

Multi-Parametric MRI Techniques for the Non-Invasive Assessment of Patients with Liver Disease

Doctoral Thesis

Michael Pavlides BSc MBBS MRCP

Keble College

September 2015

Supervisors

Prof. Stefan Neubauer MD FRCP FACC FMedSci

Prof. Eleanor Barnes BSc MBBS MRCP PhD

I dedicate this thesis to my children Sophia, Pavlos and Sebastian.

For all the times I couldn't be there when you got up in the morning or in the evenings when you went to bed.

Contents

Acknowledgements 8
Declaration 11
Conflict of Interest 11
Publications arising 12
Published abstracts 14
Abbreviations 16
Index of tables 19
Index of figures 21
Thesis abstract 23
1. Introduction 25
1.1 Liver disease epidemiology 26
1.2 Risk stratification in clinical practice 27
1.3 Need for robust surrogate end points 27
1.4 Non-invasive assessment of liver fibrosis 28
1.4.1 Ultrasound based elastography 28
1.4.2 Serum based algorithms 29
1.5 Magnetic resonance techniques for liver fibrosis assessment 29
1.5.1 MR for liver fat quantification 30
1.5.2 MR for liver iron quantification 30
1.6 Magnetic resonance elastography 30
1.6.1 MRE for liver fibrosis evaluation 31
1.6.2 MRE for NAFLD evaluation 31
1.6.3 MRE for cirrhosis and portal hypertension evaluation 32
1.7 Diffusion weighted imaging 32
1.8 T_1 relaxometry 33
1.8.1 Combination of T_1 relaxation time with other non-invasive biomarkers 37
1.8.2 $T_1\rho$ relaxation time 37
1.8.3 Dynamic contrast enhanced MRI 38
1.8.4 Molecular MR imaging 41
1.8.5 Multi-parametric MR protocols 44
1.9 Iron corrected T_1 45
1.10 Thesis aims 46
2. General methods of magnetic resonance imaging and spectroscopy 47

2.1 General scanning considerations 48
2.2 Outcome measures 48
2.3 Magnetic resonance imaging 49
2.3.1 Relaxation times 49
2.3.2 T_1 and T_2* mapping techniques 52
2.3.4 Image acquisition 53
2.3.5 Image analysis 54
2.3.6 T_1 correction for iron 55
2.3.7 LiverMultiScan^{TM} 56
2.4 Proton magnetic resonance spectroscopy 60
2.4.1 Background 60
2.4.2 Data acquisition 60
2.4.3 Data processing 61
3. Multi-parametric MR in healthy volunteers 63
3.1 Abstract 64
3.2 Introduction 65
3.3 Aims 65
3.4 Methods 66
3.4.1 Recruitment strategy 66
3.4.2 Multi-parametric MR 66
3.4.3 Measurement of the inferior vena cava and aorta cross sectional area 66
3.4.4 Other study procedures 68
3.4.5 Variability in reporting of MR scans 68
3.4.6 Statistical analysis 68
3.5 Results 70
3.5.1 Participants characteristics 70
3.5.2 Variability between methods of data processing 71
3.5.3 Intra- observer variability 72
3.5.4 Age and gender differences in liver cT_1 73
3.5.5 Additional predictors of liver cT_1 75
3.5.6 Methods for correcting for the confounding effect of liver congestion 76
3.5.7 Effect of aortic size 81
3.6 Discussion 84
3.6.1 Interpretation of results for the variation of liver cT_1 in healthy controls 84

3.6.2 Implications for the application of liver cT_1 as a biomarker 86
4. Assessment of liver fibrosis 88
4.1 Abstract 89
4.2 Introduction 90
4.3 Aims 91
4.4 Patients and methods 92
4.4.1 Study design and population 92
4.4.2 Routine clinical reporting of biopsies 92
4.4.3 Digital imaging analysis for collagen proportionate area 92
4.4.4 Variability in CPA reporting 93
4.4.5 Liver biopsy quality 95
4.4.6 Statistical analysis 96
4.5 Results 97
4.5.1 Cohort characteristics 97
4.5.2 Collagen proportionate area 99
4.5.3 Liver cT_1 histological associations 100
4.5.4 Liver cT_1 and liver disease aetiology 102
4.5.5 Multivariate models 102
4.6 Discussion 104
5. Assessment of non-alcoholic fatty liver disease 107
5.1 Abstract 108
5.2 Introduction 109
5.3 Aims 111
5.4 Patients and methods 112
5.4.1 Study design and patient population 112
5.4.2 Multi-parametric MR examination 113
5.4.3 Transient elastography 113
5.4.4 Liver histology 113
5.4.5 Statistical analysis 114
5.5 Results 116
5.5.1 Baseline characteristics 116
5.5.2 Assessment of fibrosis 120
5.5.3 Assessment of disease activity; ballooning and lobulitis 123
5.5.4 Steatosis vs NASH 125

5.5.5 Overall disease severity (mild vs significant) 125
5.6 Discussion 128
6. Assessment of portal hypertension 131
6.1 Abstract 132
6.2 Introduction 133
6.3 Aims 134
6.4 Patients and methods 136
6.4.1 Study design and patient population 136
6.4.2 Multi-parametric MR 136
6.4.3 Hepatic vein pressure gradient and transjugular liver biopsy 140
6.4.4 Statistical analysis 140
6.5 Results 142
6.5.1. Cohort characteristics 142
6.5.2 Assessment of HVPG as a continuous variable 145
6.5.3 Diagnostic accuracy against HVPG severity categories 148
6.6 Discussion 150
7. Prediction of clinical outcomes 153
7.1 Abstract 154
7.2 Introduction 155
7.3 Aims 156
7.4 Patients and methods 157
7.4.1 Study design and patient population 157
7.4.2 All-cause mortality and liver related clinical events 159
7.4.3 Multi-parametric MRI 159
7.4.4 Statistical analysis 161
7.5 Results 162
7.5.1 Baseline characteristics 162
7.5.2 Follow-up data 166
7.5.3 Cox regression analysis 170
7.6 Discussion 171
8 Conclusion and suggested further work 174
8.1 Summary of main findings 175
8.2 Spleen cT_1 179
8.3 Comparison with other non-invasive tests 179

8.4 Multi-parametric MR as a surrogate end point 182
9. References 184

Acknowledgements

First and foremost I would like to thank all the participants who took part in this study and their families. They have given up their time and in some cases travelled long distances to attend their research MRI scans, which sometimes happened in the early mornings or late evenings.

The inspiration for this project first came from my brother Costas Pavlides whose words of encouragement at a time of uncertainty gave me the final push towards pursuing a post graduate academic degree in Oxford. I am indebted to him as he has been a constant source of advice and guidance.

Meeting Dr Rajarshi Banerjee was the next catalyst in developing my interest in this subject. I thank him for his infectious enthusiasm which provided the early impetus in this work and for his constant interest in my progress.

This thesis would not have been possible without the support from all the clinical and administrative staff in the Oxford Centre for Clinical Magnetic Resonance Research and the Translational Gastroenterology Unit. I would like to thank Jane Francis and Steven Knight for all the advice and help with scanning, Lukasz Owsiak for his expertise with computers, Heidi Crook, Emma Heaton, Sophia Bell, Kathryn Chisholm, Sharon Dash and Cate Wheatley for their administrative help and for managing the study finances, Karen Whatley and Ella May in OCMR reception for so efficiently organising study participant transport when required, Kate Groves, Hanan Lamlum and Elena Villarreal for their help with gaining regulatory approvals and study monitoring, Laura Dolman and Pauline Collier for their help with organising meetings, Hayley Harvey, Joanne Sellwood, Judith Delos Santos and Carla Cascone for their practical input into patient recruitment and data gathering and Lizzie Stafford, Denise O'Donnell, Mark Ainsworth, Victoria Black, Jane

Phillips and Toby Cabe for providing a link to Oxford Hepatology and identifying potential study participants.

In OCMR I was lucky to share an open plan office with other research fellows and to work alongside more senior colleagues who have influenced this work through informal interactions. I would like to thank Eylem Levelt, Victoria Stoll, Andrew Lewis, David Holdsworth. Margaret Loudon, Malenka Bissel, Rina Ariga, Jenny Rayner, Alexander Liu, Rohan Wijesurendra, Ntobeko Ntusi, Henry Boardman, Vanessa Ferreira, Oliver Rider, Maslisa Mahmod, and Saul Myerson for all the discussion academic or not, for the exchange of ideas and for all the help with scanning.

The OCMR physics team including Matthew Robson, Stefan Piechnik, Chris Rodgers, Aaron Hess, Will Clarke, Lucian Purvis, Ferenc Mozes, Damian Tyler and Elizabeth Tunnicliffe were always willing to help with any technical questions. Particular thanks to Elizabeth Tunnicliffe for all the hours spent teaching me how to report MRI scans and to Stefan Piechnik for doing scans with me in the latter part of this project as I was developing an interest in more advanced imaging techniques.

During this project I have worked alongside many clinicians and scientists whose input has facilitated many aspects of this work. I would like to thank: Jane Collier and Jeremy Cobbold in the clinical department of hepatology and their clinical fellows and trainees including Kate Williamson and Neil Rajoria, for embracing the project and identifying potential study participants from their clinics, Bruno Sgromo and Lisa Rickers who have helped with the recruitment of patients undergoing bariatric surgery, Jane Phillips-Hughes who has expertly led the interventional radiology aspects of the study, Lai Mun Wang, Kenneth Fleming and Eve Fryer who engaged in the development of quantitative histology techniques, Jeremy Tomlinson for his support and advice and Leanne Hodson for mentorship and collaboration. I would also like to thank John Ryan, Ricki Sharma,

Michael Silva, Helen Winter, Tessa Greenhalgh, Nicole Zitzmann and Bevin Gangadharan for their interest and engagement with this study.

My supervisors Stefan Neubauer and Ellie Barnes have been a source of inspiration throughout. I found their door was always open and they made time to meet or discuss progress even when their timetable didn't really have any free time. They provided guidance beyond the scientific aspects of this thesis and I consider them as friend and life mentors.

I would like to express my gratitude to my parents Pavlos and Eleni Pavlides for the sacrifice and hardship they endured to give me the best possible education. Without their efforts I would have never been in a position to pursue this thesis.

Lastly this thesis would not have been possible without the support of my wife Mary Pavlides. My early morning starts and late evening finishes have been a lot more difficult for her rather than me. She was equally happy to share my excitement when things were going well and my frustrations during my setbacks. As always she has been the rock that supported not only me but the whole family and for this I am always grateful.

The study was funded by the Oxford NIHR BRC, the Oxfordshire Health Services Research Committee (OHSRC) and the Division of Medical Sciences of the University of Oxford.

Declaration

I declare that the work presented in this thesis is my own and has not been submitted for any other degree in this or any other University or Institute of Learning.

The studies included in this thesis were granted ethical approval (NICOLA Ethic Ref 13/SC/0243; RIAL Ethics Ref: 11/H0504/2) and Trust management approval.

The studies described here were designed by me with supervision from Professor Stefan Neubauer and Professor Eleanor Barnes. Patients in the histological validation study against collagen proportionate area (chapter 4) were recruited and scanned by Dr Rajarshi Banerjee, but I developed the digital imaging analysis technique that was used as the reference standard. Seventeen patients in the NAFLD study (chapter 5) were also from the cohort of patients recruited and scanned by Dr Banerjee. The first 74 patients in the study on prediction of clinical outcomes (chapter 7) were recruited and scanned by Dr Banerjee, but I collected the outcomes data and performed the statistical analysis. I recruited and scanned all the other volunteers and patients that took part in the studies of this thesis.

Conflict of Interest

I am a shareholder in Perspectum Diagnostics, a University of Oxford spin off company.

I have filed patent applications in the field of liver disease assessment using magnetic resonance techniques.

Publications arising

Papers published

Banerjee R*, Pavlides M*, Tunnicliffe EM, Piechnik SK, Sarania N, Philips R, Collier JD, Booth JC, Schneider JE, Wang LM, Delaney DW, Fleming KA, Robson MD, Barnes E, Neubauer S. Multi-parametric magnetic resonance for the non-invasive diagnosis of liver disease. J Hepatol. 2014 Jan;60(1):69-77. *joint first authors

Pramfalk C, Pavlides M, Banerjee R, McNeil CA, Neubauer S, Karpe F, Hodson l. Sex-specific differences in hepatic fat oxidation and synthesis may explain the higher propensity for NAFLD in men. JCEM (in press)

Papers accepted

Pavlides M*, Banerjee R*, Sellwood J, Kelly KJ, Robson MD, Booth JC, Collier J, Neubauer S, Barnes E. Multi-parametric magnetic resonance imaging predicts clinical outcomes in patients with chronic liver disease. *joint first authors

Papers submitted

Pavlides M, Banerjee R, Tunnicliffe EM, Collier J, Wang LM, Fleming KA, Cobbold JF, Robson MD, Neubauer S, Barnes E. Multi-parametric magnetic resonance imaging for the assessment of non-alcoholic steatohepatitis disease severity.

Pavlides M, Birks J, Fryer E, Delaney D, Sarania N, Banerjee R, Neubauer S, Barnes E, Fleming K, Wang LM. Inter-observer variability in histological evaluation of liver fibrosis using categorical and quantitative scores.

Mozes FE, Tunnicliffe EM, Pavlides M, Robson MD. The influence of fat on liver T1 measurements using MOLLI methods at 3 Tesla.

Levelt E, Pavlides M, Banerjee R, Mahmod M, Kelly C, Sellwood J, Thomas S, Francis J Schneider J, Rodgers C, Clarke W, Sabharwal N, Antoniades C, Clarke K, Karamitsos T, Rider O, Neubauer S. Cardiac and hepatic phenotype of diabetes in the absence and the presence of obesity - mechanistic role of ectopic/visceral adiposity

Published abstracts

Pavlides M, Banerjee R, Sellwood J, Kelly C, Robson MD, Booth J, Collier J, Neubauer S, and Barnes E. Multi-parametric MRI predicts clinical outcomes in unselected patients with liver disease-a non-invasive test that meets the AASLD criteria for clinical trials? Hepatology (2014), 60(6);1275A-1276A.

Pavlides M, Banerjee R, Tunnicliffe EM, Wang LM, Ryan JD, Cobbold JF, Neubauer S, and Barnes E. T2* magnetic resonance imaging is more sensitive than histology and serum ferritin for the assessment of iron overload in patients with NAFLD. Hepatology (2014), 60;729A-729A.

Pavlides M, Banerjee R, Tunnicliffe EM, Collier J, Wang LM, Fleming KA, Neubauer S, and Barnes E. Multi-parametric MRI can diagnose steatohepatitis and cirrhosis in patients with NAFLD. Hepatology (2014), 60; 728A-728A.

Pavlides M., Banerjee R, Tunnicliffe EM, Collier J, Wang LM, Fleming KA, Neubauer S, and Barnes E. Novel MRI techniques for the assessment of non-alcoholic fatty liver disease. Journal of Hepatology (2014), 60(1); S344.

Pavlides M., Wang LM, Fryer E, Delaney DW, Sarania N, Banerjee R, Neubauer S, Barnes E, and Fleming KA. Estimation of liver collagen proportionate area by visual inspection. Laboratory Investigation (2014), 94; 426A-426A.

Banerjee R., Rial B, Pavlides M, Osuchukwu C, Lewandowski AJ, Collier J, Schneider J, Neubauer S, and Barnes E. Multi-parametric magnetic resonance shows that obesity is strongly associated with hepatic steatosis in adults and children, and can differentiate patients with NASH from NAFLD. Journal of Hepatology (2014), 60(1);S344-S345.

Pramfalk C, Pavlides M, Banerjee R, Karpe F, Hodson L. Higher plasma triglyceride concentrations in men may be explained by higher hepatic de novo lipogenesis and lower dietary fatty acid oxidation compared with women. Diabetes (2015) 64:A41

Levelt E, Pavlides M, Kelly C, Mahmod M, Francis J, Schneider J, Clarke K, Neubauer S. Non-alcoholic fatty liver disease is a common finding in asymptomatic, uncomplicated type 2 diabetes mellitus patients assessed by multi-parametric magnetic resonance liver imaging. Journal of Hepatology, (2015), 62;S737

Ryan, J, Sgromo B, Magee C, Pavlides M, Ainsworth M, Collier J, Barnes E, and Cobbold J. Prevalence of NASH cirrhosis in morbidly obese patients undergoing bariatric surgery. Appetite (2015), 87;389.

Abbreviations

^{1}H-MRS: proton magnetic resonance spectroscopy

95% CI: 95% confidence interval

ADC: apparent diffusion coefficient

ALP: alkaline phosphatase

ALT: alanine aminotransferase

Ao: Aorta

APRI: AST to platelet ratio index

ASH: alcoholic steatohepatitis

AST: aspartate aminotransferase

AUROC: area under the receiver operating characteristic curve

BMI: body mass index

CAH: chronic active hepatitis

CK18: cytokeratin 18

CP: Child Pugh

CPA: collagen proportionate area

CSPH: clinically significant portal hypertension

cT_1: iron corrected T_1

DCE: dynamic contrast enhanced

DEN: diethyl nitrosamine

DIA: digital imaging analysis

DWI: diffusion weighted imaging

ELF: enhanced liver fibrosis

FHVP: free hepatic vein pressure

FLIP: fatty liver inhibition of progression

GGT: gamma Glutamyl transferase

HBV: hepatitis B virus

HCC: hepatocellular carcinoma

HCT: haematocrit

HCV: hepatitis C virus

HR: hazard ratio

HSC: hepatic stellate cells

HV: healthy volunteers

HVPG: hepatic vein pressure gradient

ICCC: intra-class correlation coefficient

ICG: indo-cyanine green

IQR: inter-quartile range

IVC: inferior vena cava

LIF: liver inflammation and fibrosis

LMS: LiverMultiScan^{TM}

LS: liver stiffness

MELD: model of end stage liver disease

MR: magnetic resonance

MRE: magnetic resonance elastography

MRI: magnetic resonance imaging

MRS: magnetic resonance spectroscopy

NAFLD: non-alcoholic fatty liver disease

NASH CRN: NASH clinical research network

NASH: non-alcoholic steatohepatitis

OCP: oral contraceptive pill

PH: portal hypertension

RF: radio frequency

RGB: red, green, blue

ROC: receiver operating curve

ROI: region of interest

SAF: steatosis activity fibrosis

SD: standard deviation

shMOLLI: shortened modified look locker inversion

SPH: severe portal hypertension

SSFP: single shot steady state free precession

T2DM: type 2 diabetes mellitus

TE: transient elastography

vacT_1: blood volume adjusted and iron corrected T_1

WHVP: wedged hepatic vein pressure

Index of tables

Table 1. T_1 relaxometry for the assessment of liver disease in human subjects 35

Table 2. T_1 relaxometry for the assessment of liver disease in animal models of fibrosis. 36

Table 3. Healthy volunteer characteristics 71

Table 4. Factors with significant association to liver cT_1 on univariate analysis. 75

Table 5. Independent predictors of liver cT_1 on stepwise multiple linear regression 75

Table 6. Multiple linear regression including only MR-derived variables. 77

Table 7. Characteristics of patients who had collagen proportionate area quantification .. 98

Table 8. Associations between liver cT_1 and histological variables 100

Table 9. Multiple linear regression models for prediction of collagen proportionate area. 103

Table 10. Patient demographic laboratory, and magnetic resonance data 117

Table 11. Patient histology profile 118

Table 12. Diagnostic accuracy of optimal cut-offs identified in this study for multi-parametric magnetic resonance imaging and transient elastography for the assessment of fibrosis in patients with NAFLD 122

Table 13. Diagnostic parameters of liver inflammation and fibrosis score for NAFLD assessment 127

Table 14. Clinical. magnetic resonance and laboratory characteristics 145

Table 15. Multiple linear regression models of liver cT_1 and liver fat for prediction of HVPG 147

Table 16. Multiple linear regression models of spleen cT_1 and liver fat for prediction of HVPG 148

Table 17. Accuracy of liver and spleen cT_1 for the diagnosis of HVPG severity categories 148

Table 18: Baseline demographics and MR data 163

Table 19: Baseline histological data 164

Table 20. Baseline laboratory data 165

Table 21. Clinical features of patients who had liver related events in the follow up period. 167

Table 22. Cox regression analysis for the prediction of liver related events for LIF, fat and T2* 170

Table 23. Strategies to improve the performance of liver T_1 as a biomarker of fibrosis and inflammation 178

Table 24. Comparison of multi-parametric liver MR with other non-invasive biomarkers of liver fibrosis 181

Index of figures

Figure 1. Nuclear magnetic resonance. 49

Figure 2. Application of an external magnetic field to protons. 50

Figure 3. A T_1 map acquired from a healthy volunteer. 53

Figure 4. Manual reporting of T_1 and T_2* maps. 56

Figure 5. The relationship between cT_1 and LIF 58

Figure 6. LiverMultiScan reporting interface. 59

Figure 7. Proton magnetic resonance spectroscopy voxel positioning. 60

Figure 8. Liver water suppressed MRS data processing in jMRUI. 62

Figure 9. Measurement of the inferior vena cava and aorta cross sectional diameter. 67

Figure 10. Variability between methods of data processing 72

Figure 11. Intra-observer variability in liver cT_1 reporting 73

Figure 12. Variation of liver cT_1 according to age and gender. 74

Figure 13. Predicted vs measured liver cT_1 in healthy subjects. 76

Figure 14. Frequency distribution of IVC/Ao ratio in the cohort of healthy volunteers 78

Figure 15. Developing a correction algorithm for the removal of the confounding effect of liver congestion. 79

Figure 16. Example of the application of the algorithm for fluid correction. 80

Figure 17. Application of the fluid correction algorithm to healthy volunteers. 81

Figure 18. Frequency distribution of aorta (Ao) cross sectional area in healthy volunteers. 83

Figure 19. Aortic cross sectional area plotted against age for (a) women and (b) men. 83

Figure 20. Digital Imaging Analysis for Collagen Proportionate Area. 94

Figure 21. Frequency distribution of biopsy length 95

Figure 22. Association between collagen proportionate area (CPA) and the Ishak stage 99

Figure 23. Observer dependent variability in reporting of collagen proportionate area. 101

Figure 24. Liver cT_1 vs collagen proportionate area in patients with steatohepatitis and viral hepatitis. 102

Figure 25. Study flow diagram. 112

Figure 26. Representative MR Data..... 119

Figure 27. Liver Inflammation and fibrosis score plotted against the histological fibrosis stage..... 120

Figure 28: Liver stiffness measured using transient elastography, plotted the histological fibrosis stage. 121

Figure 29 Liver inflammation and fibrosis score for the evaluation of (a) ballooning, (b) lobulitis (c) overall activity and (d) the total steatosis, activity and fibrosis score. 124

Figure 30. Liver inflammation and fibrosis score for NAFLD classification according to the Fatty Liver Inhibition of Progression consortium algorithms 126

Figure 31. Diagrammatic illustration of the relationship between the liver and the portal circulation 133

Figure 32. Correction of liver cT_1 to account for the effects of intra-vascular fluid status. .. 138

Figure 33. Measurement of spleen cT_1 using LiverMultiScan..... 139

Figure 34. Study flow diagram 142

Figure 35. Multi-parametric MR for the assessment of non-cirrhotic portal hypertension... 143

Figure 36. Relationship between the hepatic vein pressure gradient (HVPG) and liver and spleen cT_1..... 147

Figure 37. Receiver operating characteristic curves for the accuracy of liver and spleen cT_1 in the diagnosis of (a) any degree of portal hypertension, (b) clinically significant portal hypertension and (c) severe portal hypertension..... 149

Figure 38. Study flow chart. 158

Figure 39. Examples of LiverMultiScan MR data..... 160

Figure 40. Kaplan Meier curves for liver related event free survival with patients stratified according to Liver Inflammation and Fibrosis scores..... 168

Figure 41. Kaplan Meier curves for liver related event free survival with patients stratified according to severity of liver iron and fat..... 169

Figure 42. Factors that influence liver T_1 measurements..... 178

Thesis abstract

Multi-Parametric MRI Techniques for the Non-Invasive Assessment of Patients with Liver Disease

Michael Pavlides

Keble College

Thesis submitted for DPhil in Trinity term 2015

Abstract

Liver disease is a major worldwide health problem with high prevalence and poor patient outcomes. In the last decade, efforts to improve liver disease severity assessment have led to the development of many non-invasive methods, including magnetic resonance (MR) techniques.

The overarching aim of this thesis was to examine a multi-parametric MR technique in the evaluation of different aspects of liver disease. The MR protocol included T_1 and T_2* mapping techniques which were used to compute the iron corrected T_1 (cT_1) and the liver inflammation and fibrosis (LIF) score as measures of liver inflammation and fibrosis.

In healthy volunteers, it was found that premenopausal women had a higher median liver cT_1 (784ms) compared to postmenopausal women (741ms; p=0.038) or age matched men (742ms; p=0.005). Furthermore, blood volume was identified as the most significant physiological determinant of liver cT_1 and a correction algorithm was developed to compute the "volume adjusted and iron corrected T_1" ($vacT_1$), which removes the confounding effect of liver congestion and improves the performance of the test as a biomarker of liver fibrosis.

Against histology, liver cT_1 was strongly associated with liver collagen proportionate area, particularly in patients with viral hepatitis (r=0.80; p<0.0001). Furthermore, liver cT_1 / LIF had an excellent diagnostic accuracy for the assessment of multiple aspects of non-alcoholic fatty liver disease (NAFLD) with an area under the receiver operating curve of 0.89 for the identification of significant NAFLD.

