	**mCL**	**Choline**
Steatosis grade (Brunt)	Rho	0.811	-0.421	-0.404	-0.659	-0.360
	p	**<0.001**	**0.005**	**0.007**	**<0.001**	**0.018**
	n	44	43	43	43	43
Lobular inflammation (NAS)	Rho	-0.05	-0.20	0.24	-0.05	0.02
	p	0.77	0.21	0.12	0.74	0.91
	n	44	43	43	43	43
Hepatocyte ballooning (NAS)	Rho	0.28	-0.28	-0.01	-0.24	-0.17
	p	0.07	0.07	0.93	0.12	0.28
	n	44	43	43	43	43
Total NAS	Rho	0.497	-0.478	-0.087	-0.474	-0.224
	p	**0.001**	**0.001**	0.579	**0.001**	0.148
	n	44	43	43	43	43

Table 5.5-5: Spearman's correlation coefficient for the relationship between histological markers of NASH and FA characteristics

5.5.6 Staging of fibrosis in NAFLD

By both histological and non-invasive assessment, there was more severe fibrosis in the NASH group compared to the SS group. This is shown in Table 5.5-6.

	Patients with SS n=11	**Patients with NASH n=33**	**p-value**
Kleiner fibrosis stage			**0.002**
0	4 (36%)	1 (3%)	
1	5 (46%)	5 (15%)	
2	0 (0%)	8 (24%)	
3	2 (18%)	16 (49%)	
4	0 (0%)	3 (9%)	
ELF	7.36 (±0.99)	8.47 (±1.05)	**<0.001**
Median liver stiffness (kPa)	7.0 (±1.4)	11.4 (±5.7)	**<0.001**
CPA (%)	1.81 (±0.45)	5.27 (±4.32)	**<0.001**

Data reported as mean ±SD, with p-values from t tests or n (%), with p-values from Fisher's exact tests, as applicable. Bold p-values are significant at $p<0.05$

Table 5.5-6: By all assessment methods there was more severe fibrosis in those with NASH than those with SS.

There was no statistically significant difference in any fatty acid parameter between different fibrosis stages.

	Kleiner Stage 0-1 (n=15)	**Kleiner Stage 2-3 (n=26)**	**Kleiner Stage 4 (n=3)**	**p-value**
FF	0.19 (0.06-0.47)	0.20 (0.05-0.34)	0.16 (0.04-0.22)	0.582
nDB	1.73 (1.02-3.99)	1.48 (0.63-4.28)	1.57 (1.42-3.53)	0.604
PUFA	0.0 (0.0-0.34)	0.0 (0.0-0.19)	$8.6x10^{-4}$ (0.0-0.03)	0.714
mCL	22.1 (16.1-29.9)	19.5 (13.7-25.9)	21.7 (19.2-30.1)	0.523
Choline	0.0 ($0.0-4.52x10^{-4}$)	$7.58x10^{-6}$ ($0.0-6.91x10^{-4}$)	$2.18x10^{-5}$ ($2.20x10^{-6}$-$1.50x10^{-4}$)	0.472

Data reported as median (range), with p-values from the Jonckheere-Terpstra test.

Table 5.5-7: Characteristics of the lipidome in patients with NAFLD grouped by fibrosis stage.

5.6 Discussion

The data presented in this chapter demonstrates that the measurement and characterisation of hepatic lipids is feasible with in vivo ^{1}H-MRS at 3 Tesla. FF measured with ^{1}H-MRS has been shown in this study to have a strong correlation with histological assessment of steatosis. This is in line with published literature that supports ^{1}H-MRS as a reliable and reproducible method of non-invasively quantifying liver fat.[171, 254] Although it was possible to measure FF across the whole range of FF seen in patients and HV, it was not possible to characterise the lipidome if FF was less than approximately 1%. In cases with FF < 1%, even after water suppression, there is insufficient signal from fat to reliably fit the individual peaks on the spectrum.

Patients with NAFLD and patients with chronic viral hepatitis were well matched for all demographic and clinical parameters except for BMI and alcohol consumption. Patients with NAFLD had a significantly higher BMI and were more likely to consume alcohol. No patient consumed alcohol to excess. Routine blood tests showed no significant difference between patients with NAFLD and chronic viral hepatitis. HV had a significantly lower BMI and were more likely to consume alcohol than either patient group.

For the differentiation of chronic viral hepatitis and NAFLD there were significant differences in both the concentration of PUFA and choline. Published work using analytical chemistry techniques on liver biopsy samples to characterise the lipidome of patients with NAFLD have shown low PUFA and choline to be typical on NAFLD.[241, 249, 250, 255] These findings are also mechanistically plausible. It is proposed that the insulin resistance driven changes in lipid

metabolism in NAFLD lead to depletion of PUFA and this in turn is a driver for the de novo lipogenesis and reduced elimination of FAs from hepatocytes seen in NAFLD.[255-258] Choline deficiency is associated with increased de novo hepatic lipogenesis, alterations in membrane function and enhancement of inflammatory pathways, which are key features of models of NAFLD pathogenesis.[259-261] Indeed a choline deficient diet is a common way of inducing a NAFLD phenotype in animal models. It is also possible that the difference in FF between aetiologies is secondary to the pathological process of NAFLD rather than the cause. The lipotoxicity model of the pathogenesis of NAFLD (as described in section 5.2.1) places the accumulation of hepatic lipid as a consequence of NAFLD rather than the cause.[240, 246]

Insulin resistance is a major driver for the development of NAFLD and is of itself associated with altered lipid metabolism. It was investigated if type 2 diabetes influenced the lipidome of patients. Non-parametric tests have shown no statistically significant differences in the calculated lipid parameters between patients with and without type 2 diabetes suggesting the lipidomic changes are due to NAFLD and not insulin resistance without NAFLD.

As can be seen in Table 5.5-3, patients with SS and NASH patients were well matched in all demographic and blood parameters and this highlights the difficulty of differentiating NASH from SS without liver biopsy histology. Brunt grade was not different between SS and NASH but NAS was significantly higher in patients with NASH. This would be consistent with the increased disease activity in patients with NASH.

As shown in Table 5.5-4, the calculated lipid parameters did not show statistically significant differences between patients with SS and those with NASH. Studies such as those discussed

in section 5.2[241, 248-250] have characterised the hepatic lipidome in great detail by applying analytical chemistry techniques to liver tissue obtained at biopsy. These studies have suggested that lipidomic changes do exist and the inability of this study to demonstrate this difference suggests that the relatively crude measures that are possible with ^{1}H-MRS are not sufficiently sensitive to detect the subtle changes in lipidome between SS and NASH.

Fatty acid parameters correlate strongly with Brunt grade (
Table 5.5-5) with more severe steatosis being assosciated with increased fatty acid saturation, shorter fatty acid chains and a reduction in the concentration of both PUFA and choline. Although the correlation between hepatocyte ballooning and nDB approaches significance, there was no statistically significant correlation between any fatty acid parameter and either hepatocyte ballooning or lobular inflammation. The total NAS score is the sum of the score for steatosis, lobular inflammation and ballooning and the strength of the correlation with Brunt grade leads to nDB and mCL correlating with total NAS score. With the lack of significant correlation with ballooning or lobular inflammation it is not reasonable to say that any of the calculated fatty acid parameters truly correlate with disease activity.

As would be expected from natural history of NAFLD, fibrosis is more prevalent and more advanced in patients with NASH than SS. In this cohort no fatty acid parameter was statistically significantly different between histological fibrosis stages nor correlated with ELF, liver stiffness or CPA.

Although this technique has been shown to be feasible, there is a high failure rate with 11/67 (16.4%) spectra not suitable for analysis. This was due to a range of technical problems with

the ^{1}H-MRS acquisition leading to the collection of data that were of insufficient quality for analysis. The ^{1}H-MRS acquisition is very sensitive to artefact, patient movement during and, in particular, the homogeneity of the magnetic field. Tuning of the magnetic field to ensure homogeneity (a process called shimming) is, to a degree, dependant on the skill of the operator. For these reasons ^{1}H-MRS is likely to have a moderate failure rate. However, the assessment of hepatic lipids with ^{1}H-MRS is not a well-established technique and increasing familiarity with the technique is likely to reduce the failure rate over time.

The method of data analysis used in this study is based on previous work [180, 252, 253, 262, 263] and has been developed in conjunction with Mr Robert Flintham and Dr Nigel Davies in the medical physics department at Queen Elizabeth Hospital Birmingham. The use of automated spectrum analysis with TARQUIN[264] is novel, allows automated analysis and has been presented as an oral abstract at the ISMRM annual meeting in 2016.[265]

Although our method is based on previous work, there are a number of different approaches to the description of the lipidome with ^{1}H-MRS. These include using different acquisition techniques, using different peaks on the spectrum to define a particular chemical group and using different calculations to derive the fatty acid parameters. In this study the STimulated Echo Acquisition Mode (STEAM) sequence is used due to the smaller effect of T2 on the spectroscopy and water suppressed spectra are used for measurements due to the improved accuracy of fitting that is possible on water suppressed spectra. However, all methods are a compromise and have alternatives. We feel that the calculation of lipid characteristics in this work is robust and allows the characterisation of the hepatic lipidome with more detail and accuracy than with some other methods.

For example, to estimate the concentration of unsaturated fatty acids, Van Werven et al used the ratio of the area under the 5.3ppm peak and the area under the 1.3ppm peaks.[263] This simple approach has the advantage of making fewer measurements and therefore reducing the opportunity for measurement errors however; it ignores the signal from part of the lipid molecule and will tend to overestimate the amount of unsaturated fatty acid. Our calculation of chain length and number of fatty acid chains, although more complex accounts for the entire length of the fatty acid chain and should provide a more accurate estimate of fatty acid saturation.

As mentioned in Chapter 2, one potential source of error is the use of water suppression. This becomes relevant when attempting to measure the number of double bonds. Protons at 5.3ppm are found in the olefinic group (**-C**$\mathbf{H}$**=C**$\mathbf{H}$**-**) and protons at 2.0ppm are found in the groups alpha to the olefinic group (-C$\mathbf{H}_2$-CH=CH-C$\mathbf{H}_2$-) so either can be used to define a double bond. The use of the 2.0ppm peak avoids the potential reduction in measurement accuracy due to water suppression however, at 3 Tesla the 2.0ppm peak is small and overlaps the peak 2.8ppm. Indeed in many cases it is not possible to separate the 2.0 and 2.8ppm peaks. During development of the analysis method, it was found that calculations based on the 2.0ppm peak varied wildly depending on the quality of the fitting (data not shown) and so this peak was not used.

There are a number of possible improvements that could be made to the method of this study. A larger cohort of patients and controls would provide greater statistical power and therefore reduce the risk of error. The ability of ^{1}H-MRS to characterise the hepatic lipidomic could potentially be improved by using a more powerful MRI system. Resonant frequency is

dependent on field strength and the effect of this is that the ^{1}H-MRS spectrum at higher field strength shows more separation of the individual peaks. This improves the quality of fitting and thus increases the accuracy of measurement. This study was conducted using a 3 Tesla MRI system and the use of a more powerful 7 Tesla system should improve the accuracy of ^{1}H-MRS measurement. However, very high field strength systems are rare and certainly are not in routine clinical use. Lastly, the control group in this study are significantly different in several demographic aspects. This has the potential to introduce variations in the hepatic lipidome that are not due to the presence or absence of NAFLD. It would provide more reliable control if this study could be repeated with control subjects that are matched to patients with NAFLD in terms of age and obesity.

In summary, a method for characterising the hepatic lipidome with ^{1}H-MRS has been developed that appears to be robust and has demonstrated mechanistically plausible differences in the hepatic lipidome between patients with NAFLD and control subjects. In this study, calculated lipid parameters were not associated with disease activity or fibrosis and did not differentiate SS from NASH.

CHAPTER 6: MULTIPARAMETRIC MRI FOR THE ASSESSMENT OF PRIMARY SCLEROSING CHOLANGITIS

6.1 Primary Sclerosing Cholangitis

Primary sclerosing cholangitis (PSC) is a chronic immune mediated disease of intra and extrahepatic bile ducts characterised by cholestasis, bile duct inflammation and bile duct stricturing. The prevalence of PSC varies between populations but overall it is a rare condition. The greatest prevalence is in northern Europe where it can reach up to 16.2 cases per 100,000 people.[266, 267] There is a male preponderance and a median age of diagnosis of 41 years of age.[268]

There is a strong association with inflammatory bowel disease and in particular ulcerative colitis (UC). Up to 80% of patients with PSC have associated IBD[266] and there are several overlapping genetic risk loci.[269] Although PSC is known to be immune mediated the pathophysiology and aetiology of PSC is not well understood. It is proposed that, in individuals with a genetic susceptibility, intestinal inflammation leads to the creation of effector T lymphocytes that can migrate to the biliary epithelium and drive inflammation.[270, 271] This immune-mediated damage is perpetuated by the action of toxic bile salts and translocated bacterial toxins to cause ongoing and progressive biliary inflammation and stricturing.

In early stage disease a cholestatic pattern of liver enzymes derangement may be the only sign of disease and histological examination of liver tissue at this stage shows only portal inflammation focused around bile ducts. The histological lesions typically progress to form concentric periductal fibrosis of small, medium or large bile ducts leading to bile duct strictures. Strictures may be visualised with cholangiography either via endoscopic retrograde

cholangiopancreatography (ERCP) or magnetic resonance cholangiopancreatography (MRCP). The high sensitivity, specificity and patient acceptance of MRCP has made this a key diagnostic test in PSC.[272] Disease progression causes further bile duct stricturing that manifests as progressive biochemical derangement, jaundice, puritis and an increased risk of bacterial cholangitis. Inflammation and fibrosis of bile ducts leads to duct loss and hepatic fibrosis leading ultimately to a biliary cirrhosis with all the associated complications of cirrhosis and portal hypertension.

Despite the relatively low prevalence of PSC in the population, there is a significant burden of morbidity and mortality with more than 50% of patients progressing to liver transplantation within 15 years of diagnosis.[266] In addition to the disease burden of progressive fibrosis, patients with PSC have a markedly increased risk of colorectal carcinoma and cholangiocarcinoma compared to the general population. For cholangiocarcinoma the relative risk (95% CI) is 1560 (780-2793).[273] The cholangiocarcinoma risk is not linked to severity of fibrosis or disease duration.[273]

There are no established treatments for PSC that have been consistently shown to improve transplant free survival. Data do not support the use of immunosuppression and the use of ursodeoxycholic acid (UDCA) remains controversial. At high dose it has been shown to be harmful and at lower dose, despite an improvement in liver biochemistry, it has not demonstrated a survival advantage.[266] Some data suggest UDCA may offer some protection from colonic malignancy but studies are conflicting and no major guideline on the management of PSC recommends routine use of UDCA.[274-276]

With the absence of effective pharmacological therapy that halts the progression of PSC there is clear need for effective monitoring of disease severity to guide clinicians and patients about prognosis and allow timely intervention including liver transplantation. The variable clinical course of PSC makes accurate prognostication difficult however, several strategies have been proposed. These prognostic markers are explored below and further to this we assess multiparametric MRI for its ability to assess disease severity in PSC.

6.2 Current methods for assessment of disease severity in PSC

6.2.1 Alkaline phosphatase

Alkaline phosphatase (ALP) is an enzyme found in many human tissues including cholangiocytes. Elevation in ALP is seen in hepatic and biliary disease and is a hallmark of PSC. Indeed it is often an early indicator of disease and may be the first recognised abnormality. As well as being a clue to diagnosis ALP levels are known to be prognostically significant. Those with ALP levels less than 1.5 x the upper limit of normal at baseline have improved survival when compared to those with higher levels.[277, 278] In addition, those with elevated ALP at baseline who demonstrate a reduction in ALP to less than 1.5 x the upper limit of normal have been shown to have an improved transplant free survival compared to those whose level remains elevated.[277, 279-281] Despite the proven benefits of UDCA in primary biliary cholangitis (PBC), the use of UDCA remains controversial in PSC. There is some work to suggest that a response to UDCA (defined by reduction in ALP) is prognostically favourable.[280] However this conclusion is not universally recognised and other studies show that transplant free survival is independent of the use of UDCA.[277]

It can therefore be seen that ALP levels and the change in ALP levels over time are useful to clinicians to inform about prognosis. However, ALP measurements are labile and can be transiently elevated by complications of PSC such as episodes of cholangitis or in the presence of a dominant biliary stricture. This lability means that ALP levels must be viewed in context and changes may not necessarily reflect progression of disease.

6.2.2 Mayo PSC score

The revised Mayo Clinic natural history model was developed in 2000 with the aim of defining the risk of survival in PSC using readily available clinical data.[18] It built on previous models for defining natural history that all included histological assessment of fibrosis stage as part of the model.[282-285] The model includes age, bilirubin, aspartate transferase, serum albumin and the occurrence of variceal bleeding. This revised score has the clear advantage that is can be calculated without requiring liver biopsy. In the validation group from the initial study patients were categorised by the model as low, medium and high risk. For these 3 groups there was no significant difference between predicted and actual survival.[18]

Subsequent work with the Mayo PSC score has shown it to be superior to Child's-Pugh score in predicting death in patients with PSC.[286] However, the Mayo score performs best in predicting events in later stage disease.[287] and it does not entirely predict the risk of hepatobiliary or colonic cancer associated with PSC.[288]

6.2.3 Model for end stage liver disease

The model for end stage liver disease (MELD) score was developed to predict risk of death following transjugular intrahepatic portosystemic shunts (TIPSS)[289] but has found acceptance as a prognostic index more generally and for liver transplant organ prioritisation.[290, 291] It is not specific to PSC but has been shown to predict the risk of adverse outcomes in patients with PSC [266, 278] and also to predict short term survival following liver transplant.[292]

6.2.4 Cholangiography

Visualisation of the biliary tree with either ERCP or MRCP is a key diagnostic test in PSC. The two modalities are of equivalent accuracy in the detection of biliary strictures with significantly reduced risk with MRCP.[293] In addition to use as a diagnostic tool; cholangiography has relevance to prognosis. This association has been recognised since at least 1995 when a retrospective study of 94 patients with PSC demonstrated worse outcome in patients with extensive intrahepatic structuring.[294]

A scoring system for classifying cholangiographic abnormalities was developed in 1991 by Majoie et al.[295] This was initial developed as a diagnostic tool however, worsening of the score over time was demonstrated by the same group in 2001.[296] Cholangiographic scoring was further developed and its use to predict radiological progression and clinical outcomes has been demonstrated.[297-299]

6.2.5 Non-invasive markers of fibrosis

Fibrosis stage is prognostically important in PSC as the majority of morbidity and mortality in PSC is as a result of progressive fibrosis and the complications of cirrhosis.[300, 301] Fibrosis stage assessed by biopsy is a component of several PSC risk models[282-284] and clearly predicts the risk of hepatic decompensation and death.[285] However, liver biopsy is not commonly indicated for the diagnosis of PSC and the patchy nature of histological changes in PSC make sampling error a major concern. This, coupled with the inherent risk of biopsy, make biopsy unattractive for prognostication in PSC. Several non-invasive markers of liver fibrosis have

emerging evidence for a role in assessing prognosis in PSC and avoid many of the potential problems with biopsy to stage fibrosis.

Vibration controlled transient elastography

Vibration controlled transient elastography (VCTE) as measured by FibroScan has been shown to accurately predict fibrosis stage in a range of chronic liver diseases. A large study of patients with PSC looked at baseline liver stiffness measurements and change in liver stiffness over time for their ability to predict outcomes. Both a low baseline liver stiffness and slow progression were associated with improved survival.[302] The finding that a single liver stiffness reading can predict histological fibrosis stage and outcome have been replicated in more recent work.[303] The prediction was strongest for advanced disease.

Enhanced liver fibrosis test

Enhanced liver fibrosis (ELF) test has been shown to predict outcomes in PSC with higher scores having worse prognosis.[304, 305] AUROC for the prediction of adverse events (transplantation or death) in these studies was between 0.78 and 0.81.[304, 305] ELF test is attractive as it is based on blood tests and is therefore simple to perform and is not influenced by the variability in fibrosis across the liver seen in PSC. It should however be borne in mind that the data to date regarding ELF in PSC is retrospective and from a single centre. Further prospective validation would increase confidence in its reliability.

Magnetic resonance elastography

Only one large retrospective study has assessed the ability of magnetic resonance elastography (MRE) to stage fibrosis and predict outcome in PSC.[306] This 2016 study shows

that MRE can stage fibrosis with performance comparable to other aetiologies. It also seems to provide additional prognostic information in that those with high scores were at greater risk of decompensation.[306] The authors note several cases with variations in liver stiffness across the liver reflecting the patchy nature of PSC. Further work assessing the use of MRE in PSC prospectively is ongoing (NCT02446665) and should be useful in defining the role of this promising technique.

6.3 Aims

As discussed above, the variability in the clinical course of PSC and the current paucity of tests that predict clinical endpoints both contribute to the difficulty in effectively assessing, prognosticating and monitoring of disease progression or regression in PSC.

Multiparametric MRI has been shown in Chapter 3 to be useful in the assessment of fibroinflammatory liver disease. This study will investigate if multiparametric MRI can be used to assess PSC and monitor progression over time.

Specifically the aim of this study is to assess the ability of multiparametric MRI to:

- Evaluate the severity of PSC when compared to established markers of disease severity
- Track changes in PSC severity over time

6.4 Methods

6.4.1 Study overview

Participants were invited to attend for two study visits 18 months apart. Patients underwent multiparametric MRI, FibroScan and blood tests at each visit. Throughout the study, participants continued with their routine standard of care.

This prospective study was undertaken at the Queen Elizabeth Hospital Birmingham (QEHB). Patients were recruited between March and September 2014. The study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki, and was approved by the National Research Ethics Service (West Midlands – The Black Country; REC Ref: 14/WM/0010).

6.4.2 Study Participants

Male and female adult (≥18 years of age) patients with PSC were recruited from the specialist PSC clinic at QEHB. Patients seen in the clinic with PSC and without contraindication to MRI received a written invitation and information sheet. Those who responded gave written informed consent and underwent baseline study investigations. Those who completed baseline investigations were invited to attend for repeat evaluation 18 months after their baseline investigations.

6.4.3 Study Investigations

All MRI scans were performed and analysed as described in Chapter 2 with the exception that due to the noticeable heterogeneity of PSC a single, user defined, region of interest (as used in chapters 3 and 4) did not give a representative assessment of cT1 values across the liver as a whole. Figure 6.4-1 shows representative images that demonstrate this variability. Because of this variability, values for mean, median and mode cT1 across a whole liver slice were recorded.

FibroScan™ examinations were performed by one trained operator (PJE) in accordance with manufacture's guidelines. The decision on using the M probe or XL probe was made on the skin to liver capsule distance measured by the FibroScan machine. Examinations were regarded as 'possible' if at least 10 valid readings could be recorded and 'reliable' if they contained at least 10 valid readings and had interquartile range (IQR) to median ratio ≤30% (Boursier's criteria).[182]

Blood samples were analysed routinely for markers of liver disease. Simple blood biomarker panels including AST/ALT ratio, FIB-4, MELD and Mayo revised PSC score were calculated according to published formulae.[18, 49, 289] Serum was also analysed to determine the ELF score (iQur Limited, London, UK).

6.4.4 Clinical events

At the second clinical visit history was taken about a range of clinical events. The occurrence and date of: death, liver related death, liver transplant, variceal bleeding, decompensation of cirrhosis, new diagnosis of cholangiocarcinoma, episode of cholangitis, flare of IBD, emergency hospital admission due to liver disease and emergency hospital admission not related to liver disease was recorded. If the event occurred multiple times the date of the first occurrence was recorded.

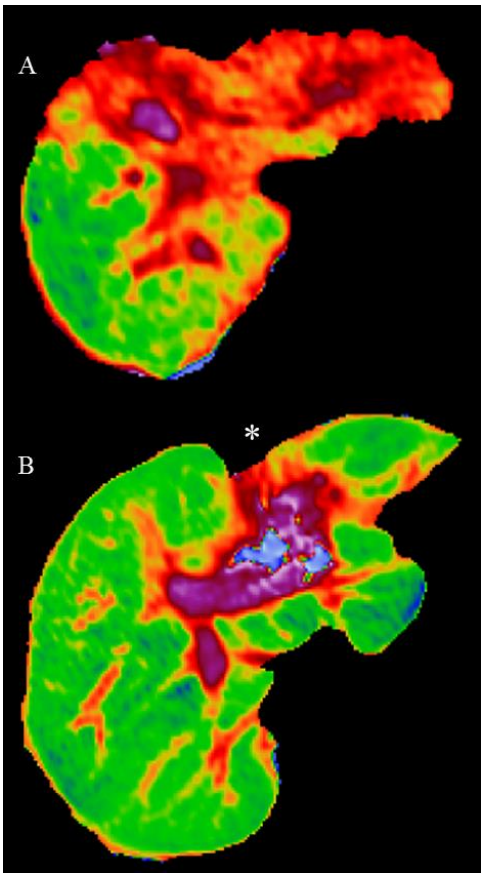


Figure 6.4-1: Representative cT1 maps showing the heterogeneity seen in PSC. A: Marked difference in cT1 between Right and left lobes of the liver. B: Liver with largely uniform cT1 values except for one area of high cT1 (*) that may represent focal scaring.

6.5 Results

6.5.1 Cohort at baseline

21 patients were consented to take part in the study and underwent baseline investigations. The characteristics of the cohort at baseline are shown in Table 6.5-1.

	Baseline cohort n=21
Age (years)	50 (18 – 73)
Male gender	15 (71%)
Caucasian	14 (95%)
BMI (kg/m^2)	23.8 (±2.8)
Small duct only phenotype	5 (24%)
Duration of disease (years)	5.8 (±3.8)
Co-existent inflammatory bowel disease (IBD)	19 (91%)
Use of azathioprine	3 (14%)
Use of UDCA	10 (48%)
UDCA dose (mg/kg)*	13.5 (±3.5)
Smoking status	
Current	1 (5%)
Previous	5 (24%)
Never	15 (71%)
Consume alcohol	7 (33%)
Alcohol intake (UK units/week)**	6 (1 – 20)

** In patients taking UDCA*

*** In patients who consume alcohol*

Data presented as mean (±SD), median (range) or n (%) as appropriate.

Table 6.5-1: Patient characteristics at baseline.

Baseline values for established non-invasive markers of fibrosis and PSC severity are shown in Table 6.5-2. Missing data was due to one patient not having a FibroScan or AST measured.

Another patient did not have a platelet count available due to a clotted sample. A summary of cT1 values at baseline is shown in Table 6.5-3

Test	n	Median (range)	Risk groups	
ALP (U/L)	21	225 (41 – 708)	≤1.5 x ULN [277]	8 (38%)
Mayo PSC score	20	-0.53 (-1.87 – 2.13)	< 0 0-2 >2 [286]	15 (75%) 4 (20%) 1 (5%)
MELD	21	6.0 (6.0 – 14.6)		
AST:ALT ratio	20	0.86 (0.44 – 1.70)		
APRI	19	0.61 (0.08 – 2.50)		
Fib-4	19	1.37 (0.30 – 5.37)		
Liver stiffness (kPa)	20	9.8 (3.0 – 36.8)		
ELF test	21	9.92 (8.50 – 13.25)		

Data presented as median (range) or n (%) as appropriate.
ULN: upper limit of normal ULN for ALP in the QEHB lab: 130 U/L

Table 6.5-2: Summary of measures of disease severity in the cohort at baseline.

Test	**n**	**Median (range)**
Mean cT1 (msec)	21	917 (736 – 1135)
Median cT1 (msec)	21	848 (697 – 1108)
Mode cT1 (msec)	21	802 (657 – 1080)

Table 6.5-3: Summary of cT1 values at baseline.

Histological assessment of PSC was not included in the study however, 8 (38%) of patients had undergone a standard of care liver biopsy prior to study entry. These liver biopsies were performed from 1 to 12 years before study recruitment and so their relevance is limited.

Histological assessment of fibrosis was: none/minimal in 2 patients, portal fibrosis only in 1 patient, bridging fibrosis in 4 patients and cirrhosis in 1 patient.

6.5.2 Correlation between non-invasive markers in PSC

The correlation between non-invasive markers of fibrosis and disease severity described in section 6.5.2 was presented as a poster at the European Association for the Study of the Liver International Liver Congress 2016.[307] cT1, ALP, AST:ALT ratio and MELD score did not significantly correlate with any other marker of fibrosis or disease severity. Analysing ELF, liver stiffness, APRI and FIB-4 in a pairwise fashion revealed statistically significant, positive correlation in all six parings (Table 6.5-4). The strongest association was between liver stiffness and ELF score (rho=0.706, p=0.001) (Figure 6.5-1). Mayo PSC score showed statistically significantly correlation with ELF (rho=0.592 p=0.006), liver stiffness (rho=0.559 p=0.010) and FIB-4 (rho=0.733 p<0.001).