In the evaluation of portal hypertension, liver cT_1 (r=0.49; p=0.038) was associated with the hepatic vein pressure gradient (HVPG), and spleen cT_1 was identified as a new biomarker of portal pressure (r=0.66; p<0.0001 vs HVPG). Furthermore, liver cT_1 / LIF was shown to be useful in predicting clinical outcomes, with a negative predictive value of 100% for a LIF<2.

Overall, this thesis demonstrated multi-parametric MR to be a robust technique that can assess the whole spectrum of liver disease.

1. Introduction

1.1 Liver disease epidemiology

Chronic liver disease is a major worldwide health problem. In the developed world, non-alcoholic fatty liver disease (NAFLD) associated with obesity and diabetes affects around a third of the population[1-3]. Viral hepatitis is estimated to affects more than 400 million people worldwide (chronic hepatitis C: 185 million[4], chronic hepatitis B: 250 million[5]). In Europe, where harmful alcohol consumption is the main cause of liver disease, there are an estimated 29 million people suffering with a chronic liver condition[6].

These worrying epidemiological facts are reflected in poor outcomes in patients with chronic liver disease. In the developing world and China, viral hepatitis is driving an epidemic of primary liver cancer, which is the second most common cause of worldwide cancer mortality, after lung cancer[7]. In the UK, over the last 40 years, mortality from liver disease has been increasing at an alarming rate, in contrast to mortality from other chronic conditions like cardiovascular disease and cancer, which are steadily declining[8]. Furthermore, liver disease is a major cause of mortality in younger patients, therefore has a greater socio-economic impact compared to other chronic conditions that affect primarily older individuals[9].

There is therefore an increasing realisation that urgent action is needed in order to address these worrying statistics. In the UK the chief medical officer has identified liver disease as a major health care priority[10] and a commission has set out recommendations to improve the care of patients with liver disease[8].

Correct diagnosis and accurate severity staging are two integral parts into any effort to address liver disease, as it allows clinicians to identify those at risk of complications and prioritise them for treatment or intensive monitoring. The hallmark of liver disease severity

is the amount of liver fibrosis (scaring) that is deposited in the liver[11]. Liver fibrosis represents a final common pathway in all chronic liver diseases and accumulates as a result of chronic liver injury. Liver fibrosis assessment is therefore crucial as this provides an indication of the overall disease severity.

1.2 Risk stratification in clinical practice

The management of chronic hepatitis C virus infection (HCV) over that last 10 -20 years has relied heavily on risk stratification based on the degree of fibrosis present. Treatment of hepatitis C, until recently, relied on pegylated interferon and ribavirin. This treatment regime was only moderately effective and had considerable side effects[12]. Therefore, in some centres interferon and ribavirin was only given to those with significant fibrosis. In routine clinical practice, the severity of fibrosis was usually assessed using liver biopsy which has considerable limitations including risk of complications, sampling variability and observer dependent variability[13]. The need to identify patients who would benefit from treatment, the large number of patients with hepatitis C and the disadvantages of liver biopsy necessitated the development of new non-invasive methods for fibrosis assessment.

1.3 Need for robust surrogate end points

The landscape of hepatitis C management is now changing, with the introduction of directly acting anti-viral medications, which offer high cure rates with very little side effects[12,14-16]. It is already being suggested that treatment of all patients with hepatitis C, irrespective of fibrosis severity, may be the most cost effective strategy[17]. This approach would therefore make the application of non-invasive techniques of fibrosis assessment for hepatitis C risk stratification obsolete. This major development in the field is shifting attention to NAFLD, which is emerging as the next challenge in hepatology and this is

now highlighting another area of urgent clinical need: the lack of robust surrogate end points in clinical trials.

There are very limited pharmacological treatment options in NAFLD[18,19] and efforts to address this are hampered by the reliance on liver biopsy as the only approved surrogate end point in clinical trials. Higher number of patients are needed in order to achieve the necessary statistical power due to the biopsy associated variability in the reporting of fibrosis stage, and patients may be less likely to participate or more likely to drop out due to the risk of complications. Validation of non-invasive biomarkers of fibrosis as surrogate end-points could therefore revolutionise this field and facilitate drug development.

1.4 Non-invasive assessment of liver fibrosis

In the last 10 – 15 years numerous non – invasive methods to assess liver fibrosis have been developed, some of which are now recommended for use in routine clinical practice[20,21]. Non-invasive tests can be broadly divided into those based on (i) elastography, (ii) serum parameters or (iii) imaging and each have their own advantages and disadvantages.

1.4.1 Ultrasound based elastography

Ultrasound based elastography techniques evaluate the propagation of mechanical waves travelling through the liver. Transient elastography[22] is the most extensively used method in this category. It can be performed at the point of care giving an instant result, but it suffers with high failure rates particularly in patients who are overweight or obese and in patients with ascites. Second generation ultrasound elastography techniques (acoustic radiation force impulse[23], shear wave elastography[24], shear wave ultrasound dispersion vibrometry[25]) have now been developed which address some of the drawbacks of

transient elastography but their main disadvantage is that they have to be performed by radiologists, usually as part of a routine liver ultrasound examination. These second generation ultrasound based elastography techniques, cannot be performed at the bedside as they require experts to perform them and are more timely compared to transient elastography.

1.4.2 Serum based algorithms

Serum based fibrosis biomarkers have shown good diagnostic accuracy for the diagnosis of advanced fibrosis in cross sectional studies comparing them to histology[26] and when used to predict clinical outcomes[27]. However, studies examining the use of serum based algorithms to assess progression or regression of fibrosis have produced disappointing results[28,29]. In the study by McHutchison *et al*[28] histologically assessed fibrosis increased by 58% without any associated change in serum based biomarkers, while in the study by Fontana *et al*[29] serum based biomarkers increased equally in treatment responders and non-responders.

1.5 Magnetic resonance techniques for liver fibrosis assessment

Magnetic resonance (MR) techniques are attractive in the assessment of liver fibrosis. Unlike ultrasound techniques, MR allows deep penetration into tissues, and unlike computed tomography it avoids the use of ionising radiation. Repeated measures can therefore be obtained without safety concerns. Furthermore, MR can assess the whole liver eliminating sampling errors and can provide additional information on anatomy. There is now extensive experience in the use of MR for the evaluation of diffuse liver disease due to liver fat or iron deposition. Addition of fibrosis assessment to these established techniques would offer the possibility of simultaneous assessment of multiple variables.

1.5.1 MR for liver fat quantification

Magnetic resonance imaging (MRI) and magnetic resonance spectroscopy (MRS) are now considered as the gold standard for liver fat quantification. Both techniques were found to be more accurate than histological assessment by a pathologist or by morphometry [30,31]. MR has been used to study the epidemiology of NAFLD[1-3] and more recently it has been used as a surrogate end-point in two NAFLD clinical trials[32,33].

1.5.2 MR for liver iron quantification

MR techniques are also regarded as the gold standard for quantification of liver iron as they have shown an excellent inverse association with biochemically quantified hepatic iron concentration[34,35]. The quantification of liver iron using MR techniques is now recommended by clinical guidelines for the management of hereditary haemochromatosis[36] and also for the management of transfusion related iron overload[37,38].

For the assessment of liver fibrosis, several promising MR techniques have been described, including the T_1 mapping technique that forms the basis of this DPhil. These are reviewed here in order to put this work into the wider context of liver fibrosis assessment using MR methodologies.

1.6 Magnetic resonance elastography

Similar to ultrasound based elastography techniques, magnetic resonance elastography (MRE) can determine liver stiffness by analysis of mechanical waves propagating through the liver. The technique requires the positioning of an external passive driver on the patient's body adjacent to the liver. The passive driver is connected to a pneumatic active driver placed outside the scanner room which generates the mechanical waves that are

propagated through the liver. MRE has now been studied in many aspects of liver disease with some encouraging results.

1.6.1 MRE for liver fibrosis evaluation

In a large prospective single centre study of 96 patients with mixed liver disease aetiologies (63% hepatitis C; 8% NAFLD), MRE had excellent diagnostic accuracy for the differentiation of all stages of fibrosis, with an area under the receiver operating characteristic curve (AUROC) for ≥F1, ≥F2, ≥F3 and F4 of 0.96, 0.99, 0.99. and 1.00 respectively. The study also showed that combinations of MRE with transient elastography (TE) or AST to platelet ratio index (APRI) also produced excellent results[39]. Similar results have been obtained in patients with chronic hepatitis B and C[40,41], and MRE was also shown to be closely related to morphometric quantification of liver fibrosis ($r=0.78$; $p<0.001$)[42].

In a meta-analysis of 697 individual patient data from 12 studies, MRE was found to have a good diagnostic performance for all stage of liver disease (AUROC for ≥F1, ≥F2, ≥F3 and F4 of 0.84, 0.88, 0.93 and 0.92, respectively)[43]. A separate meta-analysis of 13 studies containing 989 patients (but not individual patient data) found pooled AUROC of 0.95, 0.97, 0.96 and 0.98 for the diagnosis of fibrosis stage ≥F1, ≥F2, ≥F3 and F4 respectively[44].

1.6.2 MRE for NAFLD evaluation

MRE was found to be of value in the assessment of fibrosis in patients with NAFLD, with AUROC 0.84, 0.86, 0.92 and 0.89 for the diagnosis of fibrosis stage ≥F1, ≥F2, ≥F3 and F4[45]. Similar results were found in a retrospective study of 325 patients, where the AUROC for the diagnosis of bridging fibrosis (F≥3) was 0.95[46].

Animal and retrospective human studies[47,48] showed some promise for the use of MRE in the diagnosis of non-alcoholic steatohepatitis (NASH), however in the prospective study by Loomba *et al*, MRE only had a modest diagnostic accuracy for NASH (AUROC of 0.73)[45].

1.6.3 MRE for cirrhosis and portal hypertension evaluation

In a study that examined the value of MRE in predicting clinical outcomes in patients with cirrhosis, liver stiffness (LS) was independently associated with the presence of decompensation at baseline. Furthermore, in patients who had compensated disease at baseline, liver stiffness was a useful predictor of decompensation (LS≥5.8 kPa had higher cumulative incidence of decompensation compared to those with LS<5.8kPa)[49].

In animal models, liver and spleen stiffness measured using MRE correlated strongly with portal pressure measurements (spleen stiffness: r^2=0.86; p<0.01[50], liver stiffness: r=0.68; p=0.02[51]). In cirrhotic patients, liver and spleen stiffness measured with MRE were found to have a good diagnostic accuracy for the diagnosis of varices (AUROC: 0.86) and high risk varices (AUROC: 0.81)[52], and in one study the diagnostic accuracy improved with the incorporation of data from dynamic contrast enhanced MRI[53]. A feasibility study has now shown that liver and spleen stiffness could be measured simultaneously[54] and this approach may improve the clinical utility of MRE, particularly for patients with cirrhosis.

1.7 Diffusion weighted imaging

Diffusion weighted imaging (DWI) is a magnetic resonance technique that quantifies the diffusion of water molecules in tissues, and is expressed as the apparent diffusion coefficient (ADC). The rationale for using this technique to assess liver fibrosis is that the

deposition of collagen fibres in the liver would inhibit water diffusion therefore leading to a decrease in the ADC.

Some studies using DWI have produced promising results. A retrospective study of 34 patients with viral hepatitis and 25 healthy controls, showed that patients had lower ADC than healthy controls (mean ADC 1.32 vs 1.75; $p<0.001$), and there was a strong relationship between ADC and the histological fibrosis stage in the patients ($r=-0.922$; $p<0.01$)[55]. In a meta-analysis of 10 studies reporting the performance of DWI compared to histology, the pooled area under the receiver operating characteristic curve was 0.86, 0.83 and 0.86 for the diagnosis of any fibrosis (F1-4), significant fibrosis (F2-4) and bridging fibrosis (F3-4), respectively[56].

The technique has not seen widespread application as confounding factors like steatosis and perfusion also affect the ADC[57,58]. Furthermore, when compared with other MR biomarkers of liver fibrosis, DWI was inferior to MRE[56] and T_1 mapping[59], and only equivalent to transient elastography and serum based biomarkers[60].

1.8 T_1 relaxometry

T_1 relaxation time is a physical property of atoms that varies according to their electrochemical environment, and is discussed in more detail in chapter 2. The T_1 relaxation time of hydrogen atoms in water molecules is longer that the T_1 relaxation time of hydrogen atoms in long hydrocarbon chains like fatty acids, and therefore measuring T_1 can provide information about tissue composition. The historical context of using T_1 relaxation for the assessment of liver disease provides some insight into why this technique has not yet seen widespread application in clinical practice. Furthermore, this historical review will put into context the innovations of the T_1 relaxation technique used in this dissertation.

The concept of nuclear magnetic resonance which is the basis of MR imaging was first described in 1946 by Bloch[61] and Purcell[62]. The technique was first used to produce images in the mid to late 1970s[63]. The first whole body and head scans were published by 1980[64]. The application of MR for liver imaging was examined soon after, and studies published in 1981 / 2 showed that patients with diffuse liver diseases like cirrhosis, or chronic active hepatitis had longer liver T_1 relaxation compared to healthy controls[65,66]. Apart from this initial promise, subsequent studies in humans (Table 1) and animals (Table 2) did not show any improved accuracy beyond the diagnosis of cirrhosis. In particular, studies that examined the association of T_1 relaxation time with histologically defined liver disease, or with the hydroxyproline content in the case of animal studies, produced negative results. Furthermore, some of these studies provided information about confounding factors in liver T_1 measurements, which may explain the lack of diagnostic accuracy when used as a biomarker of fibrosis.

Table 1. T_1 relaxometry for the assessment of liver disease in human subjects

Study	Study design	Patients	Main findings
Smith *et al*, 1981[65]	HV vs patients with diffuse liver disease	HV (n=15) Cirrhosis (n=5) CAH* (n=1)	Longer T_1 in cirrhosis and CAH
Doyle *et al*, 1982[66]	HV vs patients with diffuse liver disease	HV (n=12) Cirrhosis (n=10)	Longer T_1 in cirrhosis
The Clinical NMR group, 1987[67]	T_1 vs histology	Patients with suspected parenchymal disease (n=55)	Tendency toward longer T_1 in cirrhosis/hepatitis
Thomsen *et al,* 1990[68]	HV vs disease controls vs cirrhosis T_1 vs histology in a subset	HV (n=7) Disease controls (n=17) Cirrhosis (n=15) Histology subset(n=10)	Longer T_1 in cirrhosis No association with histology
Keevil *et al*, 1994[69]	HV vs patients with diffuse liver disease	HV (n=42) Liver disease (n=44)	Longer T_1 in patients with cirrhosis and CAH
Heye *et al,* 2012[70]	HV vs cirrhosis	HV (n=31) Cirrhosis (n=61)	T_1 longer in cirrhosis
Kim *et al*, 2012[71]	Non cirrhotic patients vs CHB cirrhosis	Non cirrhotic patients (n=92) HBV cirrhosis (n=87)	T_1 shorter in cirrhosis
Cassinotto *et al*, 2015[59]	HV vs cirrhosis	HV (n=40) Cirrhosis (n=89)	T_1 longer in cirrhosis
Hoad *et al,* 2015[72]	T_1 vs histology	Training cohort (n=64) Validation cohort (n=46)	AUROC for cirrhosis 0.92 AUROC for advanced fibrosis 0.81

Abbreviations**. HV**: healthy volunteers; **CAH**: chronic active hepatitis, **HBV**: hepatitis B virus; **AUROC**: area under the receiver operating characteristic curve.

*Chronic active hepatitis is a historical term used to describe a hepatitis of unknown aetiology, now believed to be chronic hepatitis C. The use of the term “chronic active hepatitis” has declined since the description of the hepatitis C virus in 1989, and is no longer used.

Table 2. T_1 relaxometry for the assessment of liver disease in animal models of fibrosis.

Study	Study design	Animals	Main findings
Goldberg *et al*, 1984[73]	T_1 vs histology T_1 vs hydroxyproline Thioacetamide rat model	Controls (n=14) Treated (n=19)	No difference in T_1 between the two groups
Chamuleau *et al*, 1988[74]	T_1 vs hydroxyproline T_1 vs water content CCl_4 rat model	Controls (n=6) Treated (n=16)	T_1 increases due to drug induced oedema not fibrosis.
Aisen *et al,* 1994[75]	T_1 vs hydroxyproline T_1 vs histology CCl_4 rat model	Controls (n=6) Treated (n=17)	No difference in T_1 between the groups

The study by the Clinical NMR group[67], found significantly shorter liver T_1 in patients with high liver iron. This study also highlighted the confounding effect of inflammation on the assessment of fibrosis, as patients without cirrhosis but with liver inflammation were found to have T_1 values as high as patients with cirrhosis. In animal models of fibrosis, Chamuleau *et al*[74] found that T_1 increased after treatment with CCl_4, but the increment reflected changes in liver water from drug induced oedema rather than changes in liver fibrosis. More recently, Hoad *et al*[72] reproduced these results, showing that iron and inflammation have significant opposing effects on T_1 measurements, in patients without advanced fibrosis.

Alternative approaches became necessary in order to improve the performance of liver T_1 as a biomarker of fibrosis, and the most important of these are reviewed below.

1.8.1 Combination of T_1 relaxation time with other non-invasive biomarkers

In the study by Hoad *et al*[72], liver T_1 is proposed as a biomarker of liver inflammation, which together with liver T_2* for liver iron quantification and the enhanced liver fibrosis (ELF) panel for fibrosis[76], are combined in a decision tree to determining who should be referred for liver biopsy. In this approach, T_2* is assessed first. Patients with low T_2* (<22.6ms) suggestive of high liver iron are referred for biopsy, for the evaluation of iron overload. Patients with no MR indication of high iron are then assessed by the ELF score, and those identified as being at risk of bridging fibrosis (ELF≥9.29) are referred for biopsy. The remaining patients are assessed with liver T_1 and those with a high T_1 suggestive of significant inflammation are also referred for biopsy. Patients without features of high iron, at low risk of bridging fibrosis based on ELF and at low risk of significant inflammation are identified as having overall low risk and managed without liver biopsy. This decision tree resulted in a sensitivity of 90%, specificity of 71%, positive predictive value of 72% and negative predictive value of 89% in identifying those with significant fibrosis or inflammation.

1.8.2 $T_1\rho$ relaxation time

$T_1\rho$ relaxation time can be considered equivalent to the T_1 relaxation time measured at very low field strength and can be used to investigate tissue macromolecule composition. Overall, liver fibrosis assessment with this technique has shown some promise as detailed below, but $T_1\rho$ is yet to be tested against histologically assessed fibrosis in a meaningful patient cohort. Similar to T_1 relaxometry, measurement of $T_1\rho$ relaxation time requires no additional hardware and no injection of intravenous contrast.

$T_1\rho$ was found to be longer in the bile duct ligation rat model of liver fibrosis compared to controls and overall, $T_1\rho$ correlated with fibrosis stage[77]. In the CCl_4 mouse model of liver fibrosis, $T_1\rho$ increased by 38% by day 2 post CCl_4 administration, with the peak increase

at 6 weeks. $T_1\rho$ decreased after CCl_4 withdrawal, but did not return to baseline due to residual fibrosis which was confirmed on histology. Histological examination of the CCl_4 treated rat liver on day 2 showed extensive inflammatory infiltrate and hepatocyte swelling but no collagen deposition, which would suggest that this technique may be equally prone to confounding from inflammation as the T_1 relaxation time.

Assessment of $T_1\rho$ is feasible in humans and the technique was found to be reproducible in assessing healthy volunteers. The intra class correlation coefficient (ICCC) for the test re-test reproducibility in 10 subjects was 0.764, and the ICCC for the inter-observer agreement for reporting the same scan was 0.955[78]. Studies have also shown that $T_1\rho$ can differentiate healthy volunteers from patients with cirrhosis[79,80]. In addition, one of these studies showed that $T_1\rho$ could differentiate between patients with varying degrees of cirrhosis severity according to the Child Pugh (CP) classification (A: compensated cirrhosis – C: end stage) ($T_1\rho$ for CP-A, CP-B and CP-C were 45.4ms, 50.0ms and 54.0ms respectively; $p<0.05$ between all pairs)[79]. The latest study in this flied, found an association of $T_1\rho$ with liver function assessed using indocyanine green clearance ($r=0.46$; $p<0.05$), but no association with histologically assessed fibrosis[81]. However, this study evaluated the technique in patients with suspected liver malignancy related to chronic liver disease or malignancy in other organs, and some of the patients included in the analysis had a history of partial hepatectomy or radiofrequency ablation. This cohort is therefore, not representative of the majority of patients with chronic liver disease, and as such the results of this study are difficult to interpret.

1.8.3 Dynamic contrast enhanced MRI

Dynamic contrast enhanced (DCE) MRI requires the injection of hepatocyte specific gadolinium contrast agents like gadoxetic acid (Gd-EOB-DTPA; Primovist©; Bayer, Berlin,

Germany), which leads to a decrease in the T_1 relaxation time. After injection, the contrast agent is taken up by hepatocytes and excreted into bile. Hepatobiliary excretion accounts for about 50% of the total excretion of the contrast agent. T_1 is measured on images acquired before contrast injection ("T_1-pre") and on images acquired 20 minutes after injection (delayed phase or hepatobiliary phase; "T_1-HBP"). The absolute (T_1-pre minus T_1-HPB) and proportional change (absolute T_1 change / T_1-pre) in T_1 after contrast administration are also computed. In patients with poor liver function due to decreased hepatocyte numbers or function, more contrast is retained in the delayed phase (lower T_1-HPB and larger T_1 change from baseline) compared to patients with normal liver function. This technique gives an estimate of liver function, and has also been used to evaluate liver fibrosis, severity of cirrhosis and steatohepatitis as detailed below. Even though it is generally accepted that liver functional capacity deteriorates with advancing fibrosis, the relationship between fibrosis and liver function may not be linear and may also depend on the underlying disease aetiology.

1.8.3 (i) DCE MRI for liver function evaluation

The liver produces proteins like albumin and clotting factors and excretes bilirubin. Assessing serum levels of protein (e.g. albumin) and bilirubin levels, or clotting times (prothrombin time) provide an estimate of liver function. Global liver function can also be estimated by tests like the indocyanine green (ICG) clearance rates. ICG is exclusively excreted by the liver; therefore the rate of clearance of ICG from plasma reflects global liver function. DCE MRI has the potential advantage of providing information about the segmental liver function.

Studies have shown that parameters derived from DCE MRI correlate with liver function tests like bilirubin, albumin / pre-albumin and prothrombin time[82,83]. In a recent study, the

post contrast T_1 reduction rate was closely related to the ICG clearance rate (r=0.86; P<0.001), and the prediction model improved further when liver volume measures were also included in a regression model with post contrast T_1 reduction rate (r=0.92; p<0.001)[84].

1.8.3(ii) DCE MRI for cirrhosis severity evaluation

Cirrhosis severity is routinely assessed using clinical scores like the Child Pugh (CP) score[85], which classifies patients into 3 severity groups (CP-A: compensated disease, CP-B: advanced disease and CP-C end stage disease) or composite scores like the model of end stage liver disease (MELD)[86]. DCE MRI has been shown to differentiate between healthy controls and patients with cirrhosis and between the cirrhosis severity assessed by CP or MELD[82,87-90].

No association was found between DCE MRI and cirrhosis severity that was determined radiologically by evaluation of liver morphology[91], and this would suggest that the burden of fibrosis and the associated liver structural changes may not be directly related to liver function.

Studies have also shown the ability of DCE MRI to assess liver function in patients with primary biliary cholangitis[92] and primary sclerosing cholangitis[93]. DCE MRI may be particularly useful in the assessment of liver function in these conditions which affect the liver in an inhomogeneous distribution, and which were demonstrated by the finding of significant differences between liver segments in one patient in the primary sclerosing cholangitis study[93].

1.8.3(iii) DCE MRI for non-alcoholic steatohepatitis evaluation

DCE MRI has demonstrated good accuracy for the diagnosis of non-alcoholic steatohepatitis (NASH; AUROC 0.87) in a rabbit model of NAFLD[94], and in a retrospective human study of 81 patients (AUROC 0.85)[95].

1.8.3(iv) DCE MRI for liver fibrosis evaluation

DCE MRI was shown to have some diagnostic accuracy for the diagnosis of fibrosis in retrospective studies[96,97], and in the animal model of NAFLD (AUROC for F≥1, F≥2 and F≥3 were respectively 0.91, 0.88 and 0.95)[94]. Furthermore, in the study by Feier *et al*[97], fibrosis was the only independent histological predictor of the relative enhancement on DCE MRI.

1.8.3(v) Comparison of DCE MRI with other modalities

DCE MRI was found to be superior to diffusion weighted imaging[96] and to have similar diagnostic accuracy to liver stiffness measurement using transient elastography for the diagnosis of bridging fibrosis (F0-2 vs F3-4; AUROC 0.83 for both tests)[98]. Furthermore, use of gadolinium based contrast agents were shown to be superior to iron oxide based agents for the assessment of liver fibrosis[99], but equivalent for the assessment of liver function[100].

1.8.4 Molecular MR imaging

Molecular MR imaging represents a unique implementation of MR technology to visualise biological processes at the cellular and molecular level. In these techniques a ligand probe (e.g. monoclonal antibody, peptide) with high binding affinity to a biological molecule of interest (e.g. cell surface receptor, enzyme) is coupled to an injectable MR contrast agent (e.g. gadolinium). Contrast agents will produce a signal change between the baseline image and the image acquired after contrast administration. Therefore, when

the contrast agent is coupled to a molecular probe, preferential binding with the biological molecule of interest occurs, therefore allowing visualisation and quantification of related processes. If the biological molecule of interest is related to liver fibrogenesis or fibrosis, then this technique could help localise the process, as MR signal (before / after contrast injection) will only change in tissues where fibrogenesis and fibrosis are occurring. Furthermore, the magnitude of the signal change would be representative of the severity of fibrosis. Four molecular probes have been studied in animal models of fibrosis.

1.8.4(i) Molecular MR for type I collagen

Type I collagen is a major constituent of liver fibrosis. As it deposited in the extracellular space, it is deal target for molecular MR imaging. A molecular probe (EP-3533) to type I collagen has been developed, and is the most extensively studied probe in molecular MR imaging of liver fibrosis. When injected, this probe leads to shortening of the T_1 relaxation time.

In a feasibility study[101], EP-3533 was evaluated in the CCL_4 mouse model and the diethyl nitrosamine (DEN) rat model where it was shown to identify liver fibrosis. Animals with fibrosis were found to have greater signal intensity compared to controls (0.55 vs 0.39; $p<0.05$) when imaged after EP-3533 injection. Furthermore, EP-3533 washed out more slowly from the livers of fibrotic animals (half-life 51.3 minutes for fibrotic rats vs 42.0 minutes in controls). EP-3533 content was measured in the sacrificed animals and found to have a strong correlation with hydroxyproline levels ($r=0.74$ for rats; $r=0.77$ for mice) and with Ishak staging ($r=0.84$ for rats; $r=0.79$ for mice).

In a subsequent study[102], the group confirmed the findings in the CCl_4 model of fibrosis and also showed that imaging with EP-3533 was superior to other MR biomarkers of

fibrosis (T_1 relaxation time, $T_1\rho$, diffusion weighted imaging and magnetisation transfer techniques).

In their latest study[103], the group used the rat bile duct ligation model of fibrosis to demonstrate that EP-3533 was also useful for assessing biliary type fibrosis. The ability of this technique to monitor therapeutic effects was also investigated. Some animals were treated with rapamycin, which is known to inhibit fibrosis in this model. Control animals, or animals that responded to rapamycin developed less fibrosis and had lower enhancement after EP-3533 compared to animals that did not receive or did not respond to treatment. Furthermore, this study used a whole liver imaging acquisition and analysis technique on a clinical scanner, in contrast to the previous studies by the group where high field strength animal magnets were used.

1.8.4(ii) Molecular MR for $\alpha_v\beta_3$ integrin

Hepatic stellate cell (HSC) activation is central in the cascade that leads to liver fibrosis accumulation[11]. Integrins are transmembrane proteins that interact with the extracellular matrix[104], and intergrin $\alpha_v\beta_3$ is highly expressed in activated HSCs[105]. A molecular imaging probe, consisting of a peptide with high affinity for integrin $\alpha_v\beta_3$ coupled to an iron oxide based contrast agent was developed and tested on the CCl_4 mouse model. Iron oxide based contrast agents shorten the T_2 relaxation time, unlike gadolinium based agents which shorten T_1. This study demonstrated that a greater signal change was detected when the molecular probe was used in CCl_4 treated animals (T_2 shortening of 29.8ms), compared to CCl_4 treated animals imaged with the uncoupled iron oxide contrast alone or in control animals imaged with the uncoupled iron oxide contrast or with the full molecular probe (T_2 shortening ranged from 12.0 – 12.4 ms; $p<0.05$).

1.8.4(iii) Molecular MR for fibrin-fibronectin complexes

Fibrin and fibrinogen are known to play a role in extracellular matrix deposition particularly during short term liver injury[106]. An imaging probe that recognises fibrin-fibronectin complexes has been developed and tested in the CCl_4 mouse model of fibrosis[107]. This probe consists of a cyclical peptide with high affinity to fibrin-fibronectin complexes coupled to a gadolinium based contrast agent. Administration of the molecular imaging probe resulted in higher signal change in CCl_4 treated animals compared to controls. Increased fibronectin in fibrotic livers was also demonstrated on histology.

1.8.4(iv) Molecular MR for elastin

In a small study using 6 animals (3 controls; 3 treated with CCl_4), Ehling *et al*[108] showed some concordance between histology and MR images acquired after injection with a probe of gadolinium coupled to an elastin specific agent. As the collagen-to-elastin ratio changes in the evolution of liver fibrosis[109], with elastin accumulating in the more advanced stages, the authors speculate that combinations of imaging probes to elastin and type I collagen could be used to increase diagnostic accuracy and to assess progression or regression of fibrosis.

1.8.5 Multi-parametric MR protocols

Combinations of several techniques for liver fibrosis assessment have also been examined. In a retrospective study of patients who had undergone liver MR, combination of parameters derived from diffusion weighted imaging, DCE MRI and susceptibility weighed imaging, resulted in improved diagnostic accuracy for the classification of fibrosis (AUROC for F0 vs F1-4: 0.95; F0-1 vs F2-4: 0.95; F0-2 vs F3-4: 0.90; F0-3 vs F4: 0.93)[110]. Incorporation of MR texture analysis techniques with other MR modalities has also produced some promising early results[111,112].

1.9 Iron corrected T_1

Liver iron is one of the major confounders in the use of liver T_1 as a biomarker of liver fibrosis. For example in the study by Hoad *et al*[72], 39% of patients had high iron that excluded them for further analysis with liver T_1. Dr Banerjee and investigators at the Oxford centre for clinical magnetic resonance research (OCMR) and the Oxford translational gastroenterology unit made a significant contribution in this field by developing an algorithm to remove the confounding effect of iron from the T_1 measures yielding the iron corrected T_1 (cT_1), which makes T_1 much more widely applicable[113]. In a prospectively recruited cohort of patients undergoing liver biopsy for the assessment of fibrosis, cT_1 showed excellent diagnostic accuracy against the Ishak staging system of fibrosis with an AUROC of 0.94 for the diagnosis of patients with any degree of fibrosis (F0 vs F≥1)[114].

Liver cT_1 offers some potential advantages compared to the other T_1 relaxometry techniques summarised above. Compared to the approach by Hoad *et al*[72], iron corrected T_1 is more widely applicable without the added costs of an ELF test. Furthermore, measurement of liver cT_1 is quicker than DCE-MRI and does not require injection of intravenous contrast. In addition the study by Dr Banerjee[114] validated liver cT_1 against histology in a prospectively recruited patient cohort, something that is still lacking in many of the other MR techniques.

Liver iron corrected T_1 (cT_1) and a standardised score derived from cT_1 (liver inflammation and fibrosis score; LIF) were the main variables of interest in this thesis and are described in detail in chapter 2.

1.10 Thesis aims

Liver cT_1 had been compared to categorical histological scores of fibrosis before the start of this DPhil. Although comparison with categorical histology scores is widely reported in the literature for many non-invasive techniques, this approach is flawed due to the observer dependent variability in the reporting of categorical histology scores. Therefore, the main aim of this thesis was to validate and extend the work on liver cT_1, by using more robust references comparators. These comparators were (i) morphometrically assessed liver fibrosis (collagen proportionate area), which has been shown to suffer from very little observer dependent variability[115], (ii) hepatic venous portal gradient which gives an overall assessment of the severity of liver disease and (iii) the development of clinical end points. Liver cT_1 was also studied in a cohort of healthy volunteers and in patients with NAFLD, which is a condition where application of this technique may have a significant clinical impact. The specific aims were:

1. To study liver cT_1 in healthy volunteers in order to:
    a. Establish a normal range of liver cT_1
    b. Investigate the physiological determinants of liver cT_1
2. To investigate the association of liver cT_1 with quantitative histologic measures of liver fibrosis.
    a. This necessitated the development of digital imaging analysis methodology for the quantification of fibrosis on liver biopsy sections.
3. To assess the diagnostic accuracy of liver cT_1 in the assessment of non-alcoholic fatty liver disease, as this condition is the next major challenge in hepatology
4. To assess the association between liver cT_1 and portal hypertension and to explore spleen related MR parameters for the detection of portal hypertension.
5. To determine whether liver cT_1 had any value in predicting clinical events.

2. General methods of magnetic resonance imaging and spectroscopy

2.1 General scanning considerations

All scans were performed at the Oxford centre for clinical magnetic resonance research (OCMR) with the patients lying supine in a 3 Tesla MR scanner (Tim Trio, Siemens Healthcare, Germany). Patients attended for their MR scans having fasted for at least 4 hours. The scanning protocol included imaging (T_1 and T_2* mapping) and liver proton MR Spectroscopy (^{1}H MRS).

2.2 Outcome measures

The outcome measures from the multi-parametric technique were:

1. The iron corrected T_1 (cT_1) relaxation time (ms)
    - Derived from T_1 and T_2* maps (details below)
    - Used as a measure of liver fibrosis and inflammation
2. The Liver Inflammation and Fibrosis (LIF) score (0-4)
    - Derived from the histological association of cT_1 with liver fibrosis (details below)
    - Used as a measure of liver fibrosis and inflammation
3. The T_2* relaxation time (ms)
    - Measured from T_2* maps
    - Used as a measure of liver iron
4. Liver fat content (%)
    - Measured from ^{1}H MRS data
    - Used as a measure of liver fat

2.3 Magnetic resonance imaging

2.3.1 Relaxation times

Magnetic resonance imaging exploits fundamental properties of protons in the atomic nucleus of hydrogen. Protons revolve around their axis, in a manner similar to planets. This rotation creates a magnetic field, so protons can be considered as tiny magnets. The direction of the magnetic field at any one time is random (Figure 1).

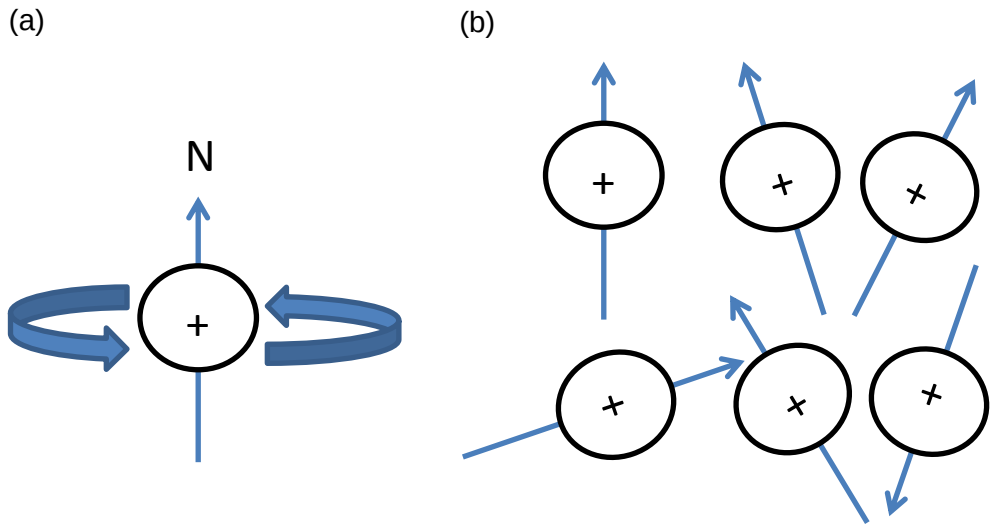


Figure 1. Nuclear magnetic resonance.

(a) Illustration of how a proton revolves around its axis to generate a magnetic field, and (b) an illustration of how the direction of the magnetic fields of protons are random at any one time.

When protons are placed in an external magnetic field such as in an MR scanner (main magnetic field, B_0), they align themselves to that field. Some protons with a lower energy state will align in the direction of the magnetic field and some with a higher energy state will align against the direction of the magnetic field. When in the magnetic field, protons are not static, but rather revolve around their axis similar to a spinning top. This characteristic motion is called precession (Figure 2).

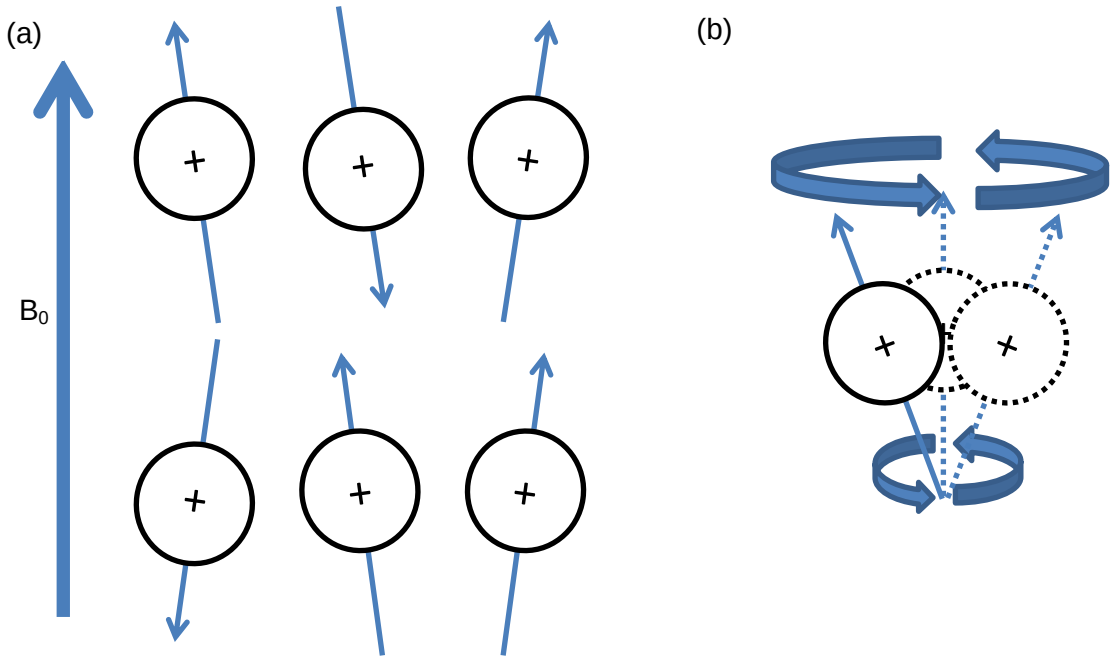


Figure 2. Application of an external magnetic field to protons.

(a) An illustration of how protons will align themselves with the main magnetic field (B_0). The majority of the protons will be in a low energy state and will align in the direction of the main magnetic field, while high energy protons will align in the opposite direction. (b) An illustration of precession. When placed in an external magnetic field, protons will precess, a process where protons rotate in a manner similar to a spinning top. Circles with dotted lines represent protons at different time points, to illustrate the rotational motion of precession.

At any one time, more protons are in a low energy state, and this generates a small net magnetic field (longitudinal magnetisation) in the same direction as the external magnetic field. The signal from this longitudinal magnetisation is not detectable as it lies in the same direction as the much stronger external magnetic field. In order to generate detectable signal, protons have to be manipulated through the application of external radiofrequency (RF) pulses.

The application of a radiofrequency pulse has two main effects. Firstly, it causes some protons to enter into a high energy state and switch their alignment into the opposite direction from the main magnetic field. After removal of the RF pulse, these protons will fall back into their low energy state. In falling back to a lower energy state, protons emit energy, and this can be detected, thus represents a useful signal. The process of falling back into the lower energy state is termed T_1 relaxation, and the time needed for this to occur is the T_1 relaxation time.

Secondly, RF excitation leads to synchronisation of the precession of protons, which results in the generation of a small magnetic field at 90° to the main magnetic field (transverse magnetisation). As this magnetic field now lies in a different direction to the main magnetic field, it can be detected, therefore generating useful signal. As the RF pulse is removed, protons become “de-synchronised”, and the transverse magnetisation is lost. This is termed T_2 relaxation and the time needed for this to occur is the T_2 relaxation time. T_2* relaxation time is derived from T_2 and also takes account magnetic field inhomogeneity.

The relaxation times of protons are influenced by their immediate electrochemical environment. Protons in water molecules which are in constant motion have longer T_1 relaxation times than protons tightly packed in hydrocarbon chains such as fatty acids. In tissues and organs, the T_1 relaxation time depends on the relative distribution of protons in different molecular structures. For example, blood which contains a lot of water but little fat has a long T_1 relaxation time, while adipose tissue with little water and a lot of fat has a comparatively shorter T_1 relaxation time. Similar principles apply to T_2 and T_2* relaxation times.

2.3.2 T_1 and T_2* mapping techniques

MR acquisition and processing techniques can be used to measure the relaxation time of interest (T_1 or T_2*) in any anatomical plane through an organ and "map" them onto a quantitative image, usually through the use of a colour scale. The resulting image or map can then be interrogated to measure the relaxation time in a region of interest. These MR mapping techniques form the foundation of the methodology used in this thesis. An example of a T_1 map is shown in Figure 3.

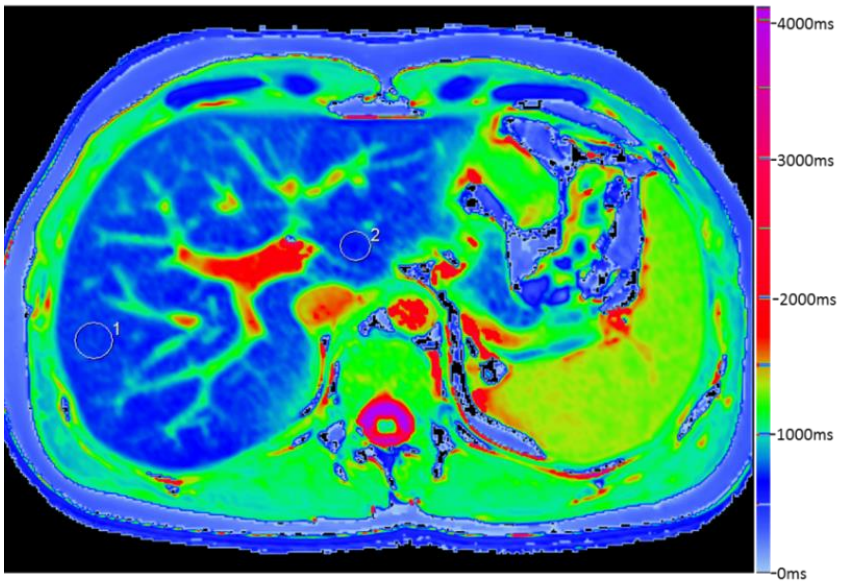


Figure 3. A T_1 map acquired from a healthy volunteer.
The colour scale to the right of the image is used to translate T_1 relaxation times from each pixel location, into colours. Subcutaneous fat has a low T_1 relaxation time and appears blue, while blood has a higher T_1 relaxation time and appears red. Examples of how regions of interest (labelled 1 and 2) can be chosen in order to measure the T_1 relaxation in the tissue of interest (the liver in this case).

2.3.4 Image acquisition

2.3.4(i) T_1 mapping

A T_1 relaxation time map was acquired using the shortened Modified Look Locker Inversion recovery (shMOLLI) sequence[116] in a transverse abdominal plane, through the liver. A subject–dependent frequency adjustment was carried out during end-expiration. The ShMOLLI sequence samples the T_1 recovery curve using single-shot steady state free precession (SSFP) acquisitions using the following parameters:

TR 2.14ms, TE 1.07ms, flip angle of 35°, field-of-view optimised per patient, acquisition matrix 192 x 134 to160, depending on patient, with GRAPPA acceleration of 2 with 24 reference lines, yielding a typical interpolated voxel size 0.9 x 0.9 x 8mm. Images were acquired 200ms after the ECG R-wave and the total time for each SSFP acquisition between 169 and 197ms, depending on the number of phase encoding steps.

2.3.4(ii) T_2* mapping

A multi-gradient-echo acquisition with radiofrequency (RF) spoiling was used to calculate a T_2* transverse abdominal map in the same plane as the T_1 map. The same field-of-view as in the T_1 mapping sequence was used, with a matrix size of 192x128-160, depending on patient, slice thickness of 3mm and 2x GRAPPA acceleration, with the same 200ms delay after the R-wave before acquisition. The image was acquired in nine segments with a TR of 26.5ms and flip angle of 20°. Echo times are selected as far as possible such that the signals from fat and water are in phase (TE = 2.46, 7.38, 12.30, 17.22 and 22.14ms). Fat-saturation and a double-inversion-recovery black blood preparation were used.

2.3.5 Image analysis

2.3.5(i) Region of interest selection

Regions of interest (ROI) were selected in the right and left liver lobes, and the two reading were averaged. There were two considerations in the ROI selection:

(a) Image quality (R^2 maps)

Both T_1 and T_2* mapping have an image quality assurance component. Each map acquisition generates an R^2 map for the fit of signal intensity to the exponential recovery curve. R^2 can be measured in each region of interest, and this work ROIs were acceptable only if R^2 was ≥ 97% in both the T_1 and T_2* maps and this was the case in all measurements.

(b) Location of ROI

The ROI was placed away from vascular and biliary structures, approximately halfway between the porta hepatis and the liver surface in order to avoid interference from the fluid filled structures in the porta hepatis and subcutaneous tissue or air, close to the liver surface.

2.3.6 T_1 correction for iron

T_1 generally reflects changes in extracellular fluid, and increases with worsening inflammation and fibrosis. In cardiac MR for example, T_1 mapping without further adjustments can provide useful information for myocardial fibrosis and inflammation[117,118]. However, in the liver, T_1 measurements are confounded by the presence of iron, which is a frequent finding in patients with liver disease[114,119], and which can lead to a false lowering of T_1. In this work, T_1 measurements were adjusted for iron, which can be accurately measured from T_2* maps[35,114], and expressed as the iron corrected T_1 (cT_1; the T_1 that would be measured if the iron concertation was normal (1.3mg/g) corresponding to T_2* times of 18.0ms to 21.0ms).

The correction algorithm can be implemented “manually”, through the separate reporting of T_1 and T_2* from corresponding locations in the organ of interest and then using look-up tables to establish the cT_1 (Figure 4), or through automated reporting using dedicated software (see below).

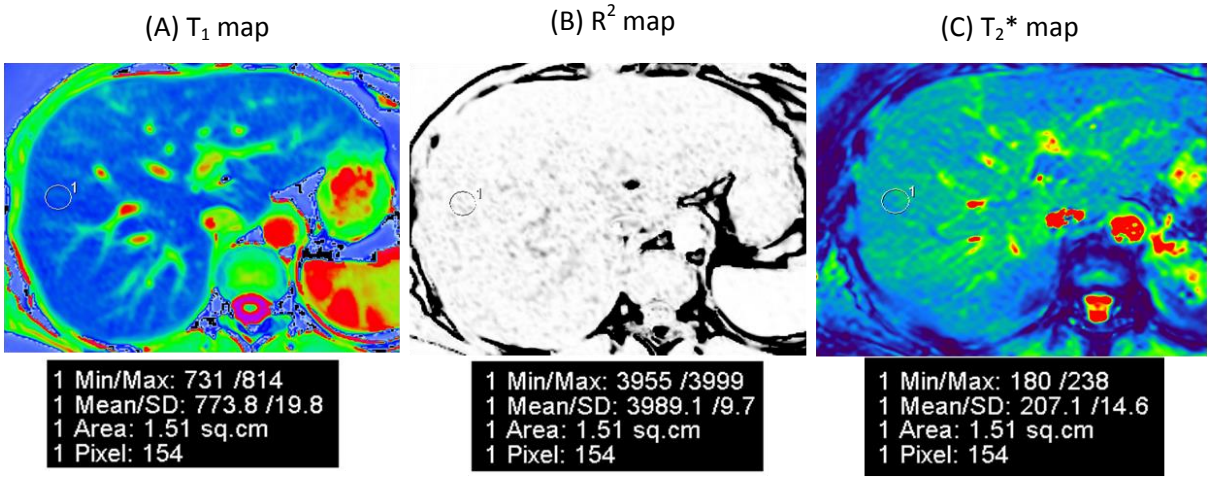


Figure 4. Manual reporting of T_1 and T_2* maps.

The region of interest (ROI) is placed in equivalent positions on the T_1, R^2 and T_2*maps. The range of values, mean and standard deviation, area (cm^2), and the number of pixels are given for each ROI. In this example, mean T_1 was 774ms, mean R^2 was 0.98 and mean T_2* 20.7ms. Look up tables are then used to compute the iron corrected T_1 which in this case was 779ms.

2.3.7 LiverMultiScan™

LiverMultiScan (LMS; Perspectum Diagnostic, Oxford, UK) is a software product specifically designed for the analysis of T_1 and T_2* maps for the measurement of cT_1 (computed from T_1 and T_2*). In an additional step, LMS reports the liver inflammation and fibrosis score (LIF). The LIF score is a continuous score (0-4) that has been derived from the histological associations of cT_1 with liver fibrosis, in a cohort of patients with mixed liver disesase[114].