		ELF	**Fib-4**	**Liver stiffness**
APRI	Rho p n	0.602 0.006 19	0.616 0.005 19	0.565 0.012 19
ELF	Rho p n		0.572 0.011 19	0.706 0.001 20
Fib-4	Rho p n			0.639 0.003 19

Table 6.5-4: Correlation between non-invasive markers of fibrosis in PSC

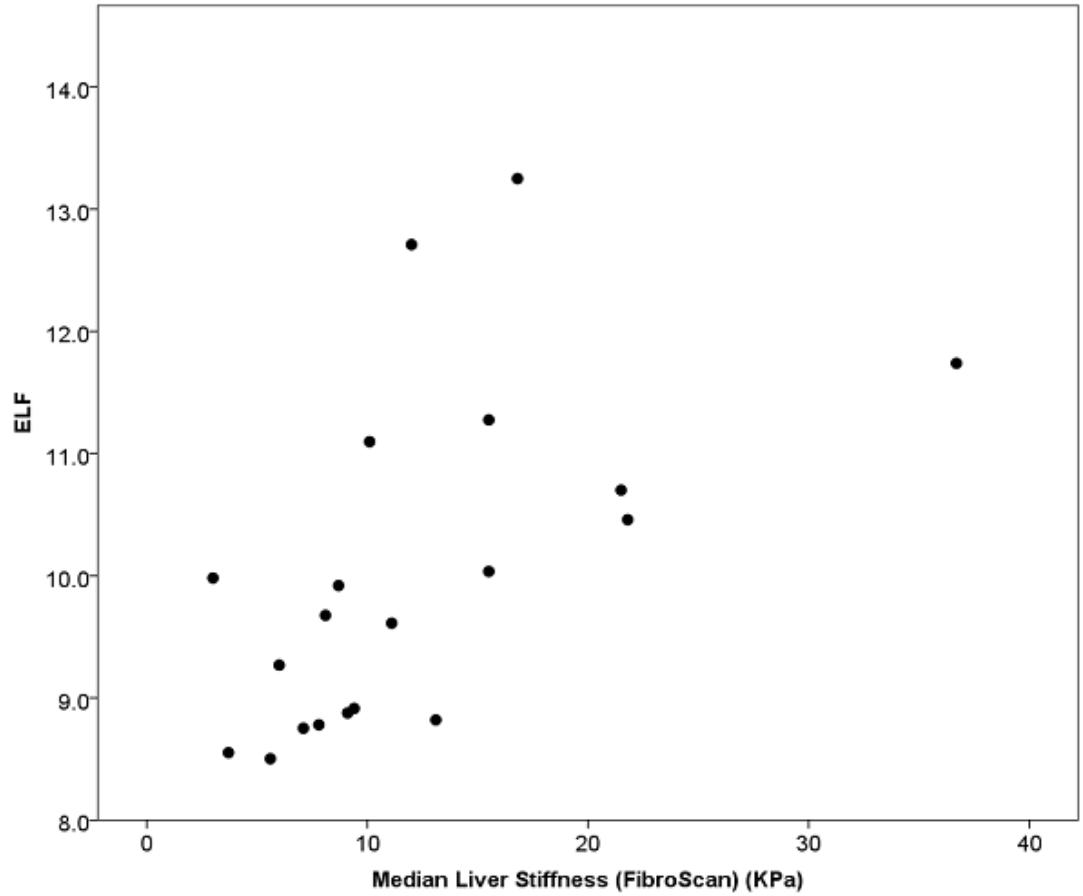


Figure 6.5-1: Scatter plot showing the correlation between ELF and liver stiffness at baseline

cT1 did not correlate with any non-invasive marker of fibrosis in a statistically significant manner. The only statistically significant correlation with a marker of disease severity was a modest negative correlation between ALP and mode cT1 (Rho = -0.447, p = 0.042) and between ALP and median cT1 (Rho = -0.444, p = 0.044). The fact that this is a negative correlation is unexpected as it would be expected that both ALP and cT1 would rise with PSC severity.

6.5.3 Cohort at time of follow-up

17 patients attended for their 2nd scan after a median (range) follow up of 18.4 (18.2-19.1) months. The reasons for dropping out of the study were withdrawal of consent: 1 patient, inability to contact patient to book 2nd scan: 2 patients, liver transplantation before 2nd scan: 1 patient. Fibrosis and diseases severity markers at the time of follow-up are shown in Table 6.5-5. The change in ALP correlates quite strongly with the change in Mayo score with Rho = 0.659, p=0.006 (Figure 6.5-2).

	Baseline		**Follow-up**	
Test	**n**	**Median (range)**	**n**	**Median (range)**
Mean cT1 (msec)	21	917 (736 – 1135)	17	898 (810 – 1006)
Median cT1 (msec)	21	848 (697 – 1108)	17	822.6 (754 – 934)
Mode cT1 (msec)	21	802 (657 – 1080)	17	780 (715 – 881)
ALP (U/L)	21	225 (41 – 708)	17	242 (52 – 655)
Mayo score	20	-0.53 (-1.87 – 2.13)	16	-0.37 (-2.85 – 0.58)
MELD	21	6.0 (6.0 – 14.6)	17	6.0 (6 – 11.1)
AST:ALT	20	0.86 (0.44 – 1.70)	16	0.89 (0.47 – 2.44)
APRI	19	0.61 (0.08 – 2.50)	16	0.60 (0.11 – 4.65)
Fib-4	19	1.37 (0.30 – 5.37)	16	2.00 (0.30 – 6.05)
Liver stiffness (kPa)	20	9.8 (3.0 – 36.8)	17	10.3 (3.3 – 35.3)
ELF test	21	9.92 (8.50 – 13.25)	17	10.09 (7.96 – 12.39)

Data presented as median (range) or n (%) as appropriate.

Table 6.5-5: Fibrosis and diseases severity markers at baseline and the time of follow-up.

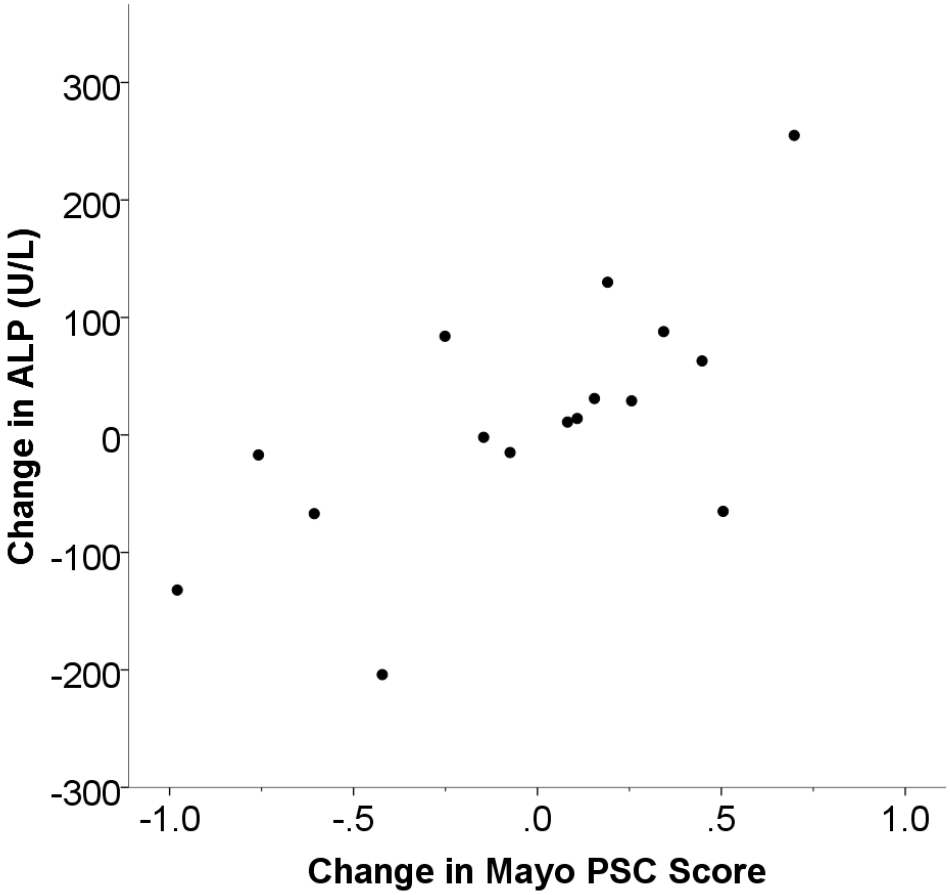


Figure 6.5-2: Scatter plot showing the correlation between the change in ALP and the change in Mayo PSC score.

6.5.4 Incidence of clinical events

Information on clinical events during follow up was available from 18 of the 21 patients who consented to take part in the study. During follow up 12 different patients had a total of 16 clinical events as outlined in Table 6.5-6.

Event	**n**
Death	0
Liver transplant	1
Variceal bleeding	0
Decompensation of cirrhosis	0
New diagnosis of cholangiocarcinoma	0
Episode of cholangitis	6
Emergency hospital admission due to liver disease (All for treatment of cholangitis)	3
Flare of IBD	4
Emergency hospital admission not related to liver disease (For treatment of IBD flare)	1
Other (elective admission for colectomy due the high grade dysplasia)	1

Table 6.5-6: Clinical events during follow-up.

Only one patient encountered a significant event (liver transplant,). This patient had significantly raised severity and fibrosis markers but near normal cT1 at baseline as shown in Table 6.5-7.

Test	Value at baseline
Bilirubin (<22 μmol/L)	161
ALP (40-130 U/L)	708
Mayo PSC score	1.77
MELD	15
UKELD	56
Child's-Pugh score	7
AST:ALT ratio	1.16
APRI	0.88
FIB-4	1.37
ELF	13.25
Liver stiffness (kPa)	16.8
Mode cT1 (msec)	690
Median cT1 (msec)	738.25
Mean cT1 (msec)	829.9

Normal range and units shown if applicable

Table 6.5-7: Baseline investigations for LAMP-B-002

When the definition of an event was expanded to include an episode of cholangitis no severity or fibrosis marker showed a statistically significant difference between those having an event than those who did not. There was a trend towards a higher ALP in those having an event, however this difference did not reach statistical significance (median: 192 vs 300U/L, p = 0.056). Excluding episodes of cholangitis that did not require a hospital admission shows a significant difference in ALP between groups (median: 188.5 vs 350U/L, p = 0.040). All patients meeting this endpoint had ALP greater than 1.5 times the upper limit of normal (Figure 6.5-3).

Neither baseline cT1, change in cT1 during follow up nor other fibrosis or severity marker showed a statistically significant difference between those with and without a clinical event regardless of the definition of an event

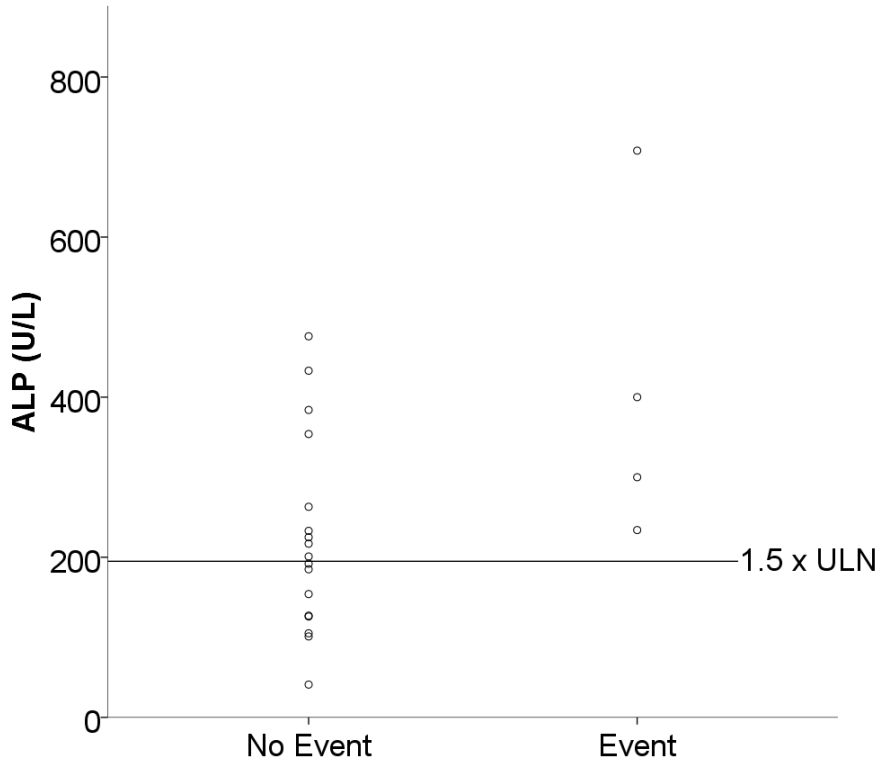


Figure 6.5-3: ALP levels in those encountering a liver related event during follow-up. Events encountered were liver transplant or cholangitis requiring hospital admission. Groups are statistically significantly different by the Mann-Whitney U test (p = 0.040). 1.5 x ULN (upper limit of normal) = 1.5 x 130 U/L = 195 U/L

6.6 Discussion

In this chapter is described a pilot study assessing the utility of multiparametric MRI in the assessment of PSC. The cohort of 21 patients with PSC has typical demographics of patients attending the PSC clinic at QEHB. No patient had evidence for an additional aetiology for chronic liver disease other than PSC. IBD is prevalent and approximately half of participants take UDCA, generally at a modest dose.

Non-invasive markers of disease severity give mixed results. The majority (75%) of patients have a Mayo score in the low risk group however, this is contradicted by ALP. The majority of patients have an ALP value >1.5x ULN indicating an increased risk of events. ALP was noted to not correlate with any other fibrosis or disease severity marker in this cohort whereas Mayo score correlated well with ELF, liver stiffness and Fib-4. Across the cohort as a whole fibrosis markers are generally modestly elevated at baseline. The median of both ELF and liver stiffness indicate moderate fibrosis. The majority of fibrosis markers correlate with each other. This good correlation between non-invasive markers of fibrosis supports their use to stratify risk in PSC for the long term follow up of patients, as well as in the clinical trial settings. Baseline cT1 however, does not correlate in any meaningful way with any other fibrosis or disease severity marker. The negative correlations between ALP and mode cT1 and ALP and median cT1 are unexpected. Both cT1 and ALP would be expected to increase with worsening disease and as these correlations are negative and only just reach significance they are likely to be spurious. I would expect that in a larger cohort it would not be statistically significant.

A notable feature of the baseline cT1 values shown in Table 6.5-3 is how low they are. The reason for this discrepancy is not clear but it can be seen across the whole cohort of patients with PSC that their cT1 values tend to be lower than seen in other cohorts studied. Despite several patients in the cohort having robust evidence of significant fibrosis, the mean (±SD) of median cT1 values in this cohort is 849.6 (±89.1) msec. In the work described in Chapter 3 patients with Ishak stage 1-2 fibrosis had median cT1 of 851.7 msec and in the published work by Banerjee et al median cT1 for patients with Ishak stage 1-2 fibrosis was 870.[163] It is possible that an unexpectedly low cT1 may be a feature of PSC although the mechanism for this is not clear. The study comparing biopsy data and cT1described in Chapter 3 contain only 2 patients with PSC and so it is not possible to show if cT1 is lower than in other aetiologies.

Over the period of follow-up, most markers showed a small increase in the median, which would suggest some progression of disease. However, within individuals there was little agreement between severity markers. The magnitude and often direction of change varied between markers. Of note, the cT1 values fell over the period of follow up. In all other situations so far examined, cT1 has risen with increasing inflammation and worsening fibrosis and so it was expected that cT1 values would rise in line with the other non-invasive markers of disease severity. It is unclear why cT1 has not behaved as expected in this cohort. This may be result of the a statistical quirk given the small sample size or it may be reflection of the behaviour of cT1 in PSC.

It has been shown that cT1 does not correlated with any other surrogate marker of fibrosis or disease severity but these markers have their own limitations. Assessing the ability of cT1 to predict clinical events is a more robust way of assessing the utility of cT1 in PSC. Clinical

events were recorded for patients in the study however, with so few patients who were mostly at an early stage in their disease, there were few significant events. This makes the assessment of clinical events difficult. Mayo score has been designed to identify patients at risk of death or transplant[18] and not cholangitis so it is not unexpected that there is no difference between those encountering an episode of cholangitis and those who did not. Again, the ALP cut off of <1.5x ULN has been shown to identify patients at risk of serious events in PSC (death, liver transplantation, or diagnosis of cholangiocarcinoma)[277, 278] and not cholangitis. However, those with an ALP value ≥1.5x ULN had an increased risk of cholangitis that required hospital admission. This may reflect that patients with more advanced liver disease are more severely affected by an episode of cholangitis rather than an increase in frequency.

As noted above, detailed statistical analysis is unlikely to provide robust results in a small cohort such as this with few events. It is worth considering if the participant who did go on to have a significant event can be instructive. The one patient in the cohort who underwent liver transplantation during follow-up had the study number LAMP-B-002. Non-invasive markers would place them in a high risk group for disease progression and requiring transplantation. Baseline investigations for this patient are shown in Table 6.5-7. It is striking from these data that the cT1 values are within the normal range and so suggest the absence of significant fibro-inflammatory liver disease. This clearly contradicts the other non-invasive markers and histological examination of the hepatectomy specimen showed significant inflammation.

6.7 Conclusions

cT1 values, as they have been measured in this experiment, have not shown significant correlation with other surrogate markers and have not behaved as expected from other work. The absolute values have been lower than expected and cT1 appears to have fallen as other severity markers have risen. It is not clear why this has occurred and it open and interesting avenue for further work.

It is possible that the unexpected cT1 results are due to the way cT1 has been summarised with mean, median and mode. This may not be an appropriate way of looking at cT1 in PSC. It was noted that the pattern of cT1 was different between patients with PSC and those with 'parenchymal' liver diseases such as chronic viral hepatitis. Measuring an 'average' cT1 value may be missing valuable information such as the distribution or pattern of cT1. Figure 6.7-1 demonstrates the heterogeneity and generally lower value of cT1 in LAMP-B-002 compared to the higher and more uniform cT1 values in a patient with chronic hepatitis C infection. Both patients were cirrhotic and similar in terms of non-invasive markers of fibrosis. Further investigation of the importance of cT1 distribution is currently underway.

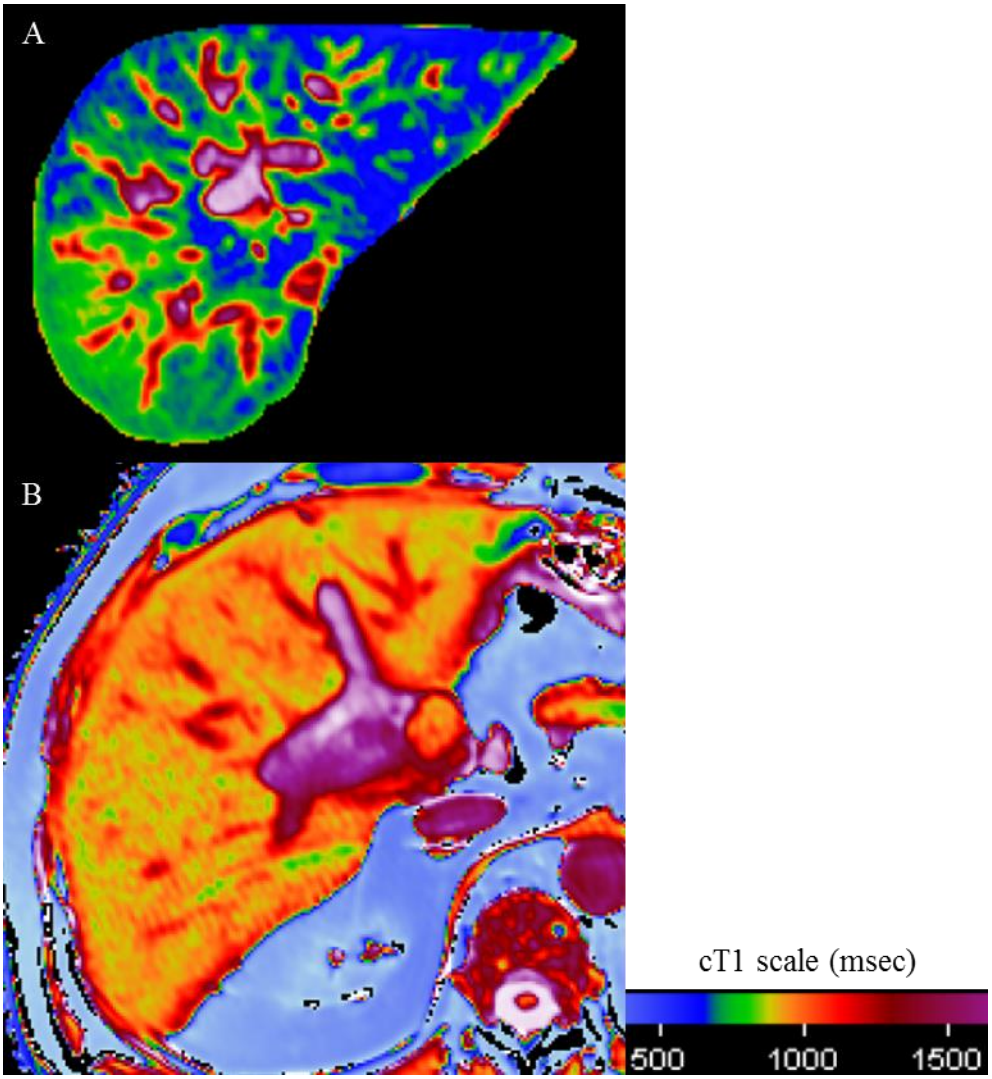


Figure 6.7-1: Representative images of: A. Patient with PSC and B. Patient with chronic hepatitis C infection. Both patients have similar stage fibrosis based on histology and non-invasive markers but cT1 is markedly different between them both in terms of distribution and average value.

6.8 Opportunities for further work – cT1 distribution

By visual inspection of cT1 maps it can be seen that the variability in cT1 values across the liver in patients with PSC is markedly different to healthy controls and patients with a 'parenchymal' liver disease such as chronic viral hepatitis or autoimmune hepatitis (AIH). It is proposed that by measuring cT1 distribution it may be possible to discriminate PSC from other liver disease and to risk stratify PSC.

The variation in cT1 values can be expressed as a histogram as seen in Figure 6.8-1. The shape of the histogram or kurtosis can be quantified as can the skewedness of the distribution (Figure 6.8-2). It is currently under investigation if the distribution of cT1 values is a more useful metric to assess the severity of PSC.

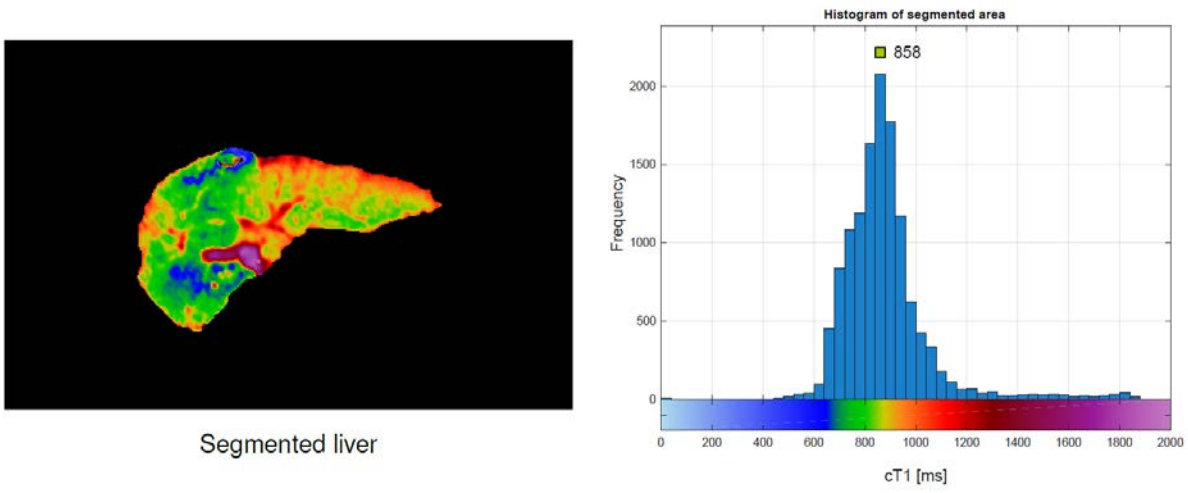


Figure 6.8-1: Liver slice and cT1 histogram from a patient with PSC.

In an extension to the current study a further 180 patients with autoimmune liver disease (PSC, AIH and primary biliary cholangitis (PBC)) stratified into low and high risk groups are being recruited to see if cT1 distribution can differentiate these diseases and assess severity.

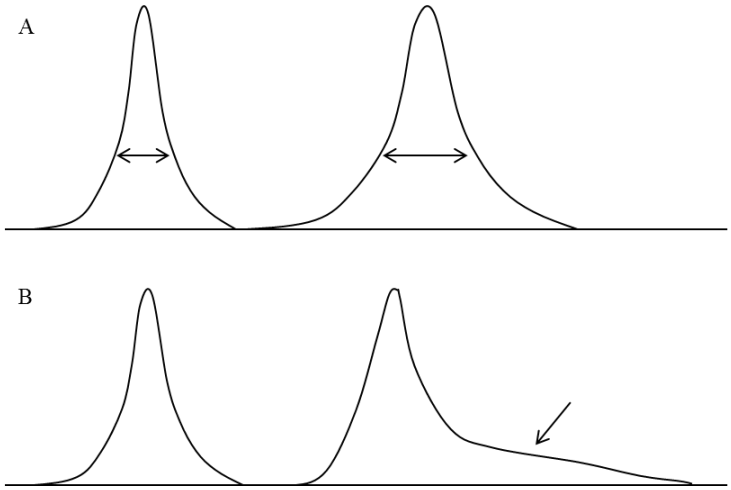


Figure 6.8-2: A. The curve on the right shows a greater kurtosis than the curve on the left. B. A skewed distribution on the right compared to a more symmetrical distribution on the left.

In addition; the inclusion of MRCP in the protocol for this extension to the existing study will allow the co-localisation of cT1 values and bile ducts to assess how cT1 values immediately adjacent to bile ducts varies between PSC, AIH and PBC. Recruitment to this extension study is currently underway.

CHAPTER 7: OVERALL CONCLUSIONS

7.1 Summary

Chronic liver disease is a major source of morbidity and mortality and with increasing prevalence there is an urgent need to find safe, accurate and cost effective methods to assess and stage chronic liver disease. This thesis presents a comprehensive assessment of the performance and utility of multiparametric MRI in several different situations with a view to establishing whether multiparametric MRI has a role is evaluating chronic liver disease severity.

A review of the current literature in Chapter 1 highlights the need for the development of biomarkers. Although liver biopsy is a valuable technique there are inherent risks and limitations that make it a less than ideal tool for the staging of hepatic fibrosis. The reviewed literature demonstrates that current non-invasive biomarkers have many strengths and can identify cirrhosis with a high degree of accuracy. However, the performance in identifying moderate fibrosis is less good and performance is generally poor for the identification of early stage fibrosis. I would suggest that the ideal biomarker is not currently available and there is a need for the development of biomarkers that can accurately and reproducibly stage fibrosis and track progression and regression of fibrosis over time.

In Chapter 2 the multiparametric MRI and ^{1}H-MRS techniques used in this work are outlined. It has been demonstrated that cT1 measured with multiparametric MRI is reproducible, does not require fasting and, in healthy individuals, is stable over time. These are important findings in the further investigation and evaluation of multiparametric MRI.

In Chapter 3 two studies are described that evaluate multiparametric MRI analysed with Liver*Multiscan*™ in the staging of hepatic fibrosis. Data from a large two centre study show that cT1 has demonstrated utility in staging fibrosis with performance comparable to established non-invasive biomarkers of fibrosis. In a separate study, the use of multiple histology samples from hepatectomy specimens were used as the reference standard with the aim of reduce the sampling error inherent in liver biopsy histology. Unfortunately the method for this study encountered difficulties that have reduced the ability of this experiment to provide robust data. However, there are data to support the use of averaged cT1 across the whole liver rather than from a single region of interest to stage fibrosis.

It is clear from the data in this chapter that inflammation is a major confounder of fibrosis assessment with multiparametric MRI. Indeed cT1 should be thought of as a marker of 'fibro-inflammatory disease' rather than a marker of either fibrosis or inflammation individually. This raises the possibility that multiparametric MRI may be useful for applications other than staging fibrosis. The sensitivity to inflammation could be used to monitor response to treatment in autoimmune hepatitis or as an indication for treatment in chronic hepatitis B infection. In both situations there is potential to avoid liver biopsy and thus reduce risk to the patient and cost to health services.