In setting up the LIF score, the cT_1 values from the histological association study were used as references for 5 anchor points. The fibrosis severity in the histology study was assessed using the Ishak staging system[120] (F0: no fibrosis, F1-2: mild fibrosis, F3-4: moderate fibrosis, F5-6: severe fibrosis). The minimum cT_1 measured in the histology study was assigned a LIF score of zero, while the maximum a score of four. Receiver

operating characteristic curve analysis was used to identify optimal cT_1 cut-off for the differentiation of no fibrosis from any degree of fibrosis, no or mild fibrosis from moderate or severe fibrosis and no, mild or moderate from severe fibrosis, and these were assigned LIF scores of 1, 2 and 3 respectively. The resulting score is a continuous numerical score (0-4) which can also be interpreted as a 4 tier system where: LIF <1 signifies no significant fibrosis, LIF1-1.99 signifies mild fibrosis, 2-2.99 moderate fibrosis and ≥3 signifies severe fibrosis.

The overall relationship between cT_1 and LIF is therefore not linear, but is defined based on the range of cT_1 values between these cut-offs. Figure 5 demonstrates the relationship between cT_1 and LIF, and Figure 6 shows the LMS reporting interface.

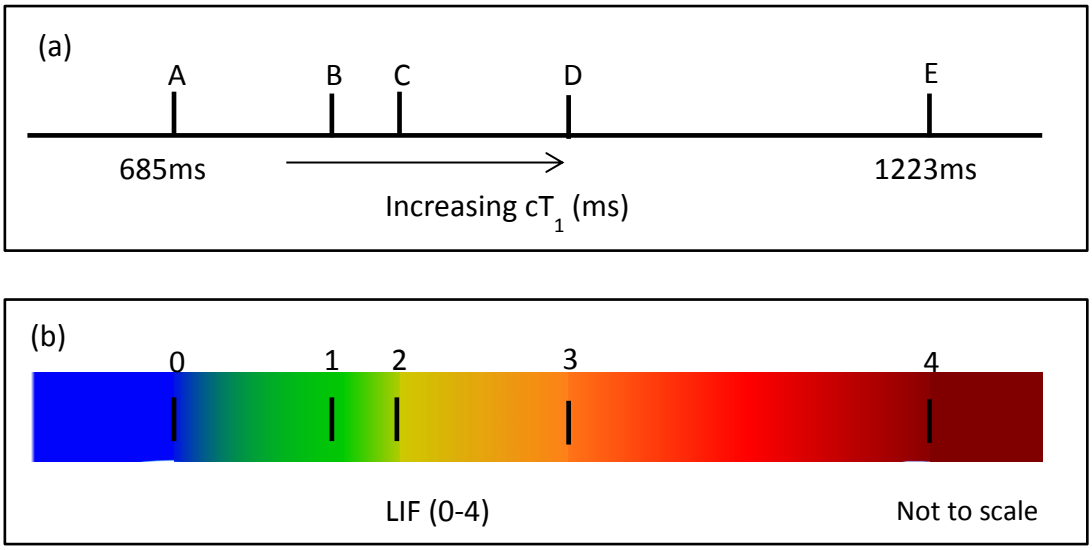


Figure 5. The relationship between cT_1 and LIF

The figure illustrates (a) how iron corrected T_1 (cT_1) cut-offs were identified based on the diagnostic accuracy of cT_1 against histological fibrosis in a cohort of patients with liver disease of mixed aetiologies,[114] where: A and E were respectively the lowest (685ms) and highest (1223ms) measured cT_1, and B, C, and D, respectively represent optimal cT_1 cut-offs for the identification of patients with no vs any degree of fibrosis, no or mild vs moderate or severe fibrosis and no, mild or moderate vs severe fibrosis. (b) An illustration of how the overall relationship between cT_1 and the liver inflammation and fibrosis score (LIF) is not linear.

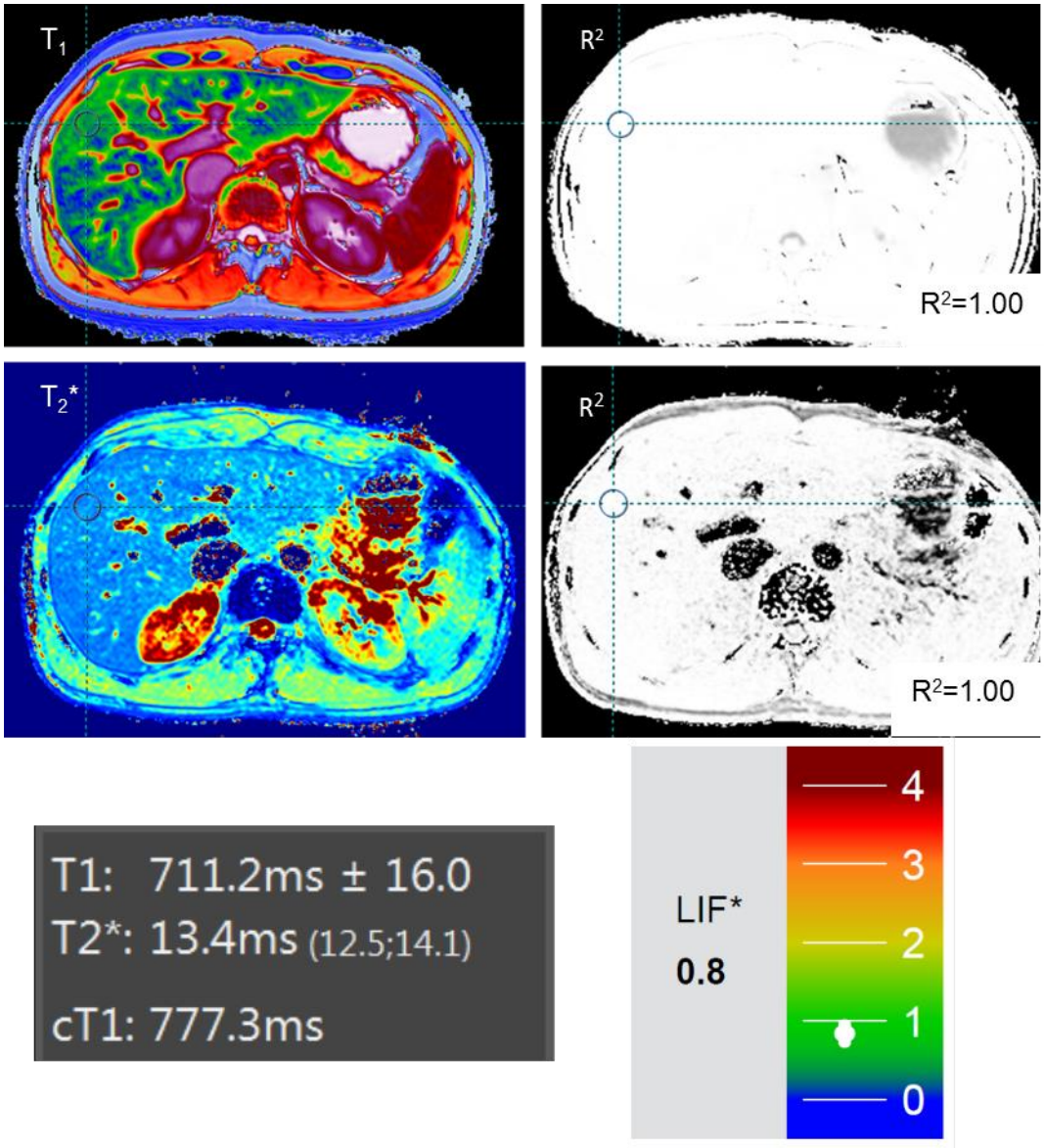


Figure 6. LiverMultiScan reporting interface.

The LiverMultiScan reporting interface allows the simultaneous examination of T_1, T_2* and R^2 maps. The cross hairs intersect at the centre of the region of interest, and can be moved over the images until a satisfactory position is achieved in all 4 maps. LMS automatically reports the cT_1 and LIF scores as illustrated here. A report showing the iron corrected T_1 map can also be produced.

2.4 Proton magnetic resonance spectroscopy

2.4.1 Background

Magnetic resonance spectroscopy utilises the properties of atomic nuclei which when placed in a fixed magnetic field, resonate at varying frequencies depending on their surrounding electrochemical environment. In proton magnetic resonance spectroscopy (^{1}H-MRS), the technique used in this thesis, the nucleus of interest is hydrogen. Hydrogen atoms in water molecules resonate at different frequency from hydrogen atoms in hydrocarbon chains such as in fatty acids. By quantifying the signal associated with each resonant frequency, ^{1}H-MRS enables the relative quantification of different hydrogen containing molecules, the basis of this technique.

2.4.2 Data acquisition

Localiser images were acquired in the transverse, coronal and sagittal plane and were used to plan a voxel (2x2x2cm) in the right liver lobe away from vascular and biliary structures (Figure 7).

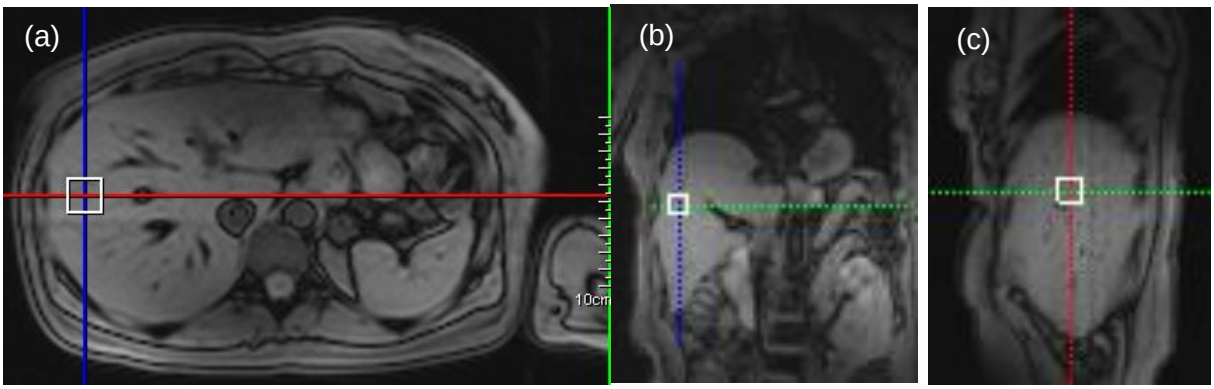


Figure 7. Proton magnetic resonance spectroscopy voxel positioning.

At 3 Tesla water resonates at a frequency of 4.7ppm and the main lipid peak (CH2-CH2; methylene) is at 1.3ppm. For more accurate quantification of the peak at 1.3ppm, data were acquired following calibration pulse sequence to determine the optimum water suppression pulse scaling factor ("water suppression on"). Five acquisitions were

obtained per breath-hold with ECG-triggering to minimise noise from variations in arterial blood flow in the liver. Three data sets with water suppression “on” were acquired per patient and one data set with water suppression “off”. The patients therefore had to hold their breath for 10-12 seconds four times.

Proton spectroscopy acquisition parameters were: TE 10ms; mixing time 7ms; 1024 points acquired at a bandwidth of 2000Hz; scan frequency 1.3ppm for water-suppressed spectra and 4.7ppm for water-unsuppressed spectra; TR 2s for water-suppressed data and 4s for water-unsuppressed data).

2.4.3 Data processing

^{1}H-MRS data were analysed offline using AMARES in the jMRUI software package and MATLAB 2012a. Signals from different coil elements in each breath-hold were combined, and individual spectra phase- and frequency- corrected prior to summation. Prior knowledge and soft constraints were used to define 7 peaks in the spectra. The water peak at 4.7ppm was quantified from water unsuppressed spectra, and the lipid peak at 1.3ppm (methylene) from water suppressed spectra. An example of data processing for a water suppressed spectrum is shown in Figure 8.

Liver fat was expressed a percentage of the total signal according to the following equation:

Liver fat = [liver fat / (liver fat + liver water)] x 100

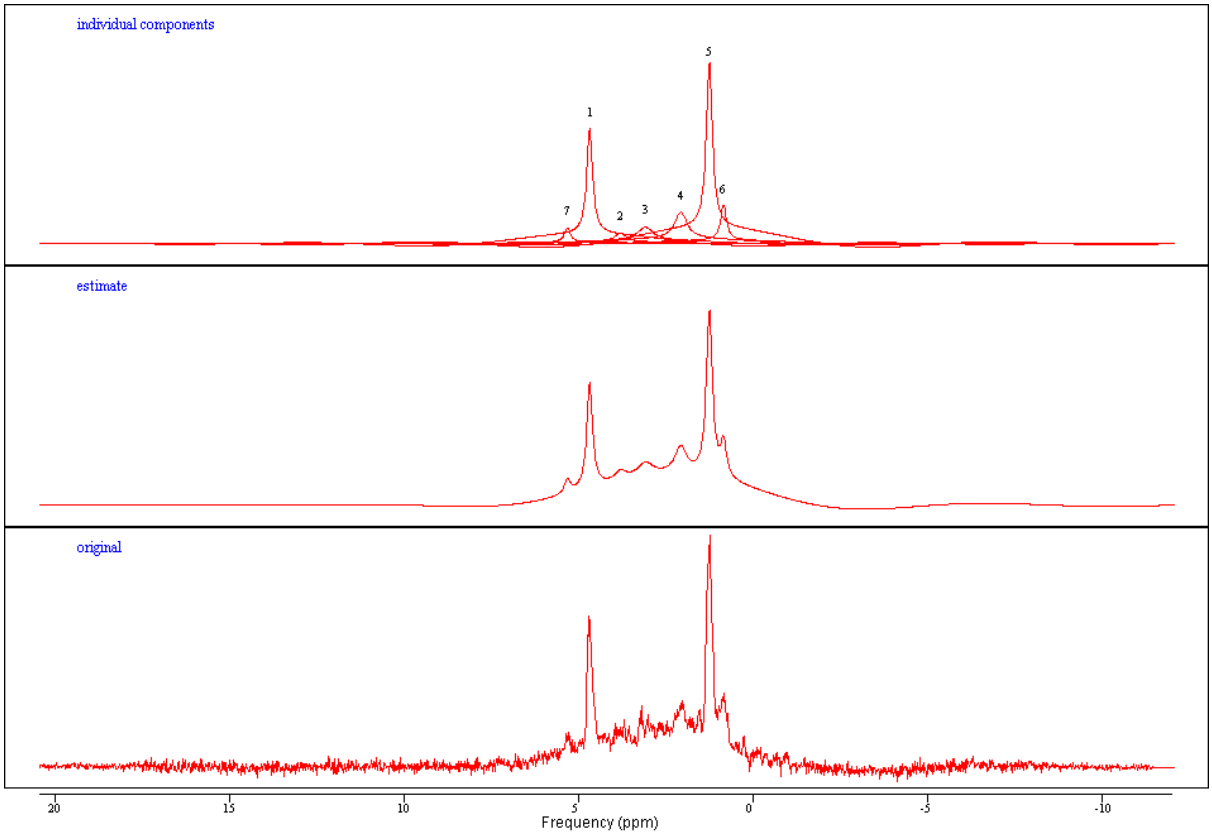


Figure 8. Liver water suppressed MRS data processing in jMRUI.

The original spectrum is shown in the bottom panel. Individual peaks were identified (middle panel) and then fitted to metabolites with known frequencies (top panel). The quantifiable peaks were:

1. Residual water,

2. Glycogen,

3: Choline,

4: Lipid; CH=CH-CH_2-CH_2; allylic,

5: Lipid; (CH_2)n; methylene,

6: Lipid; CH_3CH_2; methyl,

7: Lipid (methene)

3. Multi-parametric MR in healthy volunteers

3.1 Abstract

Background and Aims: Liver cT_1 is a newly established MR parameter that has not been sufficiently validated. The aim of this study was therefore twofold. Firstly to evaluate any technical determinants affecting the reporting of liver cT_1 and secondly to study liver cT_1 in healthy volunteers and identify its physiological determinants.

Methods: Healthy adult volunteers were recruited through poster advertisements to undergo a multi-parametric MR for liver cT_1 computation. Manual and automated data processing was performed on all scans to assess for any bias from the method of reporting. A subset of scans were reported twice to assess intra-observer variability. Associations with baseline variables were assessed and independent predictors of liver cT_1 were identified using multiple linear regression models.

Results: There was no bias in the cT_1 computation from the method of post-acquisition data processing or from intra-observer variability. Liver cT_1 varied significantly with age and gender with a median cT_1 of 784ms in women younger than 52 years compared to 741ms (p=0.038) in women older than 52 and 742ms in men younger than 52 (p=0.005). Liver cT_1 was independently associated with measures of intravascular blood volume including the ratio between the cross sectional area of the inferior vena cava and the aorta. Multivariate analysis was used to develop predictive models of liver cT_1 and based on these, a method was proposed for removing the confounding effect of liver congestion from liver cT1 measurements.

Conclusion: Liver cT_1 is a robust and reproducible measure that provides reliable results independent of the method of reporting and with no bias from intra-observer variability. Liver congestion resulting from intra-vascular blood volume expansion is the major physiological determinant of liver cT_1. A correction algorithm to remove this confounding effect may result in improved performance of liver cT_1 as a biomarker of liver fibrosis.

3.2 Introduction

Quantitative magnetic resonance (MR) T_1 techniques have been used for the characterisation of disease in the brain[121,122], heart[117,118], joints[123] and abdominal organs[124]. Similar MR techniques have shown some promise in differentiating patients with cirrhosis from healthy controls[70,71], but the application of these techniques have not seen widespread use in the assessment of diffuse liver disease. The development of a correction algorithm to remove the confounding effect of iron deposition, a frequent finding in patients with chronic liver disease, provided further diagnostic accuracy, allowing the differentiation between histological stages of fibrosis[114]. Despite these recent advances, liver T_1 and iron corrected T_1 (cT_1) have not been studied sufficiently in healthy subjects.

In order to optimise the use of liver cT_1 as a biomarker of liver disease, it was necessary to establish a normal range of cT_1 values in healthy volunteers. Furthermore, an evaluation of how physiological or technical factors influence the measurement of liver cT_1 was desirable as this could allow improved interpretation of this biomarker in disease states.

3.3 Aims

The main aim of this study was to establish a range of values of liver cT_1 in healthy volunteers and to identify any technical or physiological factors that may affect liver cT_1 in the absence of disease. The technical factors examined were variability between methods of data processing (LiverMultiScan vs manual reporting of T_1 and T_2* and use of look up tables to compute cT_1), and the intra- observer variability in the reporting of scans.

The effect of gender and age and the associations with liver fat, liver function tests, full blood count and anthropometric and anatomical data was also assessed. A further aim was to develop methods for removing the confounding effect of any physiological factors contributing to the variation liver cT_1.

3.4 Methods

3.4.1 Recruitment strategy

Participants were recruited through poster advertisements displayed at various locations in the John Radcliffe and Churchill hospitals in Oxford, UK. Participants were eligible if they were between 16 – 80 years old, with body mass index less than 25kg/m^2, with no history of liver disease or any components of the metabolic syndrome (type 2 diabetes mellitus, hypertension, hyperlipidaemia) and who did not drink alcohol in excess of the limit recommended by the UK department of Health (2-3 units /day for women, 3-4 units / day for men). Participants who responded to the poster advertisement were screened for the presence of contra-indications by telephone / email, and those who fulfilled the study criteria were invited to participate.

3.4.2 Multi-parametric MR

All participants had a multi-parametric MR as described in chapter 2, for the measurement of liver T_1, T_2*, cT_1 and liver fat content.

3.4.3 Measurement of the inferior vena cava and aorta cross sectional area

During the study, it became apparent that there was a variation in the size of the inferior vena cava (IVC) between participants and this was therefore studied systematically. Measuring tools on the scanner console were used to measure the cross sectional area of the intra-hepatic IVC and of the aorta (Ao). The Ao was measured as a reference for the size of the participant's vasculature, and the IVC/Ao ratio was computed (Figure 9).

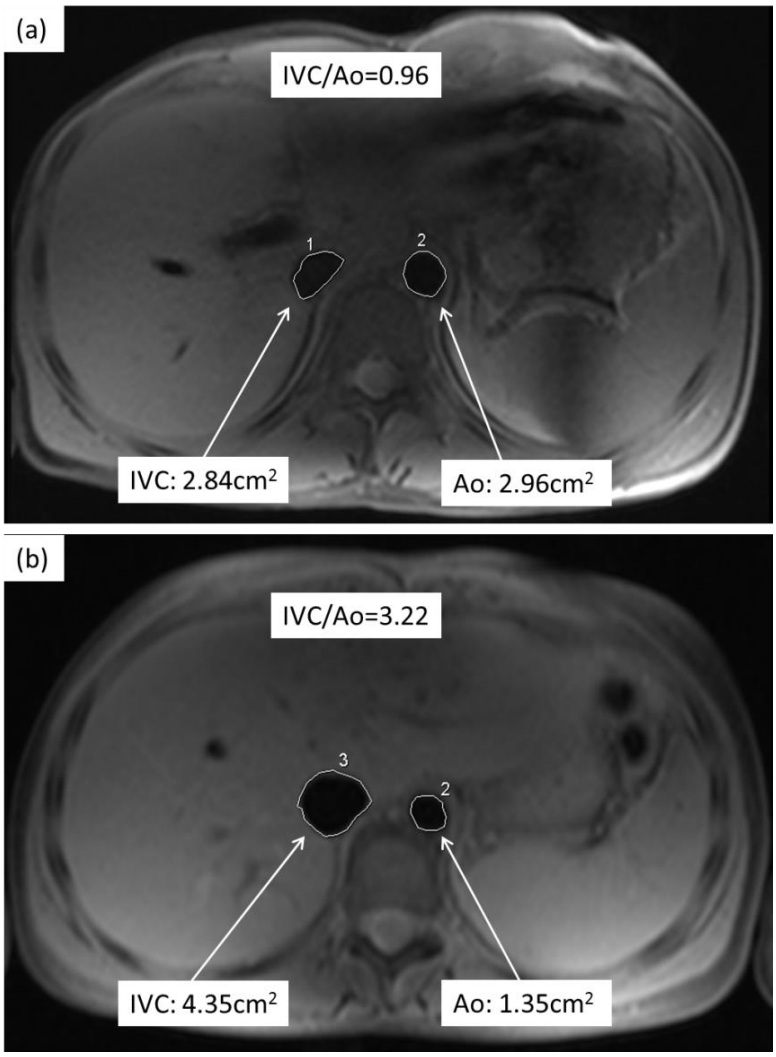


Figure 9. Measurement of the inferior vena cava and aorta cross sectional area.

The inferior vena cava (IVC) and aorta (Ao) cross sectional area was measured on the raw T_2* images, using tools on the scanner console. The figure illustrates the measurements for (a) a 30 year old male, where the IVC appears elliptical in cross section, and (b) a 23 year old female, where the IVC appears circular.

3.4.4 Other study procedures

Participants' height (m) and weight (kg) was measured and body mass index (BMI; kg/m^2) was computed. Waist and hip circumference (cm) were measured using a tape measure, and the waist to hip circumference ratio was computed. Waist circumference was measured with the subject standing and the measuring tape placed in the mid-point between the lowest rib and the anterior superior iliac spines. Hip circumference was measured as the maximum diameter around the buttocks.

Blood was collected for measurement of liver function tests (bilirubin, alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), gamma glutamyl transferase (GGT), albumin) and full blood count (haemoglobin, white cell count, platelets and haematocrit).

3.4.5 Variability in reporting of MR scans

All scans from the healthy volunteers were analysed using both automated reporting of cT_1 by LiverMultiScan (LMS; Perspectum Diagnostics, Oxford, UK) and manual reporting from separate T_1 and T_2* maps and use of look up tables to compute cT_1, as described in chapter 2. For the assessment of intra-observer agreement, 20 consecutive scans were analysed twice using LMS.

3.4.6 Statistical analysis

Descriptive statistics were used to summarise subject characteristics. Normality was tested using the Shapiro-Wilk's test. Bland Altman analysis was used for the assessment of intra-observer variability, and variability associated with the method of reporting (automated vs manual). Overall gender specific differences in liver cT_1 were tested using the Mann Whitney test. Male and female subjects further sub-classified into those less than or greater than 52 years old. The 52 year cut-off was chosen as this is the average

age of natural menopause in the UK[125]. Non-parametric analysis of variance with Dunn correction was used for the comparison between multiple groups.

Spearman's correlation analysis was carried out to identify factors associated with liver cT_1. The factors tested were: age, weight, height, waist and hip circumference, liver function tests, full blood count, liver fat content, and IVC and Ao cross sectional diameter. Factors which were significantly associated with cT_1 on the Spearman correlation analysis with $p<0.05$ were then entered into stepwise multiple linear regression models, to identify independent predictors of liver cT_1. The level of statistical significance was set at $p<0.05$. Analysis was conducted using the GraphPad Prism (v6.04) software and SPSS (v22; Armonk, NY, IBM Corp).

3.5 Results

3.5.1 Participants characteristics

Forty two participants (28 females; 67%) who fulfilled the inclusion / exclusion criteria were recruited. The median (IQR) age was 35 (27 – 55) years old and the median (IQR) BMI 22.3kg/m^2 (20.5 - 24.6). The median (IQR) values for: (a) liver fat content was 0.7% (0.5 – 1.7) and (b) T_2* was 21.3ms (17.6 – 24.4), corresponding to dry weight liver iron of 1.3mg/g (1.2 – 1.5). The median (IQR) liver cT_1 was 776ms (740 – 803). The demographics and blood results profiles of the participants are detailed in Table 3.