Another aetiology where assessment of both fibrosis and inflammation would have the ability to give insight into the disease process is non-alcoholic fatty liver disease (NAFLD). Non-invasive assessment of NAFLD is evolving and has been highlighted as a key goal for research.[80] The diagnosis of non-alcoholic steatohepatitis (NASH) and assessment of disease activity with non-invasive means would have significant impact in clinical practice and as a

surrogate endpoint in research studies. Based on the data presented in Chapter 4, there is potential for multiparametric MRI to diagnose NASH, however, cT1 did not correlate with fibrosis in this study and, given the recent work confirming the crucial role for fibrosis assessment in NAFLD[219] this is a major drawback of assessment with multiparametric MRI.

In Chapter 5 a novel application of ^{1}H-MRS was investigated for its ability to diagnose NAFLD and differentiate simple steatosis and NASH. This work has developed a novel approach to the processing and analysis of ^{1}H-MRS data from the liver. This small study suggested that there were differences in the lipid profile between patients with NAFLD and controls but could not differentiate simple steatosis and NASH. This work has potential to provide further interesting insights into the lipid abnormalities underlying NAFLD. To draw more definitive conclusions the study could be repeated at higher field strength in a larger cohort with better matched controls. With further refinement, the ability of ^{1}H-MRS to define the hepatic lipidome will increase and this may become a clinically useful technique.

In Chapter 6 the ability of cT1 to stage fibrosis and evaluate disease severity in PSC was investigated. In this study cT1 did not predict fibrosis or disease severity at baseline. cT1 was noted to be lower than would have been expected from the data resented in Chapter 3 and cT1 appeared to fall as disease severity worsened. If this finding can be repeated then it is contrary to the rise in cT1 that would be expected in the presence of worsening fibrosis and inflammation. Visual inspection of the cT1 maps from patients with PSC shows striking differences compared to the cT1 maps of patients with other aetiologies. The distribution of cT1 values across the liver may be useful in the diagnosis and assessment of PSC. An

extension to this study is currently underway and should provide further interesting information on the applicability of multiparametric MRI in PSC.

7.2 Added value and limitations of multiparametric MRI

In this thesis it has been demonstrated that cT1 correlates with histological assessment of liver fibrosis but that this ability to stage fibrosis is less good than in previous work by Banerjee et al.[163] The reason for this remains unclear and, in addition it has become clear through this work that cT1 is heavily confounded by other factors such as inflammation, steatosis and aetiology. This confounding makes a cT1 value difficult to interpret and so whether or not multiparametric MRI can find a clinical role remains to be seen. A further limitation of the current technique is the use of a single region of interest (ROI) for cT1 calculation. This technique relies on human judgement and has potential for variability and error. Data presented in Chapter 3 suggests that the average cT1 value throughout the whole liver may better predict fibrosis and this should be investigated further.

7.3 Suggested avenues for future work

The evaluation of whole liver analysis would be of value as initial data from Chapter 3 suggests this may be a more reliable marker of fibrosis. I would propose a further study comparing cT1 measured from a single ROI and from the whole liver against histological assessment of fibrosis. In this proposed study the use of a single liver disease aetiology would allow a validated score for the histological assessment of inflammation to be used (such as the Ishak grading of necro-inflammation in chronic viral hepatitis) to further define the influence of inflammation of cT1.

The work described in Chapter 6 looking at the evaluation of autoimmune liver disease has already been extended and a larger cohort of patient has been recruited with a range of autoimmune liver disease. As described in Section 6.8 this study aims to look more closely at the distribution of cT1 to evaluate if these parameters are clinically useful in evaluating autoimmune liver disease.

REFERENCES

1. Eddowes PJ, McDonald N, Davies N, et al. Utility and cost evaluation of multiparametric magnetic resonance imaging for the assessment of non-alcoholic fatty liver disease. Aliment Pharmacol Ther 2017;47:631-644.
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. The Lancet 2014;384:1953-1997.
3. Martin NK, Hickman M, Hutchinson SJ, et al. Combination interventions to prevent HCV transmission among people who inject drugs: modeling the impact of antiviral treatment, needle and syringe programs, and opiate substitution therapy. Clin Infect Dis 2013;57 Suppl 2:S39-45.
4. Brenner DA. Molecular Pathogenesis of Liver Fibrosis. Transactions of the American Clinical and Climatological Association 2009;120:361-368.
5. Bataller R, xF, Brenner DA. Liver fibrosis. The Journal of Clinical Investigation 2005;115:209-218.
6. Schuppan D, Afdhal NH. Liver Cirrhosis. Lancet 2008;371:838-851.
7. Lucero C, Brown RS, Jr. Noninvasive Measures of Liver Fibrosis and Severity of Liver Disease. Gastroenterol Hepatol (N Y) 2016;12:33-40.
8. Ishak K, Baptista A, Bianchi L, et al. Histological grading and staging of chronic hepatitis. J Hepatol 1995;22:696-9.
9. Everhart JE, Wright EC, Goodman ZD, et al. Prognostic value of Ishak fibrosis stage: findings from the hepatitis C antiviral long-term treatment against cirrhosis trial. Hepatology 2010;51:585-94.
10. Bataller R, Gao B. Liver fibrosis in alcoholic liver disease. Semin Liver Dis 2015;35:146-56.
11. Chedid A, Mendenhall CL, Gartside P, et al. Prognostic factors in alcoholic liver disease. VA Cooperative Study Group. Am J Gastroenterol 1991;86:210-6.
12. Niederau C, Fischer R, Purschel A, et al. Long-term survival in patients with hereditary hemochromatosis. Gastroenterology 1996;110:1107-19.
13. Panayi V, Froud OJ, Vine L, et al. The natural history of autoimmune hepatitis presenting with jaundice. Eur J Gastroenterol Hepatol 2014;26:640-5.
14. Angulo P, Kleiner DE, Dam-Larsen S, et al. Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology 2015;149:389-97.e10.
15. Manousou P, Dhillon AP, Isgro G, et al. Digital image analysis of liver collagen predicts clinical outcome of recurrent hepatitis C virus 1 year after liver transplantation. Liver Transpl 2011;17:178-88.
16. Yano M, Kumada H, Kage M, et al. The long-term pathological evolution of chronic hepatitis C. Hepatology 1996;23:1334-40.
17. Marbet UA, Bianchi L, Meury U, et al. Long-term histological evaluation of the natural history and prognostic factors of alcoholic liver disease. J Hepatol 1987;4:364-72.
18. Kim WR, Therneau TM, Wiesner RH, et al. A revised natural history model for primary sclerosing cholangitis. Mayo Clin Proc 2000;75:688-94.

19. Mayo MJ, Parkes J, Adams-Huet B, et al. Prediction of clinical outcomes in primary biliary cirrhosis by serum enhanced liver fibrosis assay. Hepatology 2008;48:1549-57.
20. Carbone M, Sharp SJ, Flack S, et al. The UK-PBC Risk Scores: Derivation and validation of a scoring system for long-term prediction of end-stage liver disease in primary biliary cirrhosis. Hepatology 2015.
21. Ekstedt M, Hagstrom H, Nasr P, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. Hepatology 2015;61:1547-54.
22. EASL Recommendations on Treatment of Hepatitis C 2015. Journal of Hepatology;63:199-236.
23. EASL Clinical Practice Guidelines: Management of hepatitis C virus infection. Journal of Hepatology 2014;60:392-420.
24. Tapper EB, Lok AS. Use of Liver Imaging and Biopsy in Clinical Practice. N Engl J Med 2017;377:756-768.
25. 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.
26. 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.
27. Eisenberg E, Konopniki M, Veitsman E, et al. Prevalence and characteristics of pain induced by percutaneous liver biopsy. Anesth Analg 2003;96:1392-6, table of contents.
28. Grant A, Neuberger J. Guidelines on the use of liver biopsy in clinical practice. British Society of Gastroenterology. Gut 1999;45 Suppl 4:Iv1-iv11.
29. Piccinino F, Sagnelli E, Pasquale G, et al. Complications following percutaneous liver biopsy: A multicentre retrospective study on 68 276 biopsies. Journal of Hepatology 1986;2:165-173.
30. Rockey DC, Caldwell SH, Goodman ZD, et al. Liver biopsy. Hepatology 2009;49:1017-44.
31. Scheuer PJ. Liver biopsy size matters in chronic hepatitis: Bigger is better. Hepatology 2003;38:1356-1358.
32. Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology 2005;128:1898-906.
33. Garrido MC, Hubscher SG. Accuracy of staging in primary biliary cirrhosis. J Clin Pathol 1996;49:556-9.
34. Harrison RF, Hubscher SG. The spectrum of bile duct lesions in end-stage primary sclerosing cholangitis. Histopathology 1991;19:321-7.
35. Runge JH, Bohte AE, Verheij J, et al. Comparison of interobserver agreement of magnetic resonance elastography with histopathological staging of liver fibrosis. Abdom Imaging 2014;39:283-90.
36. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. Hepatology 2003;38:1449-57.
37. 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. Journal of Hepatology 2003;39:239-244.
38. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. The French METAVIR Cooperative Study Group. Hepatology 1994;20:15-20.

39. El-Badry AM, Breitenstein S, Jochum W, et al. Assessment of hepatic steatosis by expert pathologists: the end of a gold standard. Ann Surg 2009;250:691-7.
40. Hall AR, Green AC, Luong TV, et al. The use of guideline images to improve histological estimation of hepatic steatosis. Liver Int 2014;34:1414-27.
41. Wyatt J, Hubscher S, Bellamy C. Tissue pathways for liver biopsies for the investigation of medical disease and for focal lesions. Royal College of Pathologists guidelines 2014:1-29.
42. Calvaruso V, Dhillon AP, Tsochatzis E, et al. Liver collagen proportionate area predicts decompensation in patients with recurrent hepatitis C virus cirrhosis after liver transplantation. J Gastroenterol Hepatol 2012;27:1227-32.
43. Mehta SH, Lau B, Afdhal NH, et al. Exceeding the limits of liver histology markers. Journal of hepatology 2009;50:36-41.
44. Rossi E, Adams LA, Bulsara M, et al. Assessing liver fibrosis with serum marker models. Clin Biochem Rev 2007;28:3-10.
45. Mofrad P, Contos MJ, Haque M, et al. Clinical and histologic spectrum of nonalcoholic fatty liver disease associated with normal ALT values. Hepatology 2003;37:1286-92.
46. Williams AL, Hoofnagle JH. Ratio of serum aspartate to alanine aminotransferase in chronic hepatitis. Relationship to cirrhosis. Gastroenterology 1988;95:734-9.
47. Xiao G, Zhu S, Xiao X, et al. Comparison of laboratory tests, ultrasound, or MRE to detect fibrosis in patients with non-alcoholic fatty liver disease: A meta-analysis. Hepatology 2017.
48. 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.
49. 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.
50. 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.
51. 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.
52. Dufour DR. Assessment of liver fibrosis: Can serum become the sample of choice? Clin Chem 2005;51:1763-4.
53. Soresi M, Giannitrapani L, Cervello M, et al. Non invasive tools for the diagnosis of liver cirrhosis. World J Gastroenterol 2014;20:18131-50.
54. Myers RP, Tainturier MH, Ratziu V, et al. Prediction of liver histological lesions with biochemical markers in patients with chronic hepatitis B. J Hepatol 2003;39:222-30.
55. Eminler AT, Irak K, Ayyildiz T, et al. The relation between liver histopathology and GGT levels in viral hepatitis: more important in hepatitis B. Turk J Gastroenterol 2014;25:411-5.
56. Vardar R, Vardar E, Demiri S, et al. Is there any non-invasive marker replace the needle liver biopsy predictive for liver fibrosis, in patients with chronic hepatitis? Hepatogastroenterology 2009;56:1459-65.
57. Lemoine M, Shimakawa Y, Nayagam S, et al. The gamma-glutamyl transpeptidase to platelet ratio (GPR) predicts significant liver fibrosis and cirrhosis in patients with chronic HBV infection in West Africa. Gut 2016;65:1369-1376.

58. Li Q, Song J, Huang Y, et al. The Gamma-Glutamyl-Transpeptidase to Platelet Ratio Does not Show Advantages than APRI and Fib-4 in Diagnosing Significant Fibrosis and Cirrhosis in Patients With Chronic Hepatitis B: A Retrospective Cohort Study in China. Medicine 2016;95:e3372.
59. Parkes J, Guha IN, Roderick P, et al. Performance of serum marker panels for liver fibrosis in chronic hepatitis C. J Hepatol 2006;44:462-74.
60. Chou R, Wasson N. Blood Tests to Diagnose Fibrosis or Cirrhosis in Patients With Chronic Hepatitis C Virus InfectionA Systematic Review. Annals of Internal Medicine 2013;158:807-820.
61. Morling JR, Fallowfield JA, Guha IN, et al. Using non-invasive biomarkers to identify hepatic fibrosis in people with type 2 diabetes mellitus: the Edinburgh type 2 diabetes study. J Hepatol 2014;60:384-91.
62. Rosenberg WM, Voelker M, Thiel R, et al. Serum markers detect the presence of liver fibrosis: a cohort study. Gastroenterology 2004;127:1704-13.
63. 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.
64. 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.
65. 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.
66. Orasan OH, Ciulei G, Cozma A, et al. Hyaluronic acid as a biomarker of fibrosis in chronic liver diseases of different etiologies. Clujul Med 2016;89:24-31.
67. Arriazu E, Ge X, Leung T-M, et al. Signalling via the osteopontin and high mobility group box-1 axis drives the fibrogenic response to liver injury. Gut 2017;66:1123-1137.
68. Patouraux S, Bonnafous S, Voican CS, et al. The osteopontin level in liver, adipose tissue and serum is correlated with fibrosis in patients with alcoholic liver disease. PLoS One 2012;7:e35612.
69. Matsue Y, Tsutsumi M, Hayashi N, et al. Serum osteopontin predicts degree of hepatic fibrosis and serves as a biomarker in patients with hepatitis C virus infection. PLoS One 2015;10:e0118744.
70. Duarte-Salles T, Misra S, Stepien M, et al. Circulating Osteopontin and Prediction of Hepatocellular Carcinoma Development in a Large European Population. Cancer Prev Res (Phila) 2016.
71. Tsuchiya N, Sawada Y, Endo I, et al. Biomarkers for the early diagnosis of hepatocellular carcinoma. World J Gastroenterol 2015;21:10573-83.
72. Rychlikova J, Vecka M, Jachymova M, et al. Osteopontin as a discriminating marker for pancreatic cancer and chronic pancreatitis. Cancer Biomark 2016.
73. Hu ZD, Wei TT, Yang M, et al. Diagnostic value of osteopontin in ovarian cancer: a meta-analysis and systematic review. PLoS One 2015;10:e0126444.
74. Zduniak K, Ziolkowski P, Ahlin C, et al. Nuclear osteopontin-c is a prognostic breast cancer marker. Br J Cancer 2015;112:729-38.
75. van der Leeuw J, Beulens JW, van Dieren S, et al. Novel Biomarkers to Improve the Prediction of Cardiovascular Event Risk in Type 2 Diabetes Mellitus. J Am Heart Assoc 2016;5.