Table 3. Healthy volunteer characteristics

Demographics	
Age (years)	35 (27 – 55)
Body mass index (BMI; kg / m^2)	22.3 (20.5 – 24.6)
Waist circumference (WC; cm)	72 (67 – 81)
Hip circumference (HC; cm)	95 (90 – 100)
Waist / hip circumference ratio	0.76 (0.73 – 0.85)
Alcohol (units per week)	All: 8 (2 – 14); men: 11 (4-21); women: 5 (0 – 10)
Blood biochemistry / haematology	
Bilirubin (µmol/l)	11 (10 -15)
Alanine aminotransferase (ALT; IU/l)	15 (12 – 21)
Aspartate aminotransferase (AST; IU/l)	23 (19 – 28)
Alkaline phosphatase (ALP; IU/l)	129 (101 – 155)
Gamma glutamyl transferase (GGT; IU/l)	13 (11 -18)
Albumin (g/l)	*44.6 (±2.4)*
Haemoglobin (g/dl)	*13.8 (±1.2)*
Haematocrit (%)	*0.40 (±0.03)*
White cell count (x10^9/l)	5.3 (4.6 – 6.6)
Platelets count (x10^9/l)	*236 (±50)*

Median and interquartile range quoted for non-parametric data, and mean standard deviation for normally distributed data. Normally distributed data in italics.

3.5.2 Variability between methods of data processing

There was a strong association (r=0.92; R^2=0.84; p<0.0001) between liver cT_1 measurements obtained from the manual and automated (LMS) reporting methods. The mean (SD) difference between measurements was -7ms (28) and this was not significantly different from zero, indicating good agreement between the two methods.

Two points at the lowest range of liver cT_1 measures fall outside the level of agreement, suggesting that there may be a bias between the two methods when measuring low cT_1 values. This bias is unlikely to be of any significant clinical consequence as this method is intended for use in patients, where liver cT_1 is typically above 800ms. The scatter and Bland Altman plots for manual and automated liver cT_1 measurements are shown in Figure 10.

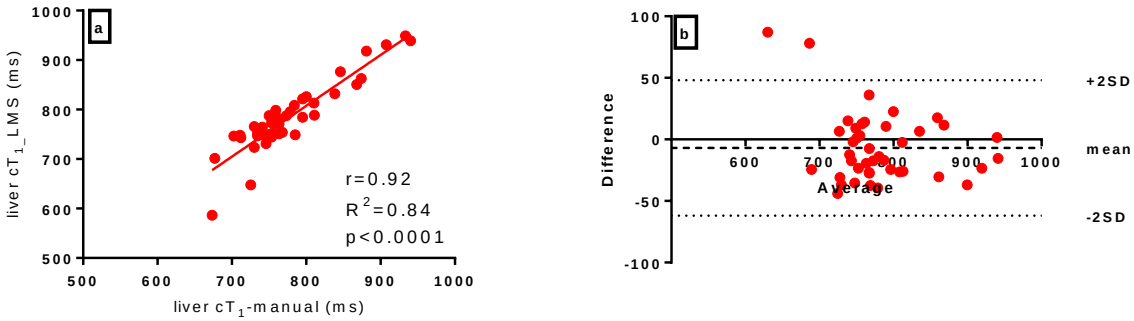


Figure 10. Variability between methods of data processing.

(a) Results from automated computation of liver iron corrected T_1 (cT_1) using LiverMultiScan (LMS) for image processing, plotted against manual computation of cT_1 from T_1 and T_2* values and look up tables. (b) Bland Altman plot for the agreement between the two methods of reporting. The difference between the manually and LMS reported liver cT_1 is plotted against the average of the two cT_1 measurements.

SD: standard deviation

3.5.3 Intra- observer variability

There was excellent intra-observer agreement for the reporting of liver cT_1 (r=0.99, R^2=0.97; p<0.0001). Scatter plots and Bland Altman plots for liver cT_1 reported by a single operator on two different occasions are shown in Figure 11.

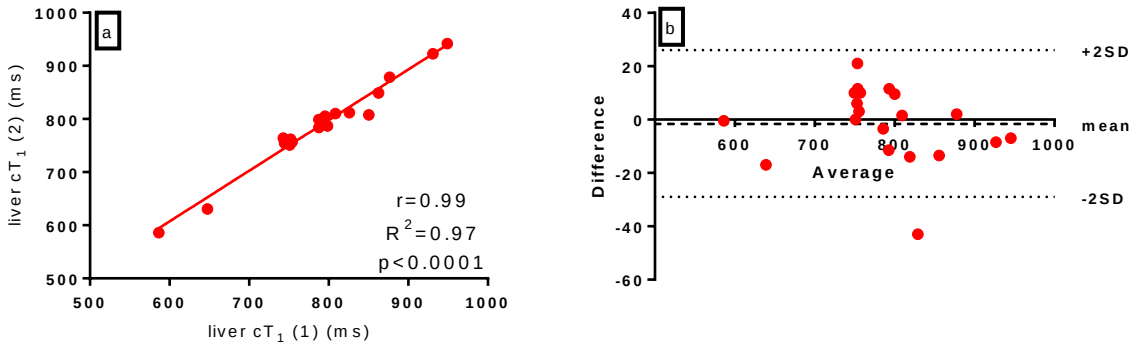


Figure 11. Intra-observer variability in liver cT_1 reporting.

(a) Scatter plot of liver iron corrected T_1 (cT_1) reported using LiverMultiScan (LMS) on two separate occasions, and (b) Bland Altman plot for intra-observer agreement in liver cT_1 reporting using LMS.

SD: standard deviation

3.5.4 Age and gender differences in liver cT_1

During the study it was noted that some premenopausal women who were otherwise healthy had high liver cT_1 values. To investigate this observation further, the cohort of healthy volunteers were subdivided according to gender and age. An age cut-off of 52 years was used as this is the average age of menopause in the UK.

There was an overall strong trend for a difference between the median (IQR) liver cT_1 of women (760ms; 745ms – 844ms) and men (752ms; 726ms – 772ms; p=0.07). When subjects were subdivided further into those younger and older than 52 years of age, women younger than 52 years old had a median liver cT_1 of 784ms, which was significantly higher that women older than 52 (741ms; p=0.038), and men younger than 52 (742ms; p=0.005). Men older than 52 (n=4) had similar liver cT_1 to women younger than 52 (790ms; p>0.05; Figure 12). Overall there was a significant negative association between age and liver cT_1 (r_s= -0.43; p=0.004), but this was driven exclusively by the relationship of age and liver cT_1 in women younger than 52 years (Figure 12).

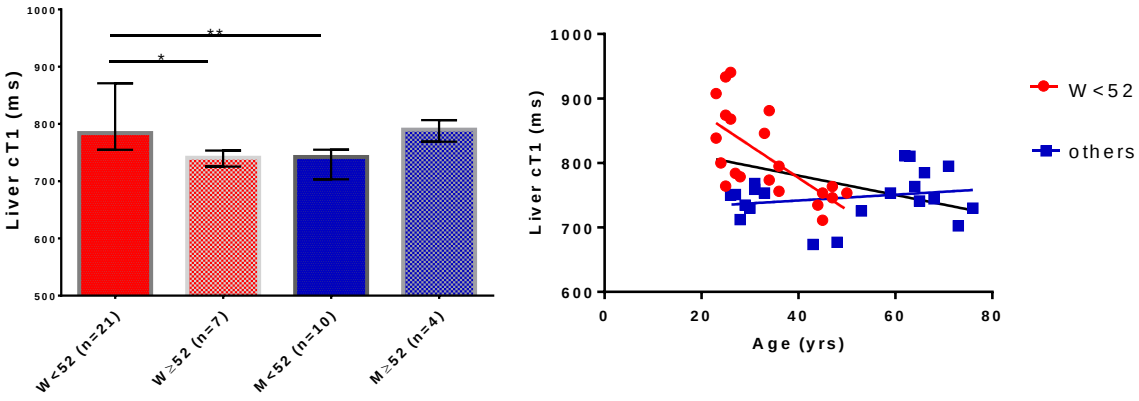


Figure 12. Variation of liver cT_1 according to age and gender.

(a) Women younger than 52 years of age (W<52) had a median (IQR) liver cT_1 of 784ms (755 – 871), which was higher than women older than 52 (W≥52; 741ms; IQR: 726 – 754; p=0.038) and men < 52 years old (M<52; 742ms; IQR: 703 – 755; p=0.005), but similar to men > 52 (M≥52, 790ms; IQR: 769 – 807; p>0.05). Bars indicate medians and interquartile range. (b) There was an overall significant association between liver cT_1 and age (r_s=-0.43; p=0.004) but this was driven by the relationship of age and liver cT_1 in women < 52years (r_s=-0.68; p<0.001). There was no significant association between liver cT_1 and age if the women younger than 52 were excluded (others; r_s=0.22; p=0.34). *p<0.05; **p<0.01

3.5.5 Additional predictors of liver cT_1

In addition to age and gender, there was a significant positive correlation with IVC cross sectional area and significant negative correlations with ALP, haemoglobin, haematocrit and Ao cross sectional diameter (Table 4).

Table 4. Factors with significant association to liver cT_1 on univariate analysis.

Variable	ALP (IU/l)	Hb (g/dl)	HCT (%)	IVC (cm^2)	Ao (cm^2)
Spearman's r	-0.56	-0.51	-0.54	0.42	-0.49
p	<0.001	0.001	0.002	0.006	0.001

Abbreviations. **ALP**: Alkaline phosphatase, **Hb**: Haemoglobin, **HCT**: haematocrit, **IVC**: inferior vena cava cross-sectional area, **Ao**: Aorta cross-sectional area.

On stepwise linear regression analysis, IVC and Ao cross sectional area and haematocrit remained independent predictors of liver cT_1 (Table 5). The graph of predicted liver cT_1 based on this model, plotted against the measured cT_1 is shown in Figure 13.

Table 5. Independent predictors of liver cT_1 on stepwise multiple linear regression

Model	B	Std. Error	Beta	t	p
(Constant)	1218	119.8		10.17	.000
HCT (%)	-1135	290.7	-0.489	-3.904	.001
IVC (cm^2)	20.38	6.422	0.386	3.157	.004
Ao (cm^2)	-18.40	8.661	-0.263	-2.124	.043

The equation for predicting linear liver cT_1 based on this regression analysis is:

Predicted liver cT_1= 1218 - 1135*HCT + 20.38*IVC - 18.40*Ao

Abbreviations: **HCT**: haematocrit, **IVC**: inferior vena cava. **Ao**: aorta

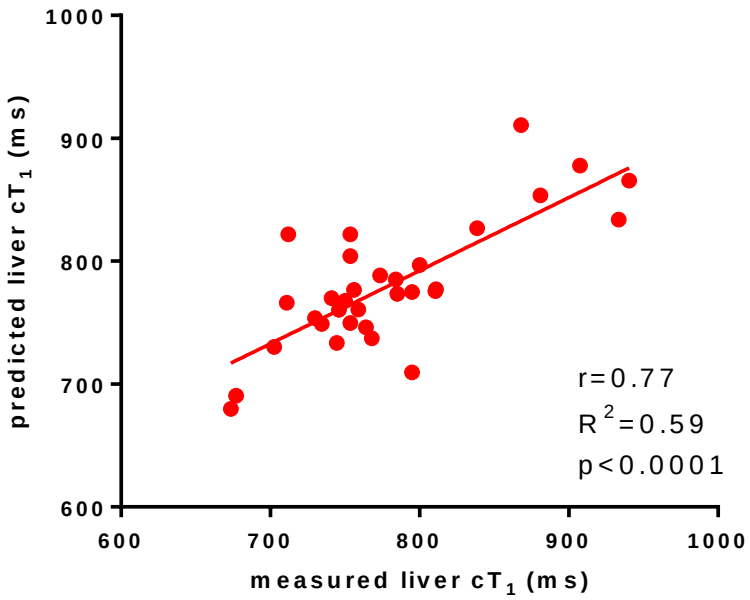


Figure 13. Predicted vs measured liver cT_1 in healthy subjects.

Predicted liver cT_1 = 1218 – 1135*haematocrit + 20.38*IVC - 18.40*Ao (Table 5)

Haematocrit was measured as a percentage, and the inferior vena cava (IVC) and aorta (Ao) cross sectional area was measured in cm^2.

3.5.6 Methods for correcting for the confounding effect of liver congestion

One of the main strengths of the liver T_1 mapping technique used in this project is the ability to remove the confounding effect of liver iron on T_1 through the application of a correction algorithm[113,114]. In a similar fashion, the linear regression equation including haematocrit, IVC and Ao diameter (Figure 13) could be used to correct for the effect of liver congestion in any liver T_1 / cT_1 measurements. This approach would probably yield the most accurate method of correction, but it may be impractical as it requires data from a blood test (haematocrit). Therefore, in order to simplify this correction algorithm, the multiple linear regression analysis was repeated, with IVC and Ao as the independent variables, or with the ratio of IVC/Ao as a single variable (Table 6).

Table 6. Multiple linear regression including only MR-derived variables.

Variable	B	P	R^2
Model 1			0.40
Constant	770	<0.001	
IVC	23.6	0.001	
Ao	-27.0	0.004	
Model 2			0.52
Constant	694	<0.001	
IVC/Ao	58.5	<0.001	

Abbreviations: **IVC**: inferior vena cava cross sectional area, **Ao**: aorta cross sectional area.

The IVC/Ao ratio was chosen as a variable to explore based on the results of the stepwise linear regression where associations with liver cT_1 was positive for IVC and negative for aorta cross sectional area (Table 5). Furthermore, in two recent studies, the IVC/Ao ratio has been used as a surrogate of fluid status in children,[126] and adults with pericardial constriction.[127] The prediction model using IVC/Ao resulted in a higher R^2 value (0.52 vs 0.40) so was explored further.

To establish which data would need correction for the effect of intravascular blood volume, the frequency distribution of IVC/Ao values was examined and found to be non-parametric, skewed by the presence of high IVC/Ao values (Figure 14). The median IVC/Ao was 1.26 and the 25th and 75th percentiles were respectively 0.78 and 1.89. It is therefore suggested that liver cT_1 should only be corrected in cases where the IVC/Ao ratio, is greater than 1.89 (75th percentile).

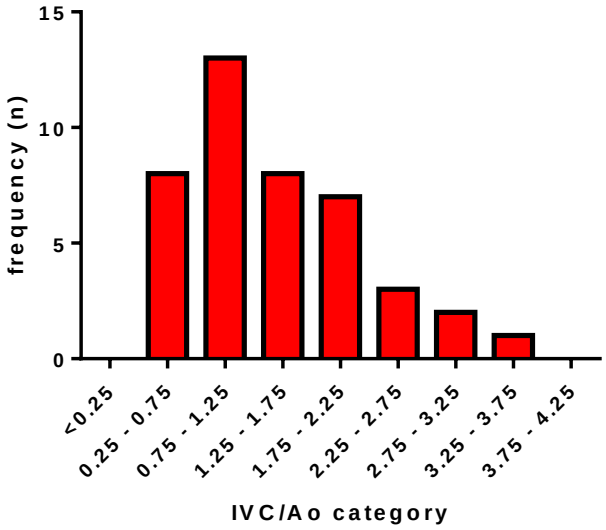


Figure 14. Frequency distribution of IVC/Ao ratio in the cohort of healthy volunteers

The purpose of the correction would be to remove the effect of liver congestion from the liver cT_1 measurement. It is suggested that the IVC/ Ao ratio is used as a measure of intravascular blood volume which in turn is an indirect measure of liver congestion. It is further proposed that the new parameter is termed "**blood volume adjusted and iron corrected T_1" ($vacT_1$),** which could be interpreted as the liver cT_1 measurement that would have been obtained if there was no liver congestion (i.e. if the IVC/Ao ratio was 1.26, which is the median value in the population from which this algorithm is derived).

The method for applying this adjustment is illustrated in Figure 15. An example of how this correction would be computed is shown in Figure 16, and the application of this to the whole cohort of healthy volunteers in shown on Figure 17.

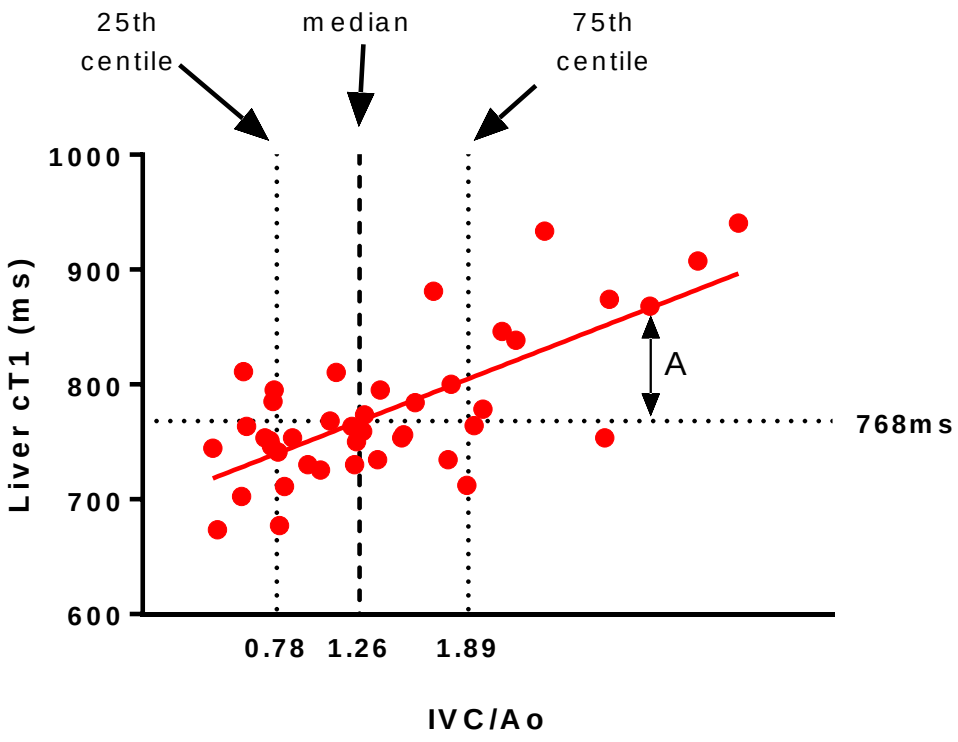


Figure 15. Developing a correction algorithm for the removal of the confounding effect of liver congestion.

In patients where the IVC/Ao ratio is greater than 1.89 (75th percentile), the measured liver cT_1 should be adjusted to yield the blood volume adjusted cT_1 ($vacT_1$), using the following equation:

$vacT_1$ = measured liver cT_1 – A, where:

A= (predicted liver cT_1 for the measured IVC/Ao) – (predicted liver cT_1 for an IVC/Ao of 1.26)

The predicted liver cT_1 for any IVC/Ao ratio is derived from the linear regression equation:

Predicted liver cT_1 = 694 + 58.5*IVC/Ao (Table 6)

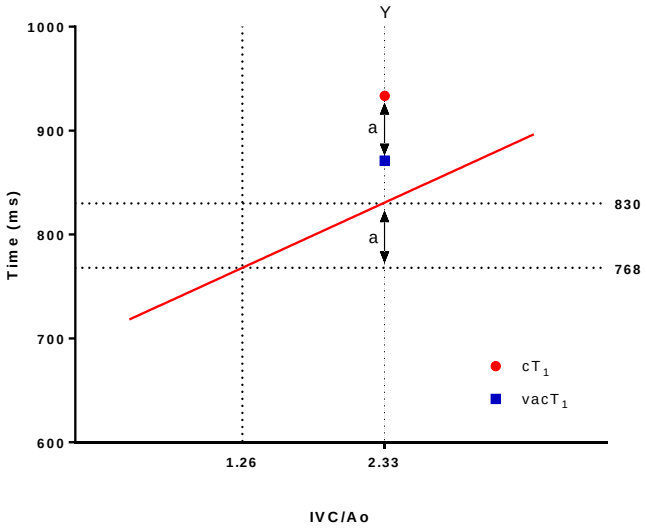


Figure 16. Example of the application of the algorithm for intravascular blood volume adjustment

The measured liver cT_1 for this participant (labelled Y) is 934ms (red dot) and the IVC/Ao ratio is 2.33. The steps in correcting this would be as follows:

1. Derive the predicted liver cT_1 for the IVC/Ao of the participant, based on the linear regression equation: Predicted liver cT_1 = 694 + 58.5 x IVC/Ao

Predicted liver cT_1 = 694+58.5x2.33

Predicted liver cT_1 = 830ms

2. Compute the difference between the predicted liver cT_1 for this participant and the predicted liver cT_1 for the median IVC/Ao ratio of 1.26 in the population from which this algorithm was derived. This is labelled as “a” on the graph and can be thought of as the degree to which liver cT_1 is affected by liver congestion.

a= 830 – (694+58.5x1.26)

a= 830 -768

a=62ms

3. Subtract “a” from the measured liver cT_1 of the participant to derive $vacT_1$. In this case the measured cT_1 is 934ms.

$vacT_1$= measured cT_1 –a

$vacT_1$ = 934 – 62

$vacT_1$ = 872ms

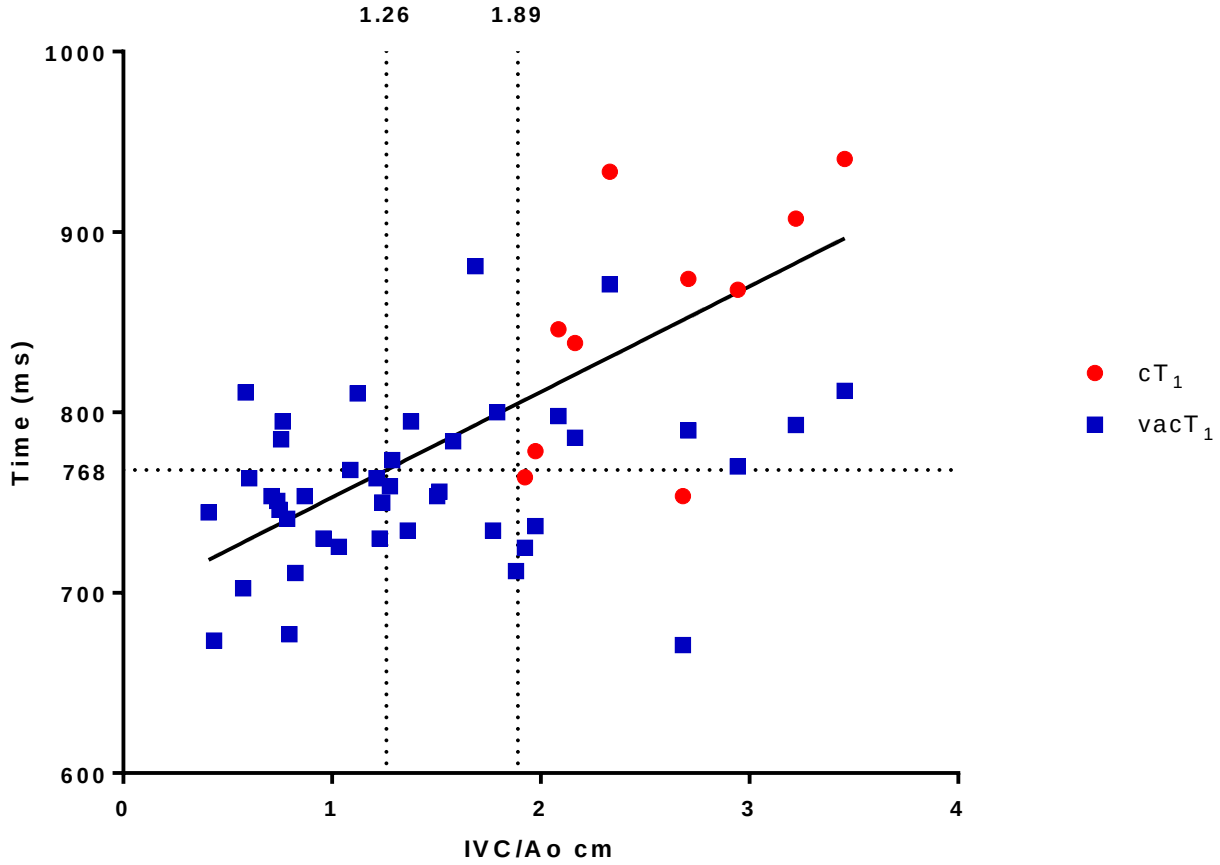


Figure 17. Application of the blood volume adjustment algorithm to healthy volunteers.

Blood volume adjusted iron corrected T_1 ($vacT_1$) has a normal distribution with a mean (±SD) of 760ms (±45). The grid lines for the median and 75th percentile IVC/Ao ratio (1.26 and 1.89 respectively) and, predicted cT_1 for the median IVC /Ao ratio (768ms) are shown for reference. The uncorrected data for those with an IVC/Ao ratio >1.89 are also shown as red dots.

3.5.7 Effect of aortic size

The size of the aorta is known to increase with age[128], and this could introduce bias in the proposed correction algorithm. In the cohort of healthy volunteers studied here, aorta cross sectional area had a non-parametric distribution, with the tail of the frequency curve

in the higher values of aortic cross sectional area (Figure 18). There was also a positive association of aortic cross sectional area with age (r_s=0.73, p<0.0001). The strength of the association was similar for men (r_s=0.75, p=0.0027) and women (r_s=0.81, p<0.0001). Women had a median (IQR) aortic cross sectional area of 2.24cm^2 (1.93 – 2.67), which was lower than men (2.69cm^2; IQR: 2.40 – 3.52; p=0.024). Three men and seven women had aortic cross sectional areas higher than the 75th percentile for their gender (Figure 19).