76. Jazwinski AB, Thompson AJ, Clark PJ, et al. Elevated serum CK18 levels in chronic hepatitis C patients are associated with advanced fibrosis but not steatosis. J Viral Hepat 2012;19:278-82.
77. Morling JR, Fallowfield JA, Williamson RM, et al. Non-invasive hepatic biomarkers (ELF and CK18) in people with type 2 diabetes: the Edinburgh type 2 diabetes study. Liver Int 2014;34:1267-77.
78. 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.
79. 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.
80. National Institute for Health and Care Excellence. Nonalcoholic fatty liver disease (NAFLD): assessment and management. NICE guidelines (NG49) 2016.
81. Yakoob R, Bozom IA, Thandassery RB, et al. Noninvasive biomarkers FibroTest and ActiTest versus liver biopsy in chronic hepatitis C patients: the Middle East experience. Ann Gastroenterol 2015;28:265-270.
82. Wang C, Cheng X, Meng C, et al. [Diagnostic value of FibroTest for liver fibrosis in patients with chronic hepatitis B]. Zhonghua Gan Zang Bing Za Zhi 2015;23:738-41.
83. Naveau S, Essoh BM, Ghinoiu M, et al. Comparison of Fibrotest and PGAA for the diagnosis of fibrosis stage in patients with alcoholic liver disease. Eur J Gastroenterol Hepatol 2014;26:404-11.
84. Dodd GD, 3rd, Baron RL, Oliver JH, 3rd, et al. Spectrum of imaging findings of the liver in end-stage cirrhosis: part I, gross morphology and diffuse abnormalities. AJR Am J Roentgenol 1999;173:1031-6.
85. Harbin WP, Robert NJ, Ferrucci JT, Jr. Diagnosis of cirrhosis based on regional changes in hepatic morphology: a radiological and pathological analysis. Radiology 1980;135:273-83.
86. Ito K, Mitchell DG, Gabata T, et al. Expanded gallbladder fossa: simple MR imaging sign of cirrhosis. Radiology 1999;211:723-6.
87. Ito K, Mitchell DG, Gabata T. Enlargement of hilar periportal space: a sign of early cirrhosis at MR imaging. J Magn Reson Imaging 2000;11:136-40.
88. Cross TJ, Joseph M, Fernando RA, et al. The liver to abdominal area ratio (LAAR): a novel imaging score for prognostication in cirrhosis. Aliment Pharmacol Ther 2013;38:1385-94.
89. Ito K, Mitchell DG, Hann HW, et al. Compensated cirrhosis due to viral hepatitis: using MR imaging to predict clinical progression. AJR Am J Roentgenol 1997;169:801-5.
90. Sarvazyan A, Hall TJ, Urban MW, et al. AN OVERVIEW OF ELASTOGRAPHY – AN EMERGING BRANCH OF MEDICAL IMAGING. Current medical imaging reviews 2011;7:255-282.
91. Armstrong MJ, Corbett C, Hodson J, et al. Operator training requirements and diagnostic accuracy of Fibroscan in routine clinical practice. Postgrad Med J 2013;89:685-92.
92. Sandrin L, Fourquet B, Hasquenoph JM, et al. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. Ultrasound Med Biol 2003;29:1705-13.

93. 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.
94. 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.
95. Shen F, Zheng RD, Shi JP, et al. Impact of skin capsular distance on the performance of controlled attenuation parameter in patients with chronic liver disease. Liver Int 2015.
96. A B. Utilization of FibroScan in Clinical Practice. 2014.
97. Mederacke I, Wursthorn K, Kirschner J, et al. Food intake increases liver stiffness in patients with chronic or resolved hepatitis C virus infection. Liver Int 2009;29:1500-6.
98. Petta S, Maida M, Macaluso FS, et al. The severity of steatosis influences liver stiffness measurement in patients with nonalcoholic fatty liver disease. Hepatology 2015;62:1101-10.
99. P. Eddowes MA, E. Tsochatzis, Q. Anstee, D. Sheridan, I. Guha, J. Cobbold, V. Paradis, P. Bedossa, P. Newsome. Performance of liver stiffness by FibroScan in a Large Prospective Multicenter UK Study: Applicability, Reliability, Diagnostic Performance and Influence of The Probe Type and Of Steatosis on the Liver Stiffness Measurement. AASLD: The Liver Meeting. Washington DC, 2017.
100. Cardoso AC, Carvalho-Filho RJ, Stern C, et al. Direct comparison of diagnostic performance of transient elastography in patients with chronic hepatitis B and chronic hepatitis C. Liver Int 2012;32:612-21.
101. Nahon P, Kettaneh A, Tengher-Barna I, et al. Assessment of liver fibrosis using transient elastography in patients with alcoholic liver disease. J Hepatol 2008;49:1062-8.
102. Corpechot C, El Naggar A, Poujol-Robert A, et al. Assessment of biliary fibrosis by transient elastography in patients with PBC and PSC. Hepatology 2006;43:1118-1124.
103. Wong VW-S, Vergniol J, Wong GL-H, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. Hepatology 2010;51:454-462.
104. 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.
105. 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). Journal of Hepatology;53:1013-1021.
106. Tsochatzis EA, Gurusamy KS, Ntaoula S, et al. Elastography for the diagnosis of severity of fibrosis in chronic liver disease: a meta-analysis of diagnostic accuracy. J Hepatol 2011;54:650-9.
107. Roulot D, Costes JL, Buyck JF, et al. Transient elastography as a screening tool for liver fibrosis and cirrhosis in a community-based population aged over 45 years. Gut 2011;60:977-84.
108. Buechter M, Kahraman A, Manka P, et al. Spleen and Liver Stiffness Is Positively Correlated with the Risk of Esophageal Variceal Bleeding. Digestion 2016;94:138-144.

109. Abraldes JG, Bureau C, Stefanescu H, et al. Noninvasive tools and risk of clinically significant portal hypertension and varices in compensated cirrhosis: The "Anticipate" study. Hepatology 2016;64:2173-2184.
110. You MW, Kim KW, Pyo J, et al. A Meta-analysis for the Diagnostic Performance of Transient Elastography for Clinically Significant Portal Hypertension. Ultrasound Med Biol 2017;43:59-68.
111. Berzigotti A. Non-invasive evaluation of portal hypertension using ultrasound elastography. J Hepatol 2017;67:399-411.
112. Cardenas A, Mendez-Bocanegra A. Report of the Baveno VI Consensus Workshop. Ann Hepatol 2016;15:289-90.
113. de Franchis R. Expanding consensus in portal hypertension: Report of the Baveno VI Consensus Workshop: Stratifying risk and individualizing care for portal hypertension. J Hepatol 2015;63:743-52.
114. Maurice JB, Brodkin E, Arnold F, et al. Validation of the Baveno VI criteria to identify low risk cirrhotic patients not requiring endoscopic surveillance for varices. J Hepatol 2016;65:899-905.
115. Bota S, Herkner H, Sporea I, et al. Meta-analysis: ARFI elastography versus transient elastography for the evaluation of liver fibrosis. Liver International 2013;33:1138-1147.
116. Lupsor M, Badea R, Stefanescu H, et al. Performance of a new elastographic method (ARFI technology) compared to unidimensional transient elastography in the noninvasive assessment of chronic hepatitis C. Preliminary results. J Gastrointestin Liver Dis 2009;18:303-10.
117. D'Onofrio M, Crosara S, De Robertis R, et al. Acoustic radiation force impulse of the liver. World J Gastroenterol 2013;19:4841-9.
118. Nierhoff J, Chavez Ortiz AA, Herrmann E, et al. The efficiency of acoustic radiation force impulse imaging for the staging of liver fibrosis: a meta-analysis. Eur Radiol 2013;23:3040-53.
119. Tachi Y, Hirai T, Kojima Y, et al. Liver stiffness measurement using acoustic radiation force impulse elastography in hepatitis C virus-infected patients with a sustained virological response. Aliment Pharmacol Ther 2016;44:346-55.
120. Attia D, Bantel H, Lenzen H, et al. Liver stiffness measurement using acoustic radiation force impulse elastography in overweight and obese patients. Aliment Pharmacol Ther 2016;44:366-79.
121. Kiani A, Brun V, Laine F, et al. Acoustic radiation force impulse imaging for assessing liver fibrosis in alcoholic liver disease. World J Gastroenterol 2016;22:4926-35.
122. Godfrey EM, Mannelli L, Griffin N, et al. Magnetic resonance elastography in the diagnosis of hepatic fibrosis. Semin Ultrasound CT MR 2013;34:81-8.
123. Park CC, Nguyen P, Hernandez C, et al. Magnetic Resonance Elastography vs Transient Elastography in Detection of Fibrosis and Noninvasive Measurement of Steatosis in Patients With Biopsy-Proven Nonalcoholic Fatty Liver Disease. Gastroenterology 2017;152:598-607 e2.
124. Cui J, Ang B, Haufe W, et al. Comparative diagnostic accuracy of magnetic resonance elastography versus eight clinical prediction rules for non-invasive diagnosis of advanced fibrosis in biopsy-proven nonalcoholic fatty liver disease: a prospective study. Alimentary pharmacology & therapeutics 2015;41:1271-1280.

125. Tan CH, Venkatesh SK. Magnetic Resonance Elastography and Other Magnetic Resonance Imaging Techniques in Chronic Liver Disease: Current Status and Future Directions. Gut and Liver 2016;10:672-686.
126. Loomba R, Wolfson T, Ang B, et al. Magnetic Resonance Elastography Predicts Advanced Fibrosis in Patients With Nonalcoholic Fatty Liver Disease: A Prospective Study. Hepatology (Baltimore, Md.) 2014;60:1920-1928.
127. Singh S, Venkatesh SK, Loomba R, et al. Magnetic Resonance Elastography for Staging Liver Fibrosis in Non-alcoholic Fatty Liver Disease: A Diagnostic Accuracy Systematic Review and Individual Participant Data Pooled Analysis. European radiology 2016;26:1431-1440.
128. Singh S, Venkatesh SK, Wang Z, et al. Diagnostic Performance of Magnetic Resonance Elastography in Staging Liver Fibrosis: A Systematic Review and Meta-analysis of Individual Participant Data. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2015;13:440-451.e6.
129. Wang QB, Zhu H, Liu HL, et al. Performance of magnetic resonance elastography and diffusion-weighted imaging for the staging of hepatic fibrosis: A meta-analysis. Hepatology 2012;56:239-47.
130. Singh S, Muir AJ, Dieterich DT, et al. American Gastroenterological Association Institute Technical Review on the Role of Elastography in Chronic Liver Diseases. Gastroenterology 2017;152:1544-1577.
131. Huwart L, Sempoux C, Vicaut E, et al. Magnetic resonance elastography for the noninvasive staging of liver fibrosis. Gastroenterology 2008;135:32-40.
132. Bonekamp S, Kamel I, Solga S, et al. Can imaging modalities diagnose and stage hepatic fibrosis and cirrhosis accurately? J Hepatol 2009;50:17-35.
133. Venkatesh SK, Yin M, Takahashi N, et al. Non-invasive detection of liver fibrosis: MR imaging features vs. MR elastography. Abdom Imaging 2015;40:766-75.
134. Chen J, Talwalkar JA, Yin M, et al. Early detection of nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease by using MR elastography. Radiology 2011;259:749-56.
135. Kim D, Kim WR, Talwalkar JA, et al. Advanced fibrosis in nonalcoholic fatty liver disease: noninvasive assessment with MR elastography. Radiology 2013;268:411-9.
136. Chen J, Yin M, Talwalkar JA, et al. Assessment of Diagnostic Performance of MR Elastography (MRE) and Vibration-Controlled Transient Elastography (VCTE) for Detecting Hepatic Fibrosis in Patients with Severe to Morbid Obesity. Radiology 2017;283:418-428.
137. Low G, Kruse SA, Lomas DJ. General review of magnetic resonance elastography. World J Radiol 2016;8:59-72.
138. Chandarana H, Taouli B. Diffusion and perfusion imaging of the liver. Eur J Radiol 2010;76:348-58.
139. Taouli B, Koh DM. Diffusion-weighted MR imaging of the liver. Radiology 2010;254:47-66.
140. Mannelli L, Bhargava P, Osman SF, et al. Diffusion-weighted imaging of the liver: a comprehensive review. Curr Probl Diagn Radiol 2013;42:77-83.
141. Anderson SW, Jara H, Ozonoff A, et al. Effect of disease progression on liver apparent diffusion coefficient and T2 values in a murine model of hepatic fibrosis at 11.7 Tesla MRI. J Magn Reson Imaging 2012;35:140-6.