The increasing aortic size will compromise the performance of the correction algorithm for vacT_1 as this relies on the IVC/Ao ratio. It is therefore proposed that if the aorta cross sectional area is greater than the gender specific 75th percentile, then the gender specific median Ao cross sectional area should be used when computing the IVC/Ao ratio. When this approach was applied to the cohort of healthy volunteers, one further patient was identified where the measured cT_1 needed to be adjusted for blood volume. This patient had a measured IVC of 5.48cm^2 and a measured Ao of 4.88cm^2. Based on the measured parameters the IVC/Ao ratio was 1.12, however when the gender specific median Ao cross sectional area was used as the denominator, the IVC/Ao ratio was calculated as 2.04. The latter value is higher than 1.89, suggesting some contribution from congestion, which would require the application of the blood volume adjustment algorithm. In this case the measured cT_1 would decrease from 810ms to a vacT_1 of 765ms.

Age related or pathological aortic dilatation, can therefore affect the accuracy of this correction algorithm, and this can be problematic when applied to patients, who tend to be older and more likely to have pathological or age related aortic dilatation. If the aorta cross sectional area is greater than the gender specific 75th percentiles established here, the liver congestion effect may be underestimated or missed. In these cases the

measured Ao cross sectional area should be substituted with the gender specific median Ao cross sectional area as the denominator when computing the IVC / Ao ratio.

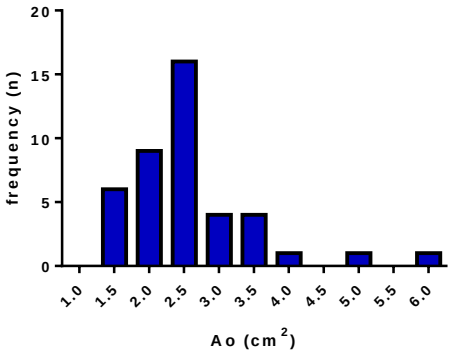


Figure 18. Frequency distribution of aorta (Ao) crosses sectional area in healthy volunteers.

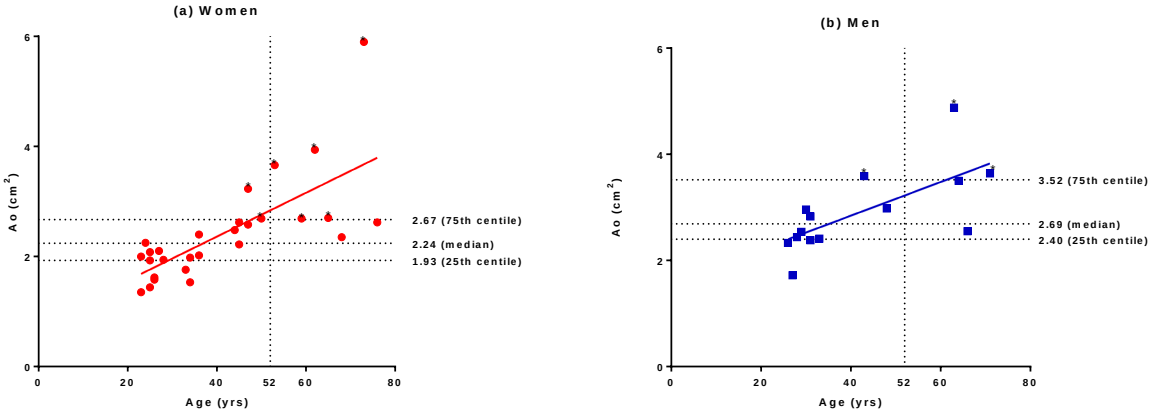


Figure 19. Aortic cross sectional area plotted against age for (a) women and (b) men.

The Ao cross sectional area is greater than the gender specific 75th percentile in seven women and three men (marked with asterisks). The median and 25th and 75th percentiles as well as the 52 year age cut-off are shown for reference.

3.6 Discussion

This work has demonstrated that liver cT_1 is a robust, highly reproducible measure, with no significant variability in the post processing of images, and excellent intra-observer agreement. Furthermore, it has established that liver congestion due to increased intravascular blood volume is the most likely physiological determinant of liver cT_1, as this can explain most of the variability in healthy volunteers. As a result of this observation, a method for removing the effect of liver congestion from the measured liver cT_1 has been described, and this is likely to improve the diagnostic accuracy of liver cT_1, as a biomarker of liver fibrosis and inflammation.

3.6.1 Interpretation of results for the variation of liver cT_1 in healthy controls

This study was conducted in a well characterised cohort, with no risk factors for liver disease, normal liver fat content measured using ^{1}H-MRS, normal liver iron measured by T_2* and normal liver function tests. Based on these observations, it can be assumed with reasonable confidence, that participants did not have liver fibrosis or inflammation that would account for high liver cT_1 values. Despite this, there was variability in the values of liver cT_1 (median: 761ms, IQR: 739 - 803), which was largely explained by the variability in premenopausal women, (women younger than 52 years old) where the median liver cT_1 (IQR) was 784ms (755 – 871), compared to the rest of the participants where median (IQR) liver cT_1 was 750ms (728 – 766).

It is well established that there are haemodynamic and plasma volume changes during the menstrual cycle, where hormonal changes in the luteal phase result in volume expansion, and haemodilution, mirroring changes of early pregnancy[129,130]. This can lead to significant weight gain, which typically peaks at the onset of menstrual flow,[131,132] and which can be perceived as ankle swelling or abdominal bloating[133]. Similarly, oral contraceptives can also lead to fluid retention and haemodilution[130]. Fluid shifts and liver

congestion during the menstrual cycle or due to hormonal contraception are therefore the most likely explanation for the variability of liver cT_1 in premenopausal women.

There has only been one other study where liver T_1 was systematically studied in healthy volunteers[134], and its findings were very similar to the ones reported here. In that study, 81 healthy volunteers (33 men with mean age of 38 and 48 women with mean age of 40) were studied using a 0.08T scanner. There was a negative association of liver T_1 with age ($r=-0.51$, $p<0.001$), and a higher variability in subjects younger than 40 years old (SD=13ms, compared to SD of 8ms in subjects older than 40), but the investigators did not do any analysis for differences between genders in the two age groups. The study also found (a) significant differences in the liver T_1 of women who were taking the oral contraceptive pill (OCP; 217ms) compared to those who were not (198ms; $p<0.0002$), (b) that in women who were taking the OCP, liver T_1 was significantly higher during the 3 weeks on the OCP (229ms) compared to the week off the OCP (211ms; $p=0.001$), and (c) that in women who were not on the OCP, there were no significant differences in liver T_1 between the week of menstruation and the rest of the menstrual cycle.

These results suggest a clear role for the OCP in determining liver T_1. The non-significant result of the study examining the variation during the menstrual cycle is difficult to interpret, as it examined the differences between the week of menstruation compared to the rest of the cycle, whereas a more physiologically meaningful comparison would be to compare the liver T_1 during the follicular (high oestrogen, low progesterone) and luteal (high progesterone, low oestrogen) parts of the cycle.

Further support for the effect of female reproductive hormones on tissue T_1 relaxation times comes from studies examining T_1 in other organs. In particular, a study of

myocardial T_1 in healthy volunteers found that women younger than 45 years had higher myocardial T_1 than age matched men.[135]

The multivariate analysis conducted here (Table 5) provides further evidence that liver congestion due to physiological fluid retention and intravascular blood volume expansion is the main determinant of liver cT_1 in the absence of fibrosis or inflammation. In this analysis, all the independent predictors of liver cT_1 were associated with blood volume changes (haematocrit, IVC, Ao). IVC cross sectional area would increases with volume expansion and was positively associated with liver cT_1, while haematocrit which would decrease with plasma volume expansion was negatively associated with liver cT_1.

Studies examining the variation of liver cT_1 and associated parameters in the luteal and follicular phases of the menstrual cycle or in women using hormonal contraception would be needed to answer this question fully, but these were beyond the scope of this project.

3.6.2 Implications for the application of liver cT_1 as a biomarker

Even though data on the menstrual cycle or hormonal contraception usage were not collected here, the inclusion of premenopausal women in this cohort has provided valuable insight into the effect of liver congestion on liver cT_1. Liver cT_1 is proposed as a biomarker of liver fibrosis and inflammation, therefore, removal of any confounding effects from liver congestion is likely to improve the diagnostic accuracy for liver fibrosis. This would be important when studying premenopausal women, where liver congestion appears to be a common feature, or in patients at risk of fluid overload from cardiac, renal or other pathological processes. Furthermore, in order to minimise this effect, premenopausal women should be preferably scanned during the follicular phase of the menstrual cycle (within 14 days from the end of menstruation), and this would be

particularly important if serial liver cT_1 measures are being taken for disease monitoring. In addition, information about use of hormonal contraception should be collected at each scan, as this could aid interpretation of the MR data.

In conclusion, the results reported here confirm some of the earlier findings and add to our knowledge of liver cT_1 in healthy volunteers. In particular, the present study established that the relationship with age is driven exclusively by the variability of liver cT_1 in premenopausal women. Furthermore, a plausible explanation is put forward for the mechanism by which female reproductive hormones may be exerting their effects through fluid shifts and changes in the intravascular blood volume reflected in liver congestion. In an additional step, a model is proposed that can be used to remove the confounding effect of liver congestion, with the aim of making liver cT_1 a more accurate biomarker of liver fibrosis and inflammation.

4. Assessment of liver fibrosis

4.1 Abstract

Background and method: Non-invasive biomarkers of liver fibrosis are generally assessed against traditional categorical scores of liver fibrosis. However, this approach is limited in part due to the drawbacks of the categorical scores including observer dependent variability. The aim of this study was to develop a method for assessing liver fibrosis using morphometry and use it as a reference comparator for liver cT_1.

Methods: Digital imaging analysis was used to measure collagen proportionate area (CPA) in liver biopsy slides stained with Sirius red. Inter-observer and intra-observer variability of CPA was assessed. The association of liver cT_1 with CPA as well as other histological variables was examined.

Results: Fifty four patients had paired MR data and liver histology slides stained with Sirius red. Most patients were male (76%) with an average age of 50.4 years. CPA was strongly associated with the Ishak stage (r_s=0.84; p<0.0001) and was not subject to bias from intra- or inter observer variability.

Liver cT_1 was significantly associated with CPA in patients with NAFLD (r=0.52; p=0.013) and patients with viral hepatitis (r=0.80; p<0.0001). Histologically assessed liver fat was also associated with liver cT_1 (r=0.42; p=0.002), indicating a confounding effect from steatosis in the application of liver cT_1 as a biomarker of liver fibrosis. The inclusion of liver fat into multi-variate predictive models improved the overall performance of multi-parametric MR.

Conclusion: Collagen proportionate area is a reliable method for assessing liver fibrosis histologically. Liver cT_1 is strongly associated with CPA particularly in patients with viral hepatitis, but the performance of liver cT_1 as a biomarker of liver fibrosis can be further improved by accounting for the confounding effect of liver fat in multivariate models.

4.2 Introduction

Fibrosis represents the final common pathway of injury in chronic liver pathologies including chronic viral hepatitis and non-alcoholic fatty liver disease (NAFLD). The assessment of liver fibrosis is therefore, central to the evaluation of patients with liver disease, irrespective of the underlying aetiology. Liver biopsy remains one of the most widely used methods to achieve this. Differences in the pathophysiology of each liver disease can result in distinct patterns of fibrosis distribution, and this has led to the development of disease specific semi-quantitative categorical scores like the Ishak score[120] and the METAVIR score[136] for viral hepatitis or the scoring system for fibrosis developed by the non-alcoholic steatohepatitis clinical research network (NASH CRN) for NAFLD[108].

Categorical scores classify biopsies according to the presence of specific histological features (e.g. fibrotic bridges between portal tracts or cirrhotic nodules). It is however increasingly recognised that these categorical scores are not sensitive to changes in the total fibrosis burden and may not be suitable for monitoring disease progression or regression. Morphometrically quantified collagen content is much more sensitive in detecting changes in fibrosis. In a study examining paired biopsies in patients with hepatitis C, fibrosis content assessed by morphometry increased by 58% without any change in the categorical fibrosis stage[137]. Morphometrically assessed liver collagen is increasingly recognised as a robust measure of liver disease severity that correlates with portal pressure[115], and predicts outcomes in patients with chronic hepatitis C[138] and cirrhosis[139].

The development of non-invasive tests for fibrosis assessment is also hampered by comparison with categorical scores. Non-invasive tests provide estimates of the amount

of fibrosis in the liver and evaluating them against categorical scores may not provide an accurate representation of their diagnostic accuracy. Comparison with quantitative fibrosis scores may therefore provide more meaningful information.

4.3 Aims

The main aims of the study were to develop and validate a digital imaging analysis technique for the morphometric quantification of liver fibrosis, and to use this as the reference standard for the comparison with liver cT_1. Secondary aims were to identify any factors that may be confounders in the assessment of liver fibrosis by cT_1, and to investigate whether multivariate models could improve the diagnostic performance of cT_1.

4.4 Patients and methods

4.4.1 Study design and population

Adult patients who were scheduled to have a clinically indicated liver biopsy for the assessment fibrosis and liver disease aetiology were prospectively recruited. Patients were eligible irrespective of the underlying liver disease aetiology, if they had liver biopsy sections stained with Sirius red for fibrosis assessment. Patients underwent multi-parametric MR to measure liver, T_1, T_2*, cT_1 and liver fat content, as described in Chapter 2. The reference standard for liver fibrosis was collagen proportionate area assessed using digital imaging analysis. MRI data were acquired and analysed before the liver biopsy. The median (IQR) interval between MRI and liver biopsy was 7 days (2 – 18).

4.4.2 Routine clinical reporting of biopsies

Biopsies were processed according to the local clinical routine, which involves review by two pathologists and a discussion in a clinico-pathological meeting before a final consensus report is issued. The Ishak staging (0-6) and grading (0-18) system was used to report fibrosis and necro-inflammation respectively[120]. Liver fat was quantified as a proportion of liver cells containing lipid droplets (%) and graded according to the NASH CRN system (0-3)[108]. Liver iron was assessed on Perl stained slides using a 5-tier grading system (0-4)[140]. The reporting pathologists were blinded to the MRI data.

4.4.3 Digital imaging analysis for collagen proportionate area

Liver histology slides stained with Sirius red were scanned using a Hamamatsu Nanozoomer 2.0 HT Digital Pathology System (Hamamatsu, Hamamatsu City, Japan) to produce high quality digital images. All images were processed using the ImageJ software (U. S. National Institute of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, v1.47).

A digital imaging analysis (DIA) technique was developed to quantify the collagen proportionate area (CPA) based on the published literature[115]. Biopsies were jointly reported with an experienced liver pathologist who was blinded to the MR data. This jointly reported CPA was used as the reference for the comparison with liver cT_1 and for the assessment of inter-observer agreement (reference-CPA). Low magnification images were used so that the whole biopsy sample was visible in one frame. The digital images were manually cropped to remove artefacts in the background and any non-liver tissue (e.g. skeletal muscle). The images were then split into the RGB channels and the green channel was used for further processing. The in-build "threshold" function of the software was used to estimate the total area of liver tissue and the total area of collagen. An outline of the collagen selection was then generated and this was super-imposed on the original image allowing the accuracy of the selection to be checked against the Sirius red staining. The area of collagen was expressed as a percentage of the total biopsy area (collagen proportionate area; CPA). Figure 20 illustrates the digital image processing steps for the estimation of CPA.

4.4.4 Variability in CPA reporting

To assess inter- observer variability in the reporting of CPA, two further liver pathologists were trained in the digital imaging analysis technique and assessed a subset of 20 slides. The training involved a 45 minute session where the use of the software and the DIA technique was demonstrated. A leaflet with step wise instructions of how to use the software was also supplied to the pathologists. The pathologists were provided with a training set of 5 slides and were asked to use these to practice the DIA technique until they were confident in the use of the software. The 5 training slides were different from the 20 used in the final analysis. One pathologist assessed the 20 slides twice to evaluate intra-observed variability.

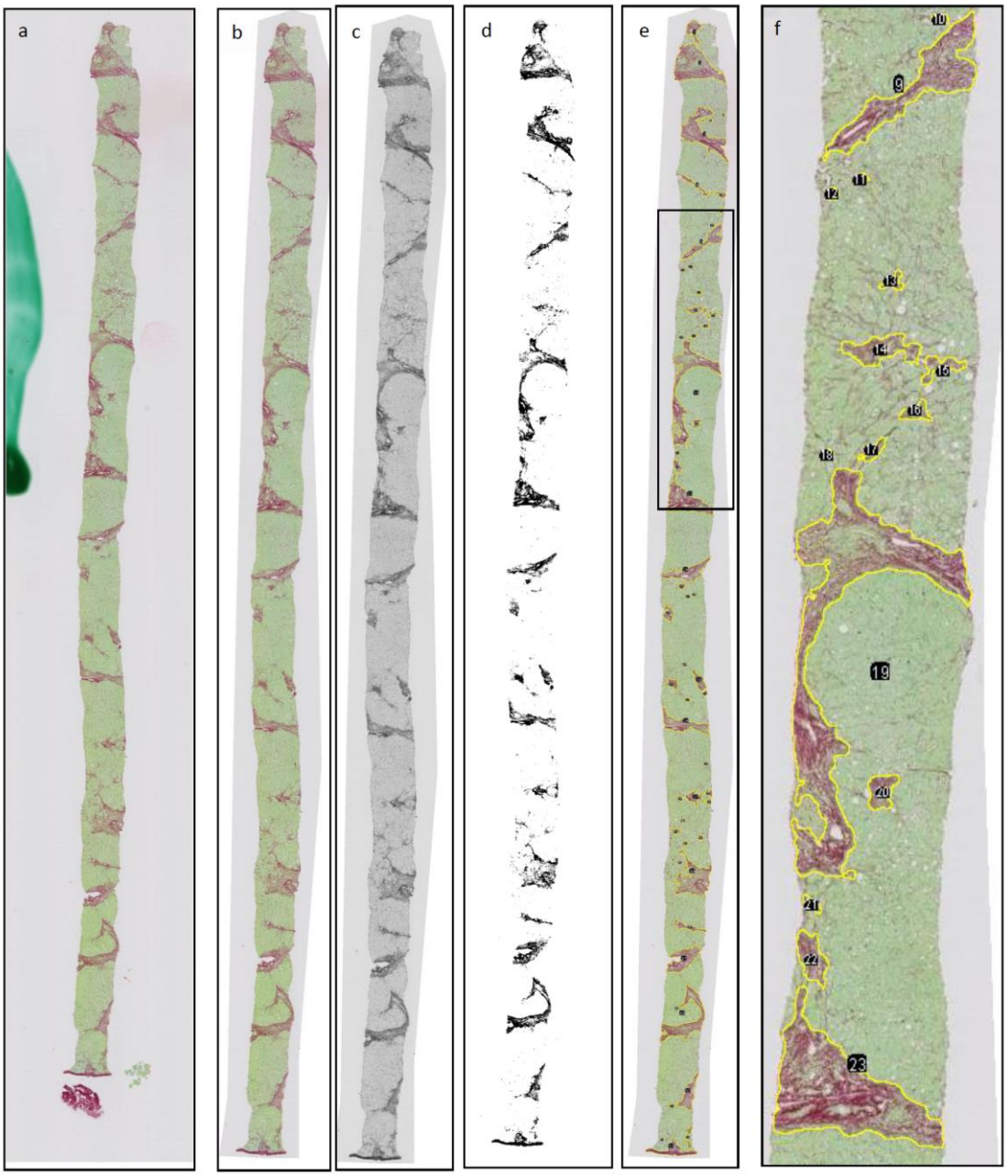


Figure 20. Digital imaging analysis for collagen proportionate area.

The figure illustrates a liver biopsy section stained with Sirius red (a) before and (b) after manual cropping, (c) the selection of the green channel of the RGB components, (d) manual thresholding to select the collagen, and a further step that allows the collagen selection (yellow outline) to be superimposed onto the original image (e and magnified section in f).

4.4.5 Liver biopsy quality

Biopsies had a median (IQR) length of 20mm (16 – 30) and contained a median (IQR) of 11 (8 – 15) portal tracts. A previous study has shown that liver biopsy length affects the accuracy of CPA measurement and that the ideal biopsy length is at least 28mm[141]. Applying this quality criterion to the present study would have excluded most of the biopsies. Therefore, to establish a minimum acceptable biopsy length for the purposes of this study, the frequency distribution of biopsy length was examined (Figure 21), and based on this, biopsies less than 12.5mm were excluded from the final analysis. With this approach, the smallest biopsies which may give unreliable CPA results were excluded, while keeping enough biopsies for the analysis to be meaningful.

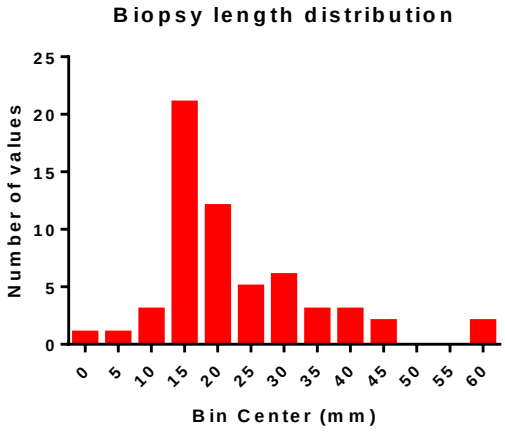


Figure 21. Frequency distribution of biopsy length

4.4.6 Statistical analysis

Descriptive statistics were used to summarise subject characteristics. Normality was tested using the Shapiro-Wilk's test. Bland Altman analysis was used for the assessment of intra- and inter-observer variability in CPA reporting. Correlations between variables were assessed using the Pearson and Spearman coefficients as appropriate. Associations of liver cT_1 with the histological variables of steatosis, iron deposition and inflammation were assessed to identify confounders. To examine whether aetiology introduced a bias, linear regression was used to evaluate and compare the relationship between CPA and cT_1 in steatohepatitis and viral hepatitis. Multiple linear regression was carried out to assess whether inclusion of all 3 MR variables (cT_1, T_2* and liver fat content), could improve the diagnostic accuracy for CPA. The level of statistical significance was set at $p<0.05$. The analysis was carried out using GraphPad Prism (v6.04) and SPSS (v22; Armonk, NY, IBM Corp).

4.5 Results

4.5.1 Cohort characteristics

Sixty patients had biopsies stained with Sirius red. One patient had severe iron overload beyond the limits of MR quantification, and this precluded computation of liver cT_1. Five patients were excluded due to poor quality biopsies (length <12.5mm). The majority of patients were male (n=41; 76%) and the mean age was 50.4 years. Most patients had mild fibrosis (Ishak F0-2; n=34; 63%) and the main liver disease aetiologies were steatohepatitis (n=22; 41%) and viral hepatitis (n=18; 33%). Table 7 summarises the cohort characteristics.

Table 7. Characteristics of patients who had collagen proportionate area quantification

Demographics	
Male, n (%)	41 (76)
Mean (±SD) age (yrs)	50.4 (±14.3)
Median (IQR) BMI	26.9 (23.9 – 29.4)
Liver disease aetiology, n (%)	
Steatohepatitis	22 (41)
Viral hepatitis	18 (33)
Other	14 (26)
Liver fibrosis stage, n (%)	
0-2	34 (63)
3-4	10 (14)
5-6	10 (14)
Biochemistry / haematology, median (IQR)	
Bilirubin (µmol/l)	11 (7 - 17)
Alanine Aminotransferase (IU/l)	60 (35 – 97)
Aspartate aminotransferase (IU/l)	49 (30 – 63)
Gamma glutamyl transferase (IU/l)	85 (37 – 163)
Alkaline phosphatase (IU/l)	209 (164 – 311)
Albumin (g/l)	44 (42 – 46)
Prothrombin time (s)	13.7 (12.9 – 14.9)
Platelets ($x10^9$/l)	193 (154 – 248)

4.5.2 Collagen proportionate area

There was a strong association between CPA and the Ishak stage (r_s=0.84; p<0.0001; Figure 22). Furthermore, there was good inter- and intra-observer agreement in reporting of CPA with the digital imaging analysis technique developed for this study (Figure 23).

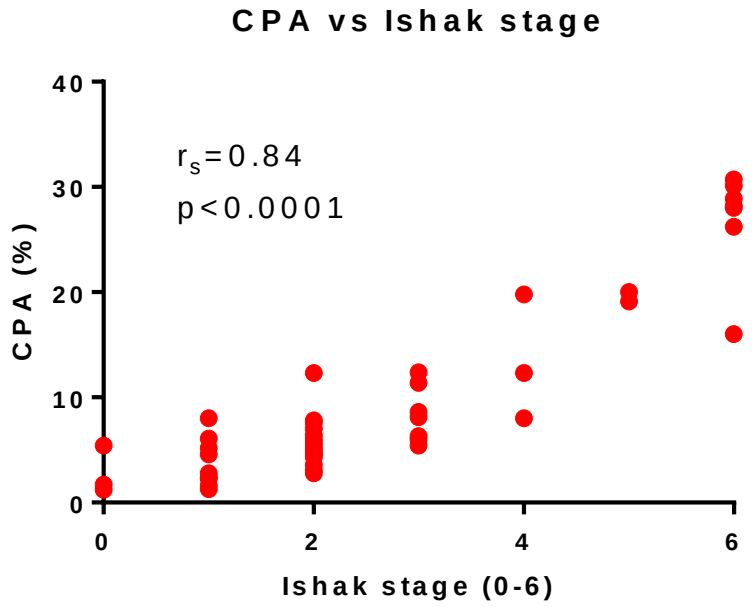


Figure 22. Association between collagen proportionate area (CPA) and the Ishak stage

4.5.3 Liver cT_1 histological associations

There was a significant association between liver cT_1 and CPA (r=0.54; p<0.001). There was also an association with the proportion of hepatocytes including lipid droplets (r=0.42; p=0.002) and the liver steatosis grade (r_s=0.42; p=0.002), but there was no association with liver iron deposition or with necro-inflammation as assessed by the Ishak grade (Table 8).

Table 8. Associations between liver cT_1 and histological variables

	CPA (%)	Steatosis (%)	Steatosis grade (0-3)	Ishak grade (0-18)	Iron deposition (0-4)
Correlation	r=0.54	r=0.42	r_s=0.42	r_s=0.04	r_s=-0.004
p value	<0.001	0.002	0.002	0.79	0.98

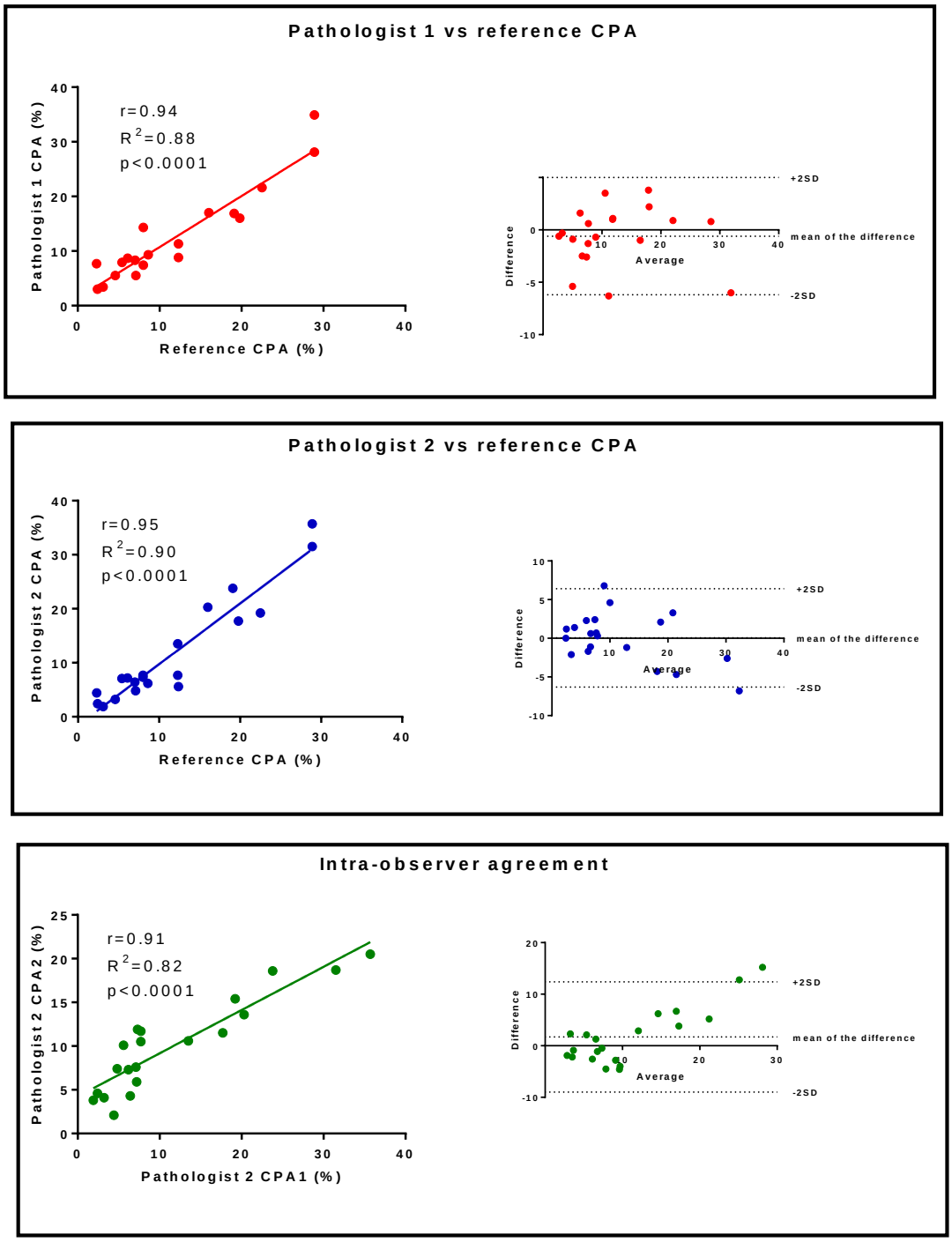


Figure 23. Observer dependent variability in reporting of collagen proportionate area.

Scatter plots and Bland Altman plots for the agreements between pathologist 1 and the reference CPA (top panel), pathologist 2 and the reference CPA (middle panel), and for CPA analysed by pathologist 2 on two occasions (bottom panel).

4.5.4 Liver cT_1 and liver disease aetiology

Liver cT_1 was associated with CPA in patients with steatohepatitis (n=22, r=0.52; p=0.013) and with viral hepatitis (n=18; r=0.80; p<0.001). Linear regression analysis showed that there was a difference between the elevations for the lines fitted for the two disease aetiologies (Figure 24).

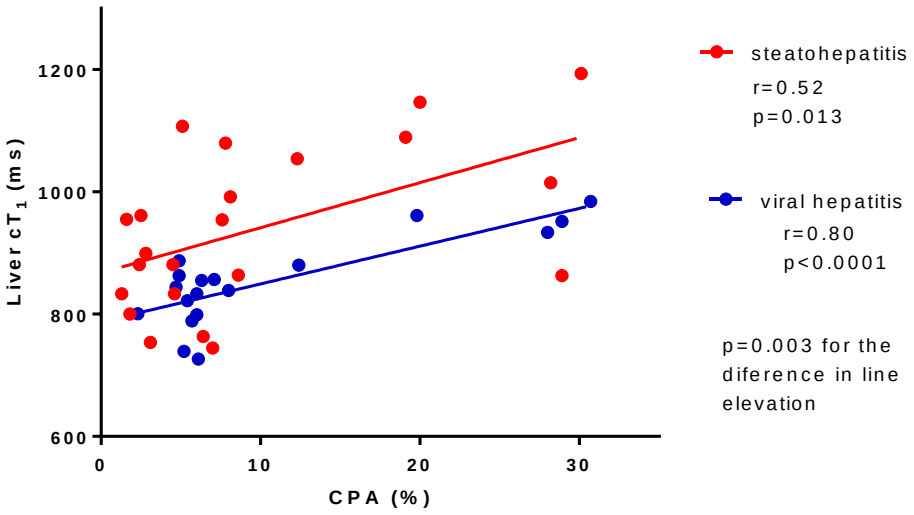


Figure 24. Liver cT_1 vs collagen proportionate area in patients with steatohepatitis and viral hepatitis.

The lines represent lines of best fit from linear regression.

4.5.5 Multivariate models

The prediction of CPA was improved when cT_1 was entered into multiple linear regression models with liver fat content, and T_2* (Table 9). Liver cT_1 remained an independent predictor of CPA in all models. In the model containing all three variables, liver fat had a strong trend for an independent effect (p=0.06), but liver T_2* was not a significant contributor to the model (p=0.17).

Table 9. Multiple linear regression models for prediction of collagen proportionate area.

Variable	B	P	r	R^2
Model 1			0.54	0.29
Constant	-28.2	0.01		
Liver cT_1	0.042	<0.001		
Model 2			0.63	0.40
Constant	-32.6	<0.001		
Liver cT_1	0.05	<0.001		
Liver fat content	-0.39	0.003		
Model 3			0.60	0.36
Constant	-36.0	<0.001		
Liver cT_1	0.43	<0.001		
Liver T_2*	0.39	0.016		
Model 4			0.64	0.41
Constant	-35.1	<0.001		
Livet cT_1	0.05	<0.001		
Liver fat content	-0.31	0.06		
Liver T_2*	0.17	0.40		

4.6 Discussion

This study demonstrated that the DIA technique developed for the morphometric assessment of liver fibrosis (CPA) represents an improvement compared to assessment using categorical scores. Unlike traditional categorical scores, CPA is quantified on a continuous scale, and is robust and reproducible, eliminating any bias from inter- and intra-observer variability. The advantages of CPA were particularly striking in this study as the two pathologists that were not involved in the development of the method, had no prior experience of the technique, and received only one training session prior to reporting the slides. DIA of histological samples for CPA quantification therefore represent a more appropriate reference standard in the development of non-invasive fibrosis biomarkers, compared to categorical scores.

Despite these improvements however, all histological methods of fibrosis quantification are limited by sampling error and for the need for a minimum biopsy size for reliable assessment. This drawback was well demonstrated here. Previous work has shown that analysis of a 28mm core taken with a 16G needle gives a 90% probability that the measured CPA will be within 5% of the true CPA[141]. The biopsies in this study were taken using an 18G needle which means that even longer biopsies would be needed for this level of accuracy. Only 16 (30%) biopsies in this study were at least 28mm long. In the study by Hall *et al*[141], analysis of a 15mm biopsy taken with a 16G needle gave a 75% probability of the measured CPA being within 5% of the true CPA. In this study, the cut-off was set at 12.5mm and 50 (93%) biopsies were longer than 15mm.

The comparison of liver cT_1 with CPA has shown a significant association between the two parameters particularly in patients with viral hepatitis ($r=0.80$; $p<0.0001$; Figure 24). Studies examining the association of liver fibrosis assessed using DIA and liver stiffness

measured using transient elastography[142], or magnetic resonance elastography[42] in patients with chronic viral hepatitis have shown similar results (r=0.77 and 0.78 for patient with hepatitis C and B respectively)[42,142].

The weaker correlation between CPA and liver cT_1, in patients with steatohepatitis (r=0.52; p<0.013; Figure 24) is most likely due to the confounding effect of liver fat, which also has a significant association with liver cT_1 (r=0.42, p=0.002; Table 8). Another possible contributor to the effect of aetiology is the fact that patients with steatohepatitis can accumulate substantial amounts of pericellular fibrosis, while in patients with viral hepatitis; fibrosis accumulates primarily in portal tracts. CPA techniques have been primarily developed and tested in patients with viral hepatitis[115,142,143] and it is not well established whether they are sensitive enough for the precise quantification of pericellular fibrosis. It is therefore possible that in patients with steatohepatitis, CPA represents a relative under-estimate of the true amount of fibrosis, hence the findings in this study.

It was encouraging to see that, in contrast to other studies[67,72], liver iron did not have a confounding effect on cT_1, something that provides further validation for the iron correction algorithm that is accounted for in the liver cT_1 score. The considerable confounding effect from liver fat is more difficult to explain. Increasing liver steatosis would be expected to result in shortening of the liver T_1 relaxation time, as fat has a shorter T_1 compared to water. However, the opposite effect is observed here, with a positive association between fat and liver cT_1. Further studies would be needed in order to elucidate this finding.

The multiple linear regression analysis highlights one other major strength of multi-parametric MR assessment as the incorporation of liver cT_1 (obtained from T_1 and T_2* maps) and liver fat content measured by ^{1}H-MRS results in improved models for the prediction of CPA (Table 9). The multivariate analysis also shows that cT_1 is the main predictor of CPA as it is remains independent in all the models.

In conclusion, work in this chapter has validated a technique for the assessment of liver collagen proportionate area on histological slides and demonstrated that liver cT_1 is closely associated with CPA particularly in patients with viral hepatitis. Furthermore, for the first time, liver fat is shown to exert a considerable confounding effect on cT_1 measurements but these can be largely accounted for in multivariate predictive models.

5. Assessment of non-alcoholic fatty liver disease

5.1 Abstract

Background and aims: The diagnosis of non-alcoholic steatohepatitis (NASH) and staging of fibrosis are central to the assessment non-alcoholic fatty liver disease (NAFLD). The aim of this study was to evaluate the utility of multi-parametric magnetic resonance (MR) in the assessment of the histological components of NASH including fibrosis.

Methods: 78 patients with suspected NAFLD were recruited within one month of liver biopsy. MR data were used to define the liver inflammation and fibrosis score (LIF 0-4). Biopsies were assessed for steatosis, lobulitis, ballooning and fibrosis and classified as NASH or simple steatosis and mild or significant NAFLD. Transient elastography (TE) was also was performed.

Results: The MR success rate was 95% vs 59% for TE ($p<0.0001$). Fibrosis stage on biopsy correlated with LIF ($r_s=0.51$, $p<0.0001$). The area under the receiver operating curve (AUROC) for the diagnosis of cirrhosis was 0.85 for LIF. The LIF score for ballooning grades 0, 1 and 2 was 1.2, 2.7 and 3.5 respectively ($p<0.05$) with an AUROC of 0.83 for the diagnosis of any grade of ballooning. Patients with simple steatosis had lower LIF (1.3) compared to patients with NASH (3.0) ($p<0.0001$); the AUROC for the diagnosis of NASH was 0.80. LIF scores for patients with mild and significant NAFLD were 1.2 and 2.9 respectively ($p<0.0001$). The AUROC of LIF for the diagnosis of significant NAFLD was 0.89.

Conclusions: Multi-parametric MR is a promising technique with good diagnostic accuracy for NAFLD histological parameters, and can potentially identify patients with NASH and advanced fibrosis.

5.2 Introduction

Non-alcoholic fatty liver disease (NAFLD) represents a disease spectrum ranging from accumulation of liver fat only (steatosis), to fat and liver inflammation (non-alcoholic steatohepatitis; NASH) which can be associated with fibrosis and cirrhosis. NAFLD is closely associated with type 2 diabetes mellitus and the metabolic syndrome [144,145], and has now reached epidemic levels in developed countries, affecting a third of the adult population[1-3]. NASH prevalence is estimated at 3-12% in the general population [146,147], and is expected to become the most common indication for liver transplantation in the US in the near future [148]. Steatosis and NASH are traditionally thought of as distinct disease entities with steatosis running a generally benign course while NASH incurs an increased risk of disease progression and cardiovascular, liver and cancer morbidity and mortality [149-152]. However, recent studies[153,154] have shown that some patients with simple steatosis can develop progressive disease, suggesting that NAFLD may be more complex than previously thought.

The diagnosis and classification of NAFLD into different sub-types (steatosis, NASH, cirrhosis) currently relies on liver biopsy. However, this is problematic due to the inherent drawbacks of this technique (sampling and observer dependent variability, risks of complications, patient discomfort and costs) [155]. Furthermore, the majority of patients with NAFLD have uncomplicated steatosis, and the reliance on biopsy means that patients need to undergo an invasive procedure to diagnose a disease that is most likely benign. There is therefore an unmet clinical need for reliable non-invasive biomarkers for the assessment of NAFLD.

The two most important aspects of NAFLD assessment are the identification of NASH and the staging of fibrosis, and the diagnostic accuracy of various non-invasive biomarkers for these features has been assessed.

Serum biomarkers have yielded mixed results that have hindered widespread clinical application. For example, cytokeratin 18 (CK18) has demonstrated moderate overall accuracy for diagnosing NASH in a meta-analysis (66% sensitivity, 82% specificity)[156], but was found to have only a limited sensitivity (58%) for the diagnosis of NASH in a large clinical study[157].

Measurement of liver stiffness (LS) using transient elastography (TE) is associated with high failure rates, particularly in obese patients (BMI>30kg/m^2) [158], where reliable measures could only be obtained in 65% of patients in one study [159]. This limits the applicability of TE for the assessment of patients with NAFLD who are often obese.

More recently measuring liver stiffness using magnetic resonance elastography (MRE) has shown promise in the evaluation of fibrosis in patients with NAFLD [45]. However, the diagnostic accuracy of MRE for the diagnosis of NASH was limited with an overall area under the receiver operating characteristic curve of 0.73. Furthermore, this technique still remains restricted to specialist centres and with considerable obstacles to widespread use, including the need for additional hardware.

Multi-parametric MR to assess liver fibrosis and inflammation has not been previously assessed in the assessment of NAFLD. As this technique showed good diagnostic

accuracy for the assessment of fibrosis in a cohort of patients with mixed liver disease aetiologies[114] it was hypothesised that it would also have utility in the assessment of patients with NAFLD.

5.3 Aims

The primary aim of this study was to evaluate the diagnostic performance of multi-parametric liver MRI specifically in the assessment of patients with non-alcoholic fatty liver disease (NAFLD) using liver histology as the reference standard. MR was also compared to TE in the assessment of fibrosis. The analysis was conducted using components of the steatosis, activity and fibrosis (SAF) score and the diagnostic categories of the fatty liver inhibition of progression (FLIP) consortium algorithm[160].

5.4 Patients and methods

5.4.1 Study design and patient population

This was a prospective study conducted at a UK tertiary centre (John Radcliffe Hospital, Oxford, UK) between May 2011 and March 2015. Adult patients (≥18 years) with suspected NAFLD scheduled to have a liver biopsy in the following month or patients with known NAFLD who had a liver biopsy within the previous month were invited to participate. Patients attended for a single visit, for multi-parametric MR examination, TE and blood sampling. Patients were recruited from general hepatology and metabolic liver disease clinics and from the bariatric surgery service. Biopsies were performed as part of normal clinical practice for the diagnosis of NAFLD/NASH and staging of fibrosis. Figure 25 summarises the recruitment details for the study.

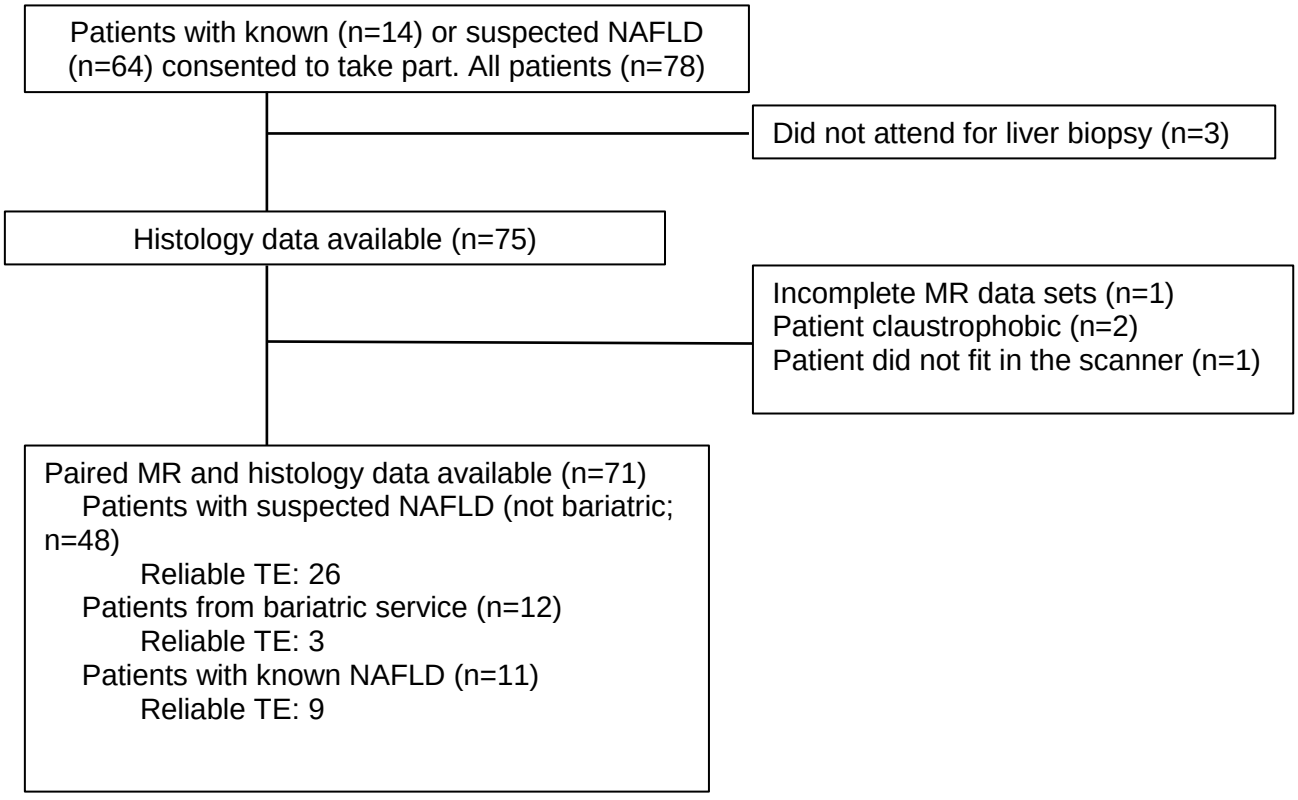


Figure 25. Study flow diagram.

Abbreviations: **NAFLD:** non-alcoholic fatty liver disease, **MR**: magnetic resonance, **TE:** transient elastography

Patients were ineligible if they had: (a) contraindications to MRI scanning, including pregnancy, implanted cardiac devices and cerebral artery aneurysm clips (b) consumed more alcohol that the current limit recommended by the UK Department of Health (2-3 units (16-24g) / day for women and 3-4 units (24-32g) / day for men), (c) had clinical or laboratory evidence of a liver diagnosis other than NAFLD including chronic viral hepatitis (positive surface antigen for hepatitis B, or positive hepatitis C antibody), cholestatic liver disease, Wilson's disease, haemochromatosis or alpha-1-antitrypsin deficiency.

5.4.2 Multi-parametric MR examination

All MR scans were performed as described in chapter 2. The liver inflammation and fibrosis (LIF) score was the primary variable of interest in this study.

5.4.3 Transient elastography

TE was performed using Fibroscan (Echosens, France) with the patient lying supine and with the right arm fully extended. The examination was attempted using the medium (M) probe and if this was not successful, measurements were repeated using the extra-large (XL) probe. Ten measurements per patient were needed for a successful scan and the manufacturer's recommendations were used to assess the validity of each examination (10 measurements in each patient, with a rate of successful measurements greater that 60% and where the interquartile range was not greater than 0.3 of the median).

5.4.4 Liver histology

Percutaneous biopsies (n=50) were performed under ultrasound guidance using 18G cutting biopsy needles (TruCut) and trans-jugular (n=9) biopsies were performed under fluoroscopic guidance using 19G needles. Patients who were having bariatric surgery (n=12) had wedge liver biopsies intra-operatively. The median (IQR) biopsy length in

patients who had needle biopsies was 18mm (14-25), including a median (IQR) of 10 (7 – 13) portal tracts. All biopsies were included in the final analysis.

Biopsies were evaluated by two experienced liver pathologists and discussed in a clinico-pathological meeting before a final consensus report was issued, and this was used as the reference standard in this study. The reporting pathologists and clinician attending the clinico-pathological meeting were blinded to the MR data.

Biopsies were assessed for steatosis, ballooning, lobulitis and fibrosis. Steatosis was assessed on a 4 tier scale (0-3), ballooning and lobulitis on a 3 tier scale (0-2) and fibrosis on a 5 tier scale (0-4).The four histological components were summed into the steatosis, activity (ballooning + lobulitis) and fibrosis (SAF) score. Patients were categorised into steatosis and NASH and for the overall disease severity into mild and significant disease according to the algorithms suggested by the fatty liver inhibition of progression (FLIP) consortium [160]. The steatosis vs NASH algorithm requires at least grade 1 steatosis (>5% of hepatocytes containing lipid inclusions) and both lobulitis and ballooning to be present for a diagnosis of NASH whilst the stage of fibrosis does not feature in this part of the algorithm. For the overall severity assessment, disease is classed as significant if either activity (ballooning + lobulitis) or fibrosis scores are ≥2 and is otherwise classified as mild [160].

5.4.5 Statistical analysis

All the analysis was carried out using GraphPad Prism software (version 6.05, July 7, 2014). Statistical significance was set at $p<0.05$. Descriptive statistics were used to summarise baseline subject characteristics. Normality was determined using the Shapiro-Wilks test. Associations were tested using the Spearman's correlation coefficient (r_s).

Differences between groups were assessed using the Mann Whitney test. Fisher's exact test was used to test for the differences in proportions between two groups. Differences between multiple groups were assessed using the Kruskal Wallis test with Dunn's correction for multiple comparisons.

Receiver operating characteristic curves (ROC) were used to determine (a) the diagnostic accuracy of multi-parametric MR for the assessment of NAFLD components (ballooning, lobulitis, activity, fibrosis, NASH vs steatosis, mild vs significant NAFLD, and (b) the diagnostic accuracy of TE in the assessment of NAFLD fibrosis. A cut-off to optimise sensitivity at 90% was reported. Ninety five percent confidence intervals (95% CI) were calculated for the parameters of diagnostic accuracy.

5.5 Results

5.5.1 Baseline characteristics

A total of 78 patients consented to participate and biopsy data were available in 75. MR examination was successful in 71 (95%) patients (Figure 25). TE was attempted in 64 (90%) patients, and valid measurements were obtained in 38 (59%) patients. The success of TE was significantly lower than multi-parametric MR ($p<0.0001$). The mean (±SD) age was 53.4 years (±11.6) and patients had a median (IQR) body mass index (BMI) of 32.7kg/m^2 (28.1 – 38.1). The majority of the patients were male (n=43; 60%) and 25 (35%) had type 2 diabetes mellitus (T2DM).