142. Koinuma M, Ohashi I, Hanafusa K, et al. Apparent diffusion coefficient measurements with diffusion-weighted magnetic resonance imaging for evaluation of hepatic fibrosis. J Magn Reson Imaging 2005;22:80-5.
143. Do RK, Chandarana H, Felker E, et al. Diagnosis of liver fibrosis and cirrhosis with diffusion-weighted imaging: value of normalized apparent diffusion coefficient using the spleen as reference organ. AJR Am J Roentgenol 2010;195:671-6.
144. Taouli B, Chouli M, Martin AJ, et al. Chronic hepatitis: role of diffusion-weighted imaging and diffusion tensor imaging for the diagnosis of liver fibrosis and inflammation. J Magn Reson Imaging 2008;28:89-95.
145. Taouli B, Tolia AJ, Losada M, et al. Diffusion-weighted MRI for quantification of liver fibrosis: preliminary experience. AJR Am J Roentgenol 2007;189:799-806.
146. Zhu NY, Chen KM, Chai WM, et al. Feasibility of diagnosing and staging liver fibrosis with diffusion weighted imaging. Chin Med Sci J 2008;23:183-6.
147. Bharwani N, Koh DM. Diffusion-weighted imaging of the liver: an update. Cancer Imaging 2013;13:171-85.
148. Soylu A, Kilickesmez O, Poturoglu S, et al. Utility of diffusion-weighted MRI for assessing liver fibrosis in patients with chronic active hepatitis. Diagn Interv Radiol 2010;16:204-8.
149. Van Beers BE, Daire JL, Garteiser P. New imaging techniques for liver diseases. J Hepatol 2014.
150. Hagiwara M, Rusinek H, Lee VS, et al. Advanced liver fibrosis: diagnosis with 3D whole-liver perfusion MR imaging--initial experience. Radiology 2008;246:926-34.
151. Chen BB, Hsu CY, Yu CW, et al. Dynamic contrast-enhanced magnetic resonance imaging with Gd-EOB-DTPA for the evaluation of liver fibrosis in chronic hepatitis patients. Eur Radiol 2012;22:171-80.
152. Haider MA, Farhadi FA, Milot L. Hepatic perfusion imaging: concepts and application. Magn Reson Imaging Clin N Am 2010;18:465-75, x.
153. Kanematsu M, Goshima S, Watanabe H, et al. Diffusion/perfusion MR imaging of the liver: practice, challenges, and future. Magn Reson Med Sci 2012;11:151-61.
154. Taouli B, Ehman RL, Reeder SB. Advanced MRI methods for assessment of chronic liver disease. AJR Am J Roentgenol 2009;193:14-27.
155. Petitclerc L, Sebastiani G, Gilbert G, et al. Liver fibrosis: Review of current imaging and MRI quantification techniques. J Magn Reson Imaging 2016.
156. Palaniyappan N, Cox E, Bradley C, et al. Non-invasive assessment of portal hypertension using quantitative magnetic resonance imaging. J Hepatol 2016;65:1131-1139.
157. Jellis CL, Kwon DH. Myocardial T1 mapping: modalities and clinical applications. Cardiovasc Diagn Ther 2014;4:126-37.
158. Stark DD, Bass NM, Moss AA, et al. Nuclear magnetic resonance imaging of experimentally induced liver disease. Radiology 1983;148:743-51.
159. Thomsen C, Christoffersen P, Henriksen O, et al. Prolonged T1 in patients with liver cirrhosis: An in vivo MRI study. Magnetic Resonance Imaging 1990;8:599-604.
160. Keevil SF, Alstead EM, Dolke G, et al. Non-invasive assessment of diffuse liver disease by in vivo measurement of proton nuclear magnetic resonance relaxation times at 0.08 T. The British Journal of Radiology 1994;67:1083-1087.
161. Thomsen C. Quantitative magnetic resonance methods for in vivo investigation of the human liver and spleen. Technical aspects and preliminary clinical results. Acta Radiol Suppl 1996;401:1-34.

162. Chow AM, Gao DS, Fan SJ, et al. Measurement of liver T(1) and T(2) relaxation times in an experimental mouse model of liver fibrosis. J Magn Reson Imaging 2012;36:152-8.
163. Banerjee R, Pavlides M, Tunnicliffe EM, et al. Multiparametric magnetic resonance for the non-invasive diagnosis of liver disease. J Hepatol 2014;60:69-77.
164. Hoad CL, Palaniyappan N, Kaye P, et al. A study of T(1) relaxation time as a measure of liver fibrosis and the influence of confounding histological factors. NMR Biomed 2015;28:706-14.
165. Pavlides M, Banerjee R, Sellwood J, et al. Multi-parametric magnetic resonance imaging predicts clinical outcomes in patients with chronic liver disease. J Hepatol 2015.
166. Blink EJ. Basic MRI Physics, 2004.
167. Henninger B, Kremser C, Rauch S, et al. Evaluation of MR imaging with T1 and T2* mapping for the determination of hepatic iron overload. Eur Radiol 2012;22:2478-86.
168. Roujol S, Weingartner S, Foppa M, et al. Accuracy, precision, and reproducibility of four T1 mapping sequences: a head-to-head comparison of MOLLI, ShMOLLI, SASHA, and SAPPHIRE. Radiology 2014;272:683-9.
169. McPherson S, Jonsson JR, Cowin GJ, et al. Magnetic resonance imaging and spectroscopy accurately estimate the severity of steatosis provided the stage of fibrosis is considered. J Hepatol 2009;51:389-97.
170. Wu CH, Ho MC, Jeng YM, et al. Quantification of Hepatic Steatosis: A Comparison of the Accuracy among Multiple Magnetic Resonance Techniques. J Gastroenterol Hepatol 2013.
171. Bannas P, Kramer H, Hernando D, et al. Quantitative magnetic resonance imaging of hepatic steatosis: Validation in ex vivo human livers. Hepatology 2015;62:1444-1455.
172. Idilman IS, Keskin O, Celik A, et al. A comparison of liver fat content as determined by magnetic resonance imaging-proton density fat fraction and MRS versus liver histology in non-alcoholic fatty liver disease. Acta Radiologica 2015.
173. Elster A. Dixon Method. Questions and Answers in MRI, 2007.
174. Moneta GL, Taylor DC, Helton WS, et al. Duplex ultrasound measurement of postprandial intestinal blood flow: effect of meal composition. Gastroenterology 1988;95:1294-301.
175. Arena U, Lupsor Platon M, Stasi C, et al. Liver stiffness is influenced by a standardized meal in patients with chronic hepatitis C virus at different stages of fibrotic evolution. Hepatology 2013;58:65-72.
176. Berzigotti A, De Gottardi A, Vukotic R, et al. Effect of meal ingestion on liver stiffness in patients with cirrhosis and portal hypertension. PLoS One 2013;8:e58742.
177. Heye T, Yang S-R, Bock M, et al. MR relaxometry of the liver: significant elevation of T1 relaxation time in patients with liver cirrhosis. European Radiology 2012;22:1224-1232.
178. Bangert; SK, Marshall WJ. Clinical biochemistry: metabolic and clinical aspects. Clinical biochemistry: metabolic and clinical aspects. Philadelphia: Churchill Livingstone/Elsevier, 2008:19.
179. America RSoN. Fat in Organs and Blood May Increase Risk of Osteoporosis. RSNA.org: RSNA, 2013.
180. Hamilton G, Yokoo T, Bydder M, et al. In vivo characterization of the liver fat (1)H MR spectrum. NMR Biomed 2011;24:784-90.

181. Wilson M, Reynolds G, Kauppinen RA, et al. A constrained least-squares approach to the automated quantitation of in vivo 1H magnetic resonance spectroscopy data. Magnetic Resonance in Medicine 2011;65:1-12.
182. Boursier J, Zarski J-P, de Ledinghen V, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. Hepatology 2013;57:1182-1191.
183. Bedossa P. Intraobserver and Interobserver Variations in Liver Biopsy Interpretation in Patients with Chronic Hepatitis C. Hepatology 1994;20:15-20.
184. Westin J, Lagging LM, Wejstål R, et al. Interobserver study of liver histopathology using the Ishak score in patients with chronic hepatitis C virus infection. Liver 1999;19:183-187.
185. Younossi ZM, Gramlich T, Liu YC, et al. Nonalcoholic fatty liver disease: assessment of variability in pathologic interpretations. Mod Pathol 1998;11:560-5.
186. Germani G, Burroughs AK, Dhillon AP. The relationship between liver disease stage and liver fibrosis: a tangled web. Histopathology 2010;57:773-84.
187. Mozes FE, Tunnicliffe EM, Pavlides M, et al. Influence of fat on liver T measurements using modified Look-Locker inversion recovery (MOLLI) methods at 3T. J Magn Reson Imaging 2016.
188. Brunt EM, Janney CG, Di Bisceglie AM, et al. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. Am J Gastroenterol 1999;94:2467-74.
189. Scheuer PJ, Williams R, Muir AR. Hepatic pathology in relatives of patients with haemochromatosis. J Pathol Bacteriol 1962;84:53-64.
190. Standish RA, Cholongitas E, Dhillon A, et al. An appraisal of the histopathological assessment of liver fibrosis. Gut 2006;55:569-78.
191. Hall A, Germani G, Isgro G, et al. Fibrosis distribution in explanted cirrhotic livers. Histopathology 2012;60:270-7.
192. Addison T. Observations on fatty degeneration of the liver. Guys Hospital Reports 1836;1:476-485.
193. Goldberg M, Thompson CM. Acute Fatty Metamorphosis of the Liver. Annals of Internal Medicine 1961;55:416-432.
194. Thaler H. Die Fettleber und ihre pathogenetische Beziehung zur Lebercirrhose. Virchows Archiv für pathologische Anatomie und Physiologie und für klinische Medizin 1962;335:180-210.
195. Adler MS, F. Fatty liver hepatitis and cirrhosis in obese patients. The American Journal of Medicine 1979;67:A56.
196. Ludwig J, Viggiano TR, McGill DB, et al. Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease. Mayo Clin Proc 1980;55:434-8.
197. Matteoni CA, Younossi ZM, Gramlich T, et al. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. Gastroenterology 1999;116:1413-9.
198. 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.
199. Dowman JK, Tomlinson JW, Newsome PN. Systematic review: the diagnosis and staging of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis. Alimentary Pharmacology & Therapeutics 2011;33:525-540.
200. Machado M, Marques-Vidal P, Cortez-Pinto H. Hepatic histology in obese patients undergoing bariatric surgery. J Hepatol 2006;45:600-6.

201. 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.
202. Dam-Larsen S, Franzmann M, Andersen IB, et al. Long term prognosis of fatty liver: risk of chronic liver disease and death. Gut 2004;53:750-5.
203. 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.
204. Teli MR, James OFW, Burt AD, et al. The natural history of nonalcoholic fatty liver: A follow-up study. Hepatology 1995;22:1714-1719.
205. McCullough AJ. The clinical features, diagnosis and natural history of nonalcoholic fatty liver disease. Clin Liver Dis 2004;8:521-33, viii.
206. Argo CK, Northup PG, Al-Osaimi AMS, et al. Systematic review of risk factors for fibrosis progression in non-alcoholic steatohepatitis. Journal of Hepatology;51:371-379.
207. Byrne CD, Targher G. Time to Replace Assessment of Liver Histology With MR-Based Imaging Tests to Assess Efficacy of Interventions for Nonalcoholic Fatty Liver Disease. Gastroenterology 2016;150:7-10.
208. Boutari C, Lefkos P, Athyros VG, et al. Nonalcoholic fatty liver disease vs. nonalcoholic steatohepatitis: pathological and clinical implications. Curr Vasc Pharmacol 2017.
209. Townsend SA, Newsome PN. Review article: new treatments in non-alcoholic fatty liver disease. Aliment Pharmacol Ther 2017.
210. Machado MV, Cortez-Pinto H. Non-invasive diagnosis of non-alcoholic fatty liver disease. A critical appraisal. J Hepatol 2013;58:1007-1019.
211. Kwon H, Pessin JE. Adipokines Mediate Inflammation and Insulin Resistance. Frontiers in Endocrinology 2013;4:71.
212. Salman A, Hegazy M, AbdElfadl S. Combined Adiponectin Deficiency and Resistance in Obese Patients: Can It Solve Part of the Puzzle in Nonalcoholic Steatohepatitis. Open Access Maced J Med Sci 2015;3:298-302.
213. Musso G, Gambino R, Durazzo M, et al. Adipokines in NASH: Postprandial lipid metabolism as a link between adiponectin and liver disease. Hepatology 2005;42:1175-1183.
214. Jamali R, Arj A, Razavizade M, et al. Prediction of Nonalcoholic Fatty Liver Disease Via a Novel Panel of Serum Adipokines. Medicine (Baltimore) 2016;95:e2630.
215. Shen J, Chan HL, Wong GL, et al. Non-invasive diagnosis of non-alcoholic steatohepatitis by combined serum biomarkers. J Hepatol 2012;56:1363-70.
216. Chen J, Zhu Y, Zheng Q, et al. Serum cytokeratin-18 in the diagnosis of non-alcoholic steatohepatitis: A meta-analysis. Hepatol Res 2013.
217. 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 2013.
218. Brunt EM, Kleiner DE, Wilson LA, et al. The NAS and The Histopathologic Diagnosis in NAFLD: Distinct Clinicopathologic Meanings. Hepatology (Baltimore, Md.) 2011;53:810-820.
219. 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.

220. Sanyal AJ, Neuschwander-Tetri BA, Tonascia J. End Points Must Be Clinically Meaningful for Drug Development in Nonalcoholic Fatty Liver Disease. Gastroenterology 2016;150:11-13.
221. 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.
222. Jaruvongvanich V, Wijarnpreecha K, Ungprasert P. The utility of NAFLD fibrosis score for prediction of mortality among patients with nonalcoholic fatty liver disease: A systematic review and meta-analysis of cohort study. Clin Res Hepatol Gastroenterol 2017.
223. 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.
224. Banerjee R, Piechnik S, Robson M, et al. Multi-Parametric Magnetic Resonance Diagnosis & Staging of Liver Disease (US2014330106), 2014.
225. Pavlides M, Banerjee R, Tunnicliffe EM, et al. Multiparametric magnetic resonance imaging for the assessment of non-alcoholic fatty liver disease severity. Liver Int 2017;37:1065-1073.
226. Noureddin M, Lam J, Peterson MR, et al. Utility of magnetic resonance imaging versus histology for quantifying changes in liver fat in nonalcoholic fatty liver disease trials. Hepatology 2013;58:1930-40.
227. Tang A, Tan J, Sun M, et al. Nonalcoholic fatty liver disease: MR imaging of liver proton density fat fraction to assess hepatic steatosis. Radiology 2013;267:422-31.
228. Permutt Z, Le TA, Peterson MR, et al. Correlation between liver histology and novel magnetic resonance imaging in adult patients with non-alcoholic fatty liver disease - MRI accurately quantifies hepatic steatosis in NAFLD. Aliment Pharmacol Ther 2012;36:22-9.
229. St Pierre TG, Clark PW, Chua-anusorn W, et al. Noninvasive measurement and imaging of liver iron concentrations using proton magnetic resonance. Blood 2005;105:855.
230. Younossi ZM, Stepanova M, Rafiq N, et al. Pathologic criteria for nonalcoholic steatohepatitis: Interprotocol agreement and ability to predict liver-related mortality. Hepatology 2011;53:1874-1882.
231. Castera L. Non-invasive diagnosis of steatosis and fibrosis. Diabetes Metab 2008;34:674-9.
232. Fedchuk L, Nascimbeni F, Pais R, et al. Performance and limitations of steatosis biomarkers in patients with nonalcoholic fatty liver disease. Aliment Pharmacol Ther 2014;40:1209-22.
233. 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.
234. Hamilton G, Middleton MS, Bydder M, et al. Effect of PRESS and STEAM sequences on magnetic resonance spectroscopic liver fat quantification. J Magn Reson Imaging 2009;30:145-52.
235. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. Hepatology 2012;55:2005-23.