Categorised using the FLIP algorithm [160], 25 (35%) patients had steatosis and 46 (65%) had NASH. For overall disease severity, 13 (18%) patients were classed as having mild disease and 58 (82%) as having significant disease. The number of patients with fibrosis stages 0, 1, 2, 3, and 4, were 5 (7%), 20 (28%), 20 (28%), 15 (21%) and 11 (15%), respectively. The subject characteristics of the whole cohort (demographics, liver function tests, metabolic profile, histology and MR data) are presented in Table 10 and Table 11. Representative MR images from patients with varying NAFLD disease severity are shown in Figure 26.

Table 10. Patient demographic laboratory, and magnetic resonance data

	All patients (n=71)
Demographics	
Age (years; mean ± SD)	53.4 (±11.6)
Male (n, %)	43 (60)
BMI (kg/m^2; median; IQR)	32.7 (28.1 – 38.1)
Type 2 Diabetes Mellitus (n, %)	25 (35)
Liver function tests; median (IQR)	
Bilirubin (µmol/l)	10 (7 – 16)
ALT (IU/l)	54 (30 – 76)
ALP (IU/l)	172 (138 – 233)
Albumin (g/l)	45 (44 – 47)
GGT (IU/l)	66 (36 – 118)
AST (IU/l)	40 (31 – 56)
Haematological tests; median (IQR)	
Platelet count (x10^9/l)	208 (166 – 278)
Prothrombin time (s)	13.6 (13.1 – 14.5)
Metabolic profile; median (IQR)	
Glucose (mmol/l)	5.1 (4.8 – 6.3)
Cholesterol (mmol/l)	4.7 (3.8 – 5.6)
HDL (mmol/l)	1.1 (1.0 – 1.2)
LDL (mmol/l)	2.7 (2.0 – 3.5)
Triglycerides (mmol/l)	1.6 (1.2 – 2.4)
Non-invasive scores; median (IQR)	
cT_1 (ms)	923 (832 – 1035)
LIF score	2.6 (1.5 – 3.3)
Liver stiffness (kPa; n=38)	7.5 (5.1 -13.3)

Abbreviations: **MR:** magnetic resonance, **cT_1**: iron corrected T_1, **LIF**: liver inflammation and fibrosis

Table 11. Patient histology profile

Histology (n, %)	All patients (n=71)
Fibrosis	
0	5 (7)
1	20 (28)
2	20 (28)
3	15 (21)
4	11 (15)
Ballooning	
0	17 (24)
1	46 (65)
2	8 (11)
Lobulitis	
0	12 (17)
1	58 (82)
2	1 (1)
Steatosis	
0	4 (6)*
1	8 (11)
2	17 (24)
3	42 (59)
FLIP algorithm classification (n; %)	
Steatosis	25 (35)
NASH	46 (65)
Mild disease	13 (18)
Significant disease	58 (82)

*Four patients with suspected NAFLD were found to have no steatosis on liver biopsy but still included as the absence of NAFLD could not have been predicted without liver biopsy.

Abbreviations: **FLIP**: fatty liver inhibition of progression consortium

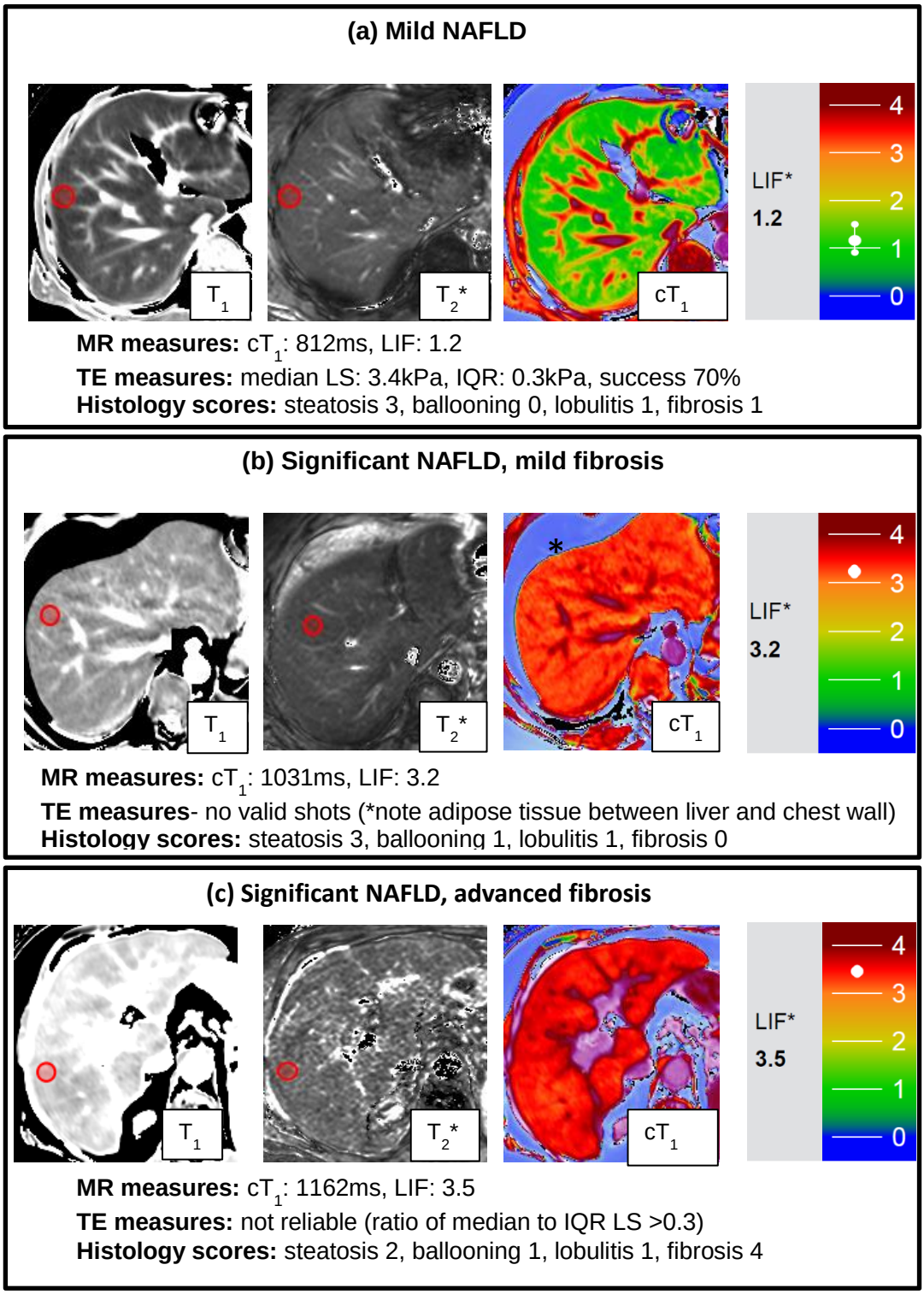


Figure 26. Representative MR Data

Magnetic resonance data (T_1, T_2*, cT_1 maps and LIF scores) from patients classified using the Fatty Liver Inhibition of Progression (FLIP) algorithm as having mild disease (top panel), significant disease with mild fibrosis (middle panel) and significant disease with advanced fibrosis (bottom panel).

5.5.2 Assessment of fibrosis

5.5.2(i) Multi-parametric magnetic resonance

There was a significant association between histological fibrosis and LIF (r_s= 0.51, p<0.0001; Figure 27). The median LIF for patients without cirrhosis (F0-3) was 2.2 and for patients with cirrhosis (F4) 3.3 (p<0.0001). The AUROC for the diagnosis of cirrhosis was 0.85 (95% CI: 0.76 – 0.95; p=0.0002). A LIF cut-off of 3.0 had a sensitivity 91% (95% CI: 59% - 100%) and specificity 73% (55% - 80%) for the diagnosis of cirrhosis; Table 12 and Table 13).

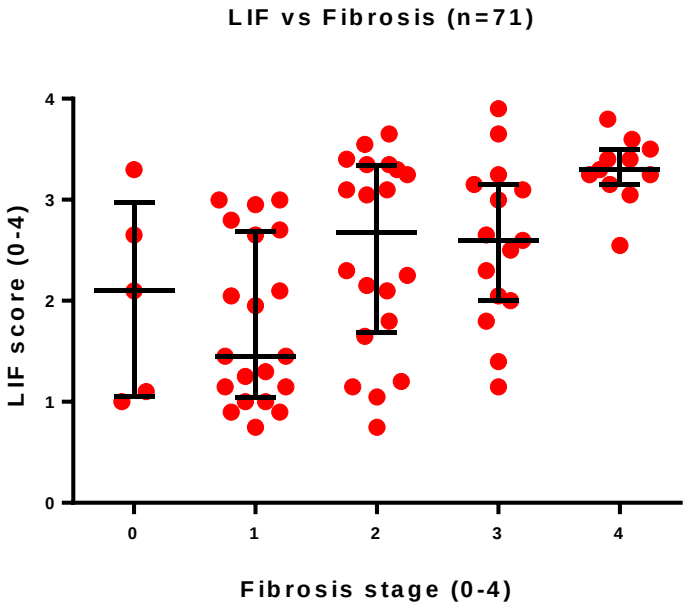


Figure 27. Liver Inflammation and fibrosis score plotted against the histological fibrosis stage.

Lines and error bars indicate the median and inter-quartile range

5.5.2(ii) Transient elastography

There was a significant association between LS and histological fibrosis stage (r_s=0.56; p=0.0003, Figure 28). Patients with cirrhosis had higher median LS compared to patients without cirrhosis (27.0kPa vs 7.0kPa; p=0.002) and the AUROC of TE for the diagnosis of cirrhosis was 0.93 (95% CI: 0.85 – 1.00; p=0.005). A LS cut-off of 14.7kPa had a sensitivity of 100% (95% CI: 40% - 100%) and specificity 91% (95%CI: 76% - 98%).

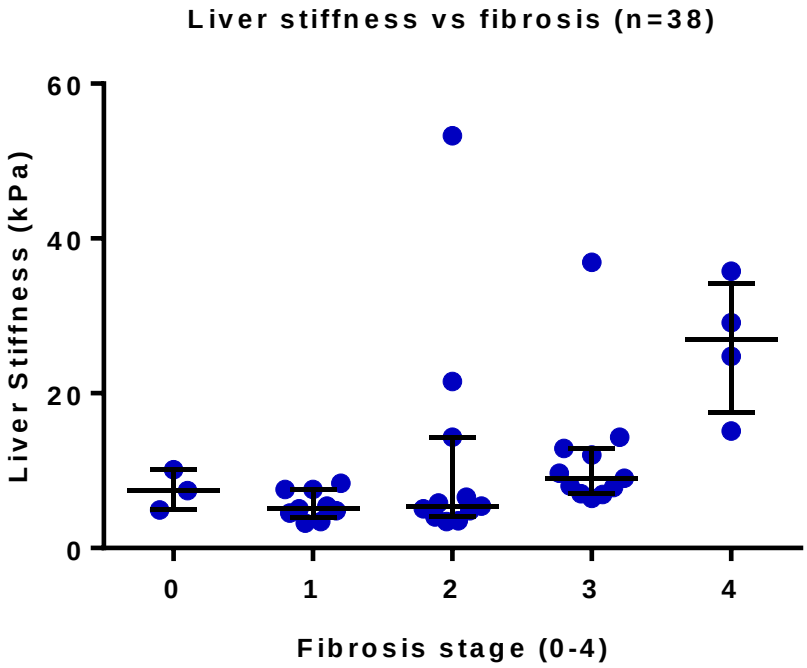


Figure 28: Liver stiffness measured using transient elastography, plotted the histological fibrosis stage.

Lines and error bars indicate the median and inter-quartile range

The diagnostic accuracy for optimal cut-offs identified in this study for the diagnosis of significant (F2-4) and bridging fibrosis (F3-4) and cirrhosis (F4) are summarised in Table 12.

Table 12. Diagnostic accuracy of optimal cut-offs identified in this study for multi-parametric magnetic resonance imaging and transient elastography for the assessment of fibrosis in patients with NAFLD

	LiverMultiScan (n=71)				Transient elastography(n=38)			
	AUROC (95% CI)	Cut-off (0.0-4.0)	Se. (%)	Sp. (%)	AUROC (95% CI)	Cut-off (kPa)	Se. (%)	Sp. (%)
F≥2	0.78 (0.67 – 0.89)	1.2	91	36	0.73 (0.58 – 0.90)	3.8	92	17
F≥3	0.73 (0.61 – 0.85)	1.7	92	40	0.83 (0.7 – 0.96)	6.8	93	65
F4	0.85 (0.76 – 0.95)	3.0	91	73	0.93 (0.85 – 1.00)	14.7	100	91

Abbreviations**. AUROC**: area under the receiver operating characteristic curve, **95% CI**: 95% confidence interval, **Se**: sensitivity**, Sp**: specificity

5.5.3 Assessment of disease activity; ballooning and lobulitis

There was a significant association between histological ballooning grade and the LIF score (r_s=0.59; p<0.0001). The median LIF scores for patients with ballooning grades 0, 1, and 2 were 1.2, 2.7 and 3.5, respectively, with significant differences between all the groups. The area under the receiver operating characteristic curve (AUROC) for the diagnosis of no ballooning vs ballooning grades 1 and 2 was 0.83 (95% CI: 0.72 – 0.93; p<0.0001). A LIF cut-off of 1.2 had a sensitivity of 91% (95% CI: 82% – 98%) and specificity of 53% (95% CI: 28% - 77%) for the diagnosis of ballooning grade >0 (Table 13).

The median LIF scores of patients with no lobulitis and lobulitis grade >0 were 1.5 and 2.7 respectively (p=0.024, Figure 29). There was an association between LIF and overall activity (sum of ballooning + lobulitis grades; r_s 0.58; p<0.0001). The median LIF scores for patients with activity scores 0-1 and 2-4 were 1.3 and 3.0 respectively (p<0.0001; Figure 29). The AUROC for the diagnosis of patients with activity ≥2 was 0.83 (95% CI: 0.73 to 0.93; p<0.0001). A LIF cut-off of 1.6 had a sensitivity of 90% (95% CI: 77% - 97%) and specificity of 61% (95% CI: 39% - 80%) for the diagnosis of an overall activity grade ≥2 (Table 13).

Overall there was a strong association between the total SAF score and LIF score (r_s= 0.70, p<0.0001; Figure 29).

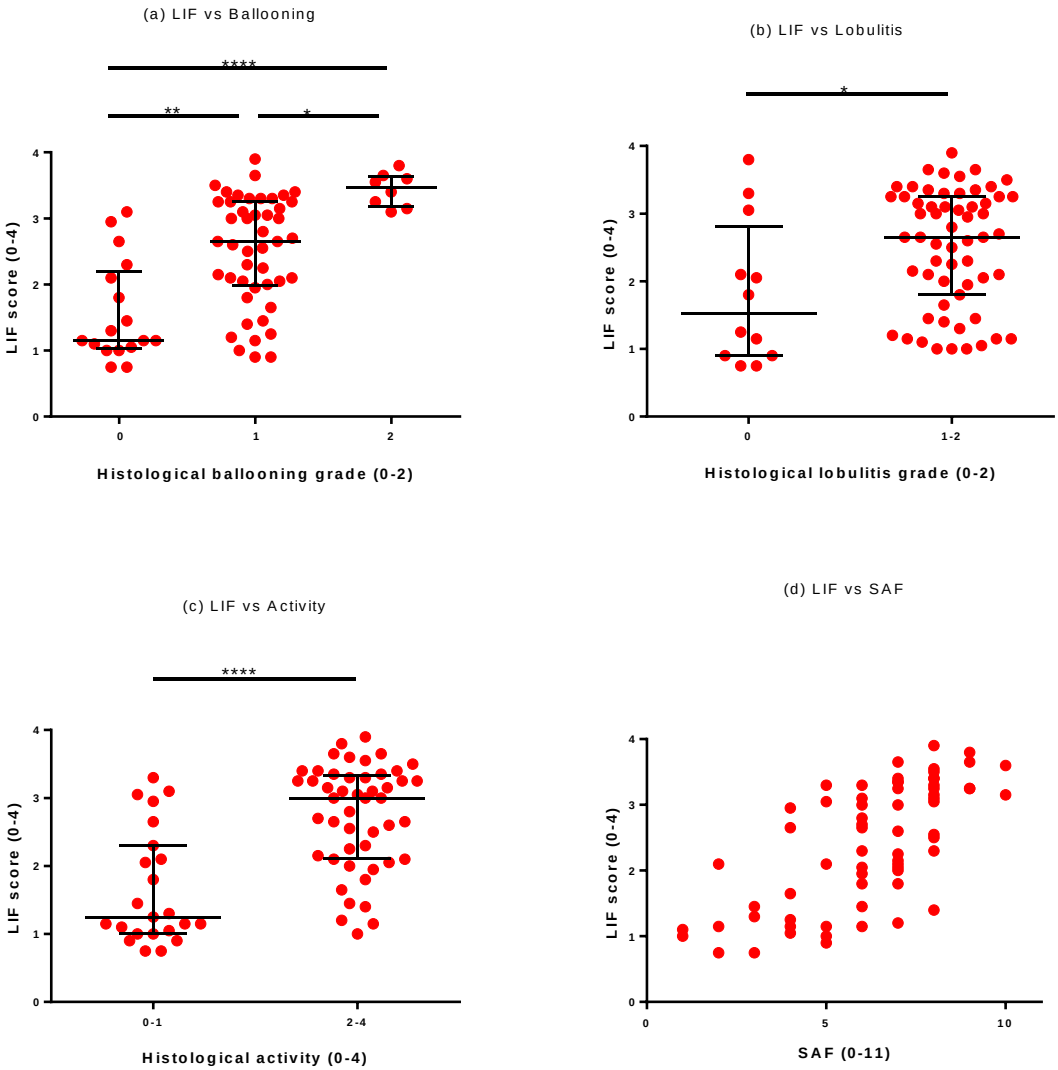


Figure 29. Liver inflammation and fibrosis score for the evaluation of (a) ballooning, (b) lobulitis (c) overall activity and (d) the total steatosis, activity and fibrosis score.

(a)The median liver inflammation and fibrosis scores (LIF) for patients with no ballooning, and ballooning grades 1 and 2 were respectively 1.2, 2.7 and 3.5 ($p<0.05$)

(b) Patients with no lobulitis had a median LIF of 1.6 compared to 2.7 in patients with lobulitis grade >0 ($p=0.02$)

(c) Patients with mild activity (activity grade 0-1) had a lower median LIF than patients with significant activity (2-4) (1.3 vs 3.0; $p<0.0001$).

(d) There was a strong association between the liver inflammation and fibrosis score (LIF) and the overall histological severity scored by the steatosis, activity and fibrosis score (SAF; $r_s=0.70$; $p<0.0001$)

Lines and error bars indicate the median and interquartile range on all graphs.

5.5.4 Steatosis vs NASH

Patients categorised using the FLIP algorithm into steatosis and NASH had a median LIF of 1.3 and 3.0 respectively ($p<0.0001$, Figure 30). The AUROC for the diagnosis of steatosis vs NASH was 0.80 (95% CI: 0.69 – 0.92; $p<0.0001$). A LIF cut-off of 1.4 had a sensitivity of 91% (95% CI: 79 - 98%) and specificity 52% (95% CI: 31% - 72%; Table 13).

5.5.5 Overall disease severity (mild vs significant)

Stratified according to the FLIP consortium algorithm, patients with mild and significant disease had a median LIF 1.2 and 2.9 respectively ($p<0.0001$; Figure 30). The AUROC for the diagnosis of mild vs significant NAFLD was 0.89 (95% CI: 0.80% - 0.98%; $p<0.0001$) and a LIF cut-off of 1.4 had a sensitivity of 90% (95% CI: 79% - 96%) and specificity of 77% (95% CI: 46% - 95%; Table 13).

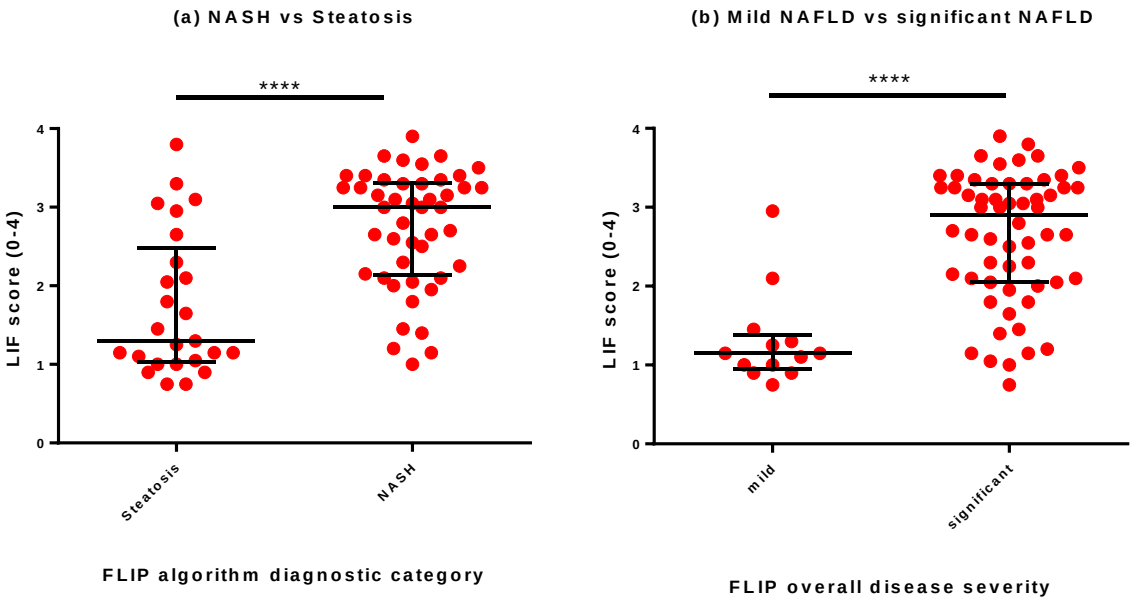


Figure 30. Liver inflammation and fibrosis score for NAFLD classification according to the Fatty Liver Inhibition of Progression consortium algorithms

Liver inflammation and fibrosis scores (LIF) are plotted against the diagnostic categories of the fatty liver inhibition of progression (FLIP) algorithm.

(a) Patients with steatosis had lower median LIF scores compared to patients NASH (1.3 vs 3.0; $p<0.0001$) and (b) patients with mild disease had lower median LIF scores compared to patients with severe disease (2.2 vs 3.3; $p<0.0001$). Lines and error bars indicate the median and interquartile range on both graphs.

Table 13. Diagnostic parameters of liver inflammation and fibrosis score for NAFLD assessment

	AUROC (95% CI)	p value	LIF cut-off	Sensitivity (%; 95% CI)	Specificity (%; 95% CI)
SAF score components*					
Ballooning grade 0 vs >0	0.83 (0.72 – 0.93)	<0.0001	1.2	91 (82 - 98)	53 (28 – 77)
Activity grade 0-1 vs >1	0.83 (0.73 – 0.93)	<0.0001	1.6	90 (77 – 97)	61 (39 – 80)
Fibrosis stage 0-3 vs 4	0.85 (0.76 – 0.95)	0.0002	3.0	91 (59 – 100)	68 (55 – 80)
FLIP algorithm categories*					
Steatosis vs NASH	0.80 (0.69 – 0.92)	<0.0001	1.4	91 (79 – 98)	52 (31 – 72)
Mild vs significant disease	0.89 (0.80 – 0.98)	<0.0001	1.4	90 (79 – 96)	77 (46 – 95)

*In the steatosis, activity and fibrosis (SAF) score, biopsies are reported for steatosis (0-3), Activity (0-4; sum of ballooning (0-2) and lobulitis (0-2)) and fibrosis (0-4). The fatty liver inhibition of progression algorithms categorise patients as NASH if all of steatosis, ballooning and lobulitis are graded as 1 or higher and as steatosis if this criterion is not met. The overall disease severity is also classified as mild (fibrosis <2 and activity <2) or significant (fibrosis or activity ≥2).

Abbreviations: **AUROC**: area under the receiver operating characteristic curve, **LIF:** liver inflammation and fibrosis score

5.6 Discussion

This prospective study has demonstrated that the newly established multi-parametric MRI technique can be used to assess the overall disease severity in patients with NAFLD, with an AUROC of 0.89 for the detection of significant disease as defined by the FLIP consortium algorithm. Furthermore, this technique had high diagnostic accuracy for the diagnosis of NASH (AUROC of 0.80) and hepatocyte ballooning (AUROC 0.83), a lesion that is considered central to the diagnosis of NASH[108]. The diagnostic accuracy for the diagnosis of cirrhosis was also good with an AUROC of 0.85. Overall, the results suggest that multi-parametric MR is sensitive and accurate in quantifying both the inflammatory (NASH / ballooning) and fibrotic components of NAFLD with an excellent diagnostic accuracy for the assessment of overall severity.

Several aspects of this technique would make it attractive as a surrogate end point in NAFLD clinical trials. The high diagnostic accuracy for both the inflammatory and fibrotic components of NAFLD, combined with the low reporting variability, could mean that small early changes in response to therapy may be more readily detected. This would be particularly important in early trials evaluating anti-inflammatory and anti-fibrotic therapies for NAFLD. Recruitment and retention to clinical trials in NAFLD, which rely on liver biopsy to assess endpoints, have been problematic