236. 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.
237. Petta S, Wong VW, Camma C, et al. Serial combination of non-invasive tools improves the diagnostic accuracy of severe liver fibrosis in patients with NAFLD. Aliment Pharmacol Ther 2017.
238. Cui J, Ang B, Haufe W, et al. Comparative diagnostic accuracy of magnetic resonance elastography vs. eight clinical prediction rules for non-invasive diagnosis of advanced fibrosis in biopsy-proven non-alcoholic fatty liver disease: a prospective study. Aliment Pharmacol Ther 2015;41:1271-80.
239. Germani G, Hytiroglou P, Fotiadu A, et al. Assessment of fibrosis and cirrhosis in liver biopsies: an update. Semin Liver Dis 2011;31:82-90.
240. Peverill W, Powell LW, Skoien R. Evolving concepts in the pathogenesis of NASH: beyond steatosis and inflammation. Int J Mol Sci 2014;15:8591-638.
241. Puri P, Baillie RA, Wiest MM, et al. A lipidomic analysis of nonalcoholic fatty liver disease. Hepatology 2007;46:1081-90.
242. Howe FA, Opstad KS. 1H MR spectroscopy of brain tumours and masses. NMR in Biomedicine 2003;16:123-131.
243. Ren J, Dimitrov I, Sherry AD, et al. Composition of adipose tissue and marrow fat in humans by 1H NMR at 7 Tesla. Journal of Lipid Research 2008;49:2055-2062.
244. Shen ZW, Cao Z, You KZ, et al. Quantification of choline concentration following liver cell apoptosis using (1)H magnetic resonance spectroscopy. World J Gastroenterol 2012;18:1130-6.
245. Kotronen A, Seppänen-Laakso T, Westerbacka J, et al. Comparison of Lipid and Fatty Acid Composition of the Liver, Subcutaneous and Intra-abdominal Adipose Tissue, and Serum. Obesity 2010;18:937-944.
246. Neuschwander-Tetri BA. Hepatic lipotoxicity and the pathogenesis of nonalcoholic steatohepatitis: the central role of nontriglyceride fatty acid metabolites. Hepatology 2010;52:774-88.
247. Pagadala M, Kasumov T, McCullough AJ, et al. Role of ceramides in nonalcoholic fatty liver disease. Trends Endocrinol Metab 2012;23:365-71.
248. Araya J, Rodrigo R, Videla LA, et al. Increase in long-chain polyunsaturated fatty acid n - 6/n - 3 ratio in relation to hepatic steatosis in patients with non-alcoholic fatty liver disease. Clin Sci (Lond) 2004;106:635-43.
249. Yamada K, Mizukoshi E, Sunagozaka H, et al. Characteristics of hepatic fatty acid compositions in patients with nonalcoholic steatohepatitis. Liver Int 2014.
250. Gorden DL, Myers DS, Ivanova PT, et al. Biomarkers of NAFLD progression: a lipidomics approach to an epidemic. J Lipid Res 2015;56:722-36.
251. Leporq B, Lambert SA, Ronot M, et al. Quantification of the triglyceride fatty acid composition with 3.0 T MRI. NMR Biomed 2014;27:1211-21.
252. van Werven JR, Marsman HA, Nederveen AJ, et al. Hepatic lipid composition analysis using 3.0-T MR spectroscopy in a steatotic rat model. Magn Reson Imaging 2012;30:112-21.
253. Hamilton G, Schlein AN, Middleton MS, et al. In vivo triglyceride composition of abdominal adipose tissue measured by 1 H MRS at 3T. J Magn Reson Imaging 2016.
254. Artz NS, Haufe WM, Hooker CA, et al. Reproducibility of MR-based liver fat quantification across field strength: Same-day comparison between 1.5T and 3T in obese subjects. Journal of Magnetic Resonance Imaging 2015;42:811-817.

255. Tapia G, Valenzuela R, Espinosa A, et al. N-3 long-chain PUFA supplementation prevents high fat diet induced mouse liver steatosis and inflammation in relation to PPAR-α upregulation and NF-κB DNA binding abrogation. Molecular Nutrition & Food Research 2014;58:1333-1341.
256. Araya J, Rodrigo R, Videla LA. Increase in long-chain polyunsaturated fatty acid n - 6/n - 3 ratio in relation to hepatic steatosis in patients with non-alcoholic fatty liver disease. Clin Sci (Lond) 2004;106.
257. Juárez-Hernández E, Chávez-Tapia NC, Uribe M, et al. Role of bioactive fatty acids in nonalcoholic fatty liver disease. Nutrition Journal 2016;15:72.
258. Jump DB, Depner CM, Tripathy S. Potential for dietary omega-3 fatty acids to prevent nonalcoholic fatty liver disease and reduce the risk of primary liver cancer. Adv Nutr 2015;6.
259. Sherriff JL, O'Sullivan TA, Properzi C, et al. Choline, Its Potential Role in Nonalcoholic Fatty Liver Disease, and the Case for Human and Bacterial Genes. Advances in Nutrition 2016;7:5-13.
260. Walker AK, Jacobs RL, Watts JL, et al. A conserved SREBP-1/phosphatidylcholine feedback circuit regulates lipogenesis in metazoans. Cell 2011;147:840-852.
261. Li Z, Agellon LB, Allen TM, et al. The ratio of phosphatidylcholine to phosphatidylethanolamine influences membrane integrity and steatohepatitis. Cell Metab 2006;3:321-31.
262. Strobel K, van den Hoff J, Pietzsch J. Localized proton magnetic resonance spectroscopy of lipids in adipose tissue at high spatial resolution in mice in vivo. J Lipid Res 2008;49:473-80.
263. van Werven JR, Schreuder TC, Nederveen AJ, et al. Hepatic unsaturated fatty acids in patients with non-alcoholic fatty liver disease assessed by 3.0T MR spectroscopy. Eur J Radiol 2010;75:e102-7.
264. Wilson M, Reynolds G, Kauppinen RA, et al. A constrained least-squares approach to the automated quantitation of in vivo (1)H magnetic resonance spectroscopy data. Magn Reson Med 2011;65:1-12.
265. Flintham RE, P. J.; Semple, S.; McDonald,N.; Fallowfield, J.; Kendall, T.; Hübscher, S. G.; Newsome, P. N.; Hirschfield, G. M.; Davies, N. P. Non-invasive quantification and characterisation of liver fat in non-alcoholic fatty liver disease (NAFLD) using automated analysis of MRS correlated with histology, In ISMRM 24th Annual Meeting & Exhibition, Singapore, 2016.
266. Hirschfield GM, Karlsen TH, Lindor KD, et al. Primary sclerosing cholangitis. The Lancet 2013;382:1587-1599.
267. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: A systematic review. Journal of Hepatology 2012;56:1181-1188.
268. Molodecky NA, Kareemi H, Parab R, et al. Incidence of primary sclerosing cholangitis: A systematic review and meta-analysis. Hepatology 2011;53:1590-1599.
269. Ellinghaus D, Folseraas T, Holm K, et al. Genome-wide association analysis in primary sclerosing cholangitis and ulcerative colitis identifies risk loci at GPR35 and TCF4. Hepatology 2013;58:1074-83.
270. Grant AJ, Lalor PF, Salmi M, et al. Homing of mucosal lymphocytes to the liver in the pathogenesis of hepatic complications of inflammatory bowel disease. The Lancet 2002;359:150-157.
271. Trivedi PJ, Adams DH. Gut-liver immunity. J Hepatol 2016;64:1187-9.

272. Dave M, Elmunzer BJ, Dwamena BA, et al. Primary Sclerosing Cholangitis: Meta-Analysis of Diagnostic Performance of MR Cholangiopancreatography. Radiology 2010;256:387-396.
273. Burak K, Angulo P, Pasha TM, et al. Incidence and Risk Factors for Cholangiocarcinoma in Primary Sclerosing Cholangitis. The American Journal of Gastroenterology 2004;99:523-526.
274. European Association for the Study of the L. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. J Hepatol 2009;51:237-67.
275. Chapman R, Fevery J, Kalloo A, et al. Diagnosis and management of primary sclerosing cholangitis. Hepatology 2010;51:660-78.
276. Lindor KD, Kowdley KV, Harrison ME, et al. ACG Clinical Guideline: Primary Sclerosing Cholangitis. Am J Gastroenterol 2015;110:646-59; quiz 660.
277. Lindström L, Hultcrantz R, Boberg KM, et al. Association Between Reduced Levels of Alkaline Phosphatase and Survival Times of Patients With Primary Sclerosing Cholangitis. Clinical Gastroenterology and Hepatology 2013;11:841-846.
278. Watanabe T, Hirano K, Tada M, et al. Short-term prognostic factors for primary sclerosing cholangitis. J Hepatobiliary Pancreat Sci 2015;22:486-90.
279. Rupp C, Rössler A, Halibasic E, et al. Reduction in alkaline phosphatase is associated with longer survival in primary sclerosing cholangitis, independent of dominant stenosis. Alimentary Pharmacology & Therapeutics 2014;40:1292-1301.
280. Al Mamari S, Djordjevic J, Halliday JS, et al. Improvement of serum alkaline phosphatase to <1.5 upper limit of normal predicts better outcome and reduced risk of cholangiocarcinoma in primary sclerosing cholangitis. Journal of Hepatology 2013;58:329-334.
281. Hilscher M, Enders FB, Carey EJ, et al. Alkaline phosphatase normalization is a biomarker of improved survival in primary sclerosing cholangitis. Ann Hepatol 2016;15:246-53.
282. Broome U, Olsson R, Loof L, et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. Gut 1996;38:610-615.
283. Dickson ER, Murtaugh PA, Wiesner RH, et al. Primary sclerosing cholangitis: Refinement and validation of survival models. Gastroenterology 1992;103:1893-1901.
284. Farrant JM, Hayllar KM, Wilkinson ML, et al. Natural history and prognostic variables in primary sclerosing cholangitis. Gastroenterology 1991;100:1710-1717.
285. Wiesner RH, Grambsch PM, Dickson ER, et al. Primary sclerosing cholangitis: Natural history, prognostic factors and survival analysis. Hepatology 1989;10:430-436.
286. Kim WR, Poterucha JJ, Wiesner RH, et al. The relative role of the child-pugh classification and the mayo natural history model in the assessment of survival in patients with primary sclerosing cholangitis. Hepatology 1999;29:1643-1648.
287. Ponsioen CY, Chapman RW, Chazouilleres O, et al. Surrogate endpoints for clinical trials in primary sclerosing cholangitis: Review and results from an International PSC Study Group consensus process. Hepatology 2016;63:1357-67.
288. Floreani A, Zancan L, Melis A, et al. Primary sclerosing cholangitis (PSC): clinical, laboratory and survival analysis in children and adults. Liver 1999;19:228-33.
289. Malinchoc M, Kamath PS, Gordon FD, et al. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. Hepatology 2000;31:864-71.

290. Kamath PS, Wiesner RH, Malinchoc M, et al. A model to predict survival in patients with end-stage liver disease. Hepatology 2001;33:464-70.
291. Kamath PS, Kim WR. The model for end-stage liver disease (MELD). Hepatology 2007;45:797-805.
292. Hoffmann K, Hinz U, Hillebrand N, et al. The MELD score predicts the short-term and overall survival after liver transplantation in patients with primary sclerosing cholangitis or autoimmune liver diseases. Langenbecks Arch Surg 2014;399:1001-9.
293. Vitellas KM, Keogan MT, Freed KS, et al. Radiologic manifestations of sclerosing cholangitis with emphasis on MR cholangiopancreatography. Radiographics 2000;20:959-75; quiz 1108-9, 1112.
294. Olsson RG, Asztely MS. Prognostic value of cholangiography in primary sclerosing cholangitis. Eur J Gastroenterol Hepatol 1995;7:251-4.
295. Majoie CB, Reeders JW, Sanders JB, et al. Primary sclerosing cholangitis: a modified classification of cholangiographic findings. AJR Am J Roentgenol 1991;157:495-7.
296. Rajaram R, Ponsioen CY, Majoie CB, et al. Evaluation of a modified cholangiographic classification system for primary sclerosing cholangitis. Abdom Imaging 2001;26:43-7.
297. Ponsioen CY, Vrouenraets SME, Prawirodirdjo W, et al. Natural history of primary sclerosing cholangitis and prognostic value of cholangiography in a Dutch population. Gut 2002;51:562-566.
298. Ponsioen CY, Reitsma JB, Boberg KM, et al. Validation of a cholangiographic prognostic model in primary sclerosing cholangitis. Endoscopy 2010;42:742-7.
299. Ruiz A, Lemoinne S, Carrat F, et al. Radiologic course of primary sclerosing cholangitis: assessment by three-dimensional magnetic resonance cholangiography and predictive features of progression. Hepatology 2014;59:242-50.
300. Garioud A, Seksik P, Chrétien Y, et al. Characteristics and clinical course of primary sclerosing cholangitis in France: a prospective cohort study. European Journal of Gastroenterology & Hepatology 2010;22:842-847.
301. Ponsioen CY. Natural history of primary sclerosing cholangitis and prognostic value of cholangiography in a Dutch population. Gut 2002;51:562-566.
302. 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.
303. Ehlken H, Wroblewski R, Corpechot C, et al. Validation of Transient Elastography and Comparison with Spleen Length Measurement for Staging of Fibrosis and Clinical Prognosis in Primary Sclerosing Cholangitis. PLOS ONE 2016;11:e0164224.
304. de Vries EM, Farkkila M, Milkiewicz P, et al. Enhanced liver fibrosis test predicts transplant-free survival in primary sclerosing cholangitis, a multi-centre study. Liver Int 2017.
305. Vesterhus M, Hov JR, Holm A, et al. Enhanced liver fibrosis score predicts transplant-free survival in primary sclerosing cholangitis. Hepatology 2015;62:188-97.
306. Eaton JE, Dzyubak B, Venkatesh SK, et al. Performance of magnetic resonance elastography in primary sclerosing cholangitis. J Gastroenterol Hepatol 2016;31:1184-90.
307. Eddowes PJ TP, Corrigan M, Ferguson J, Hirschfield GM. Correlation of Non-Invasive Markers of Liver Fibrosis in Patients with Primary Sclerosing Cholangitis: a Validation Cohort. EASL International Liver Congress. Barcelona, 2016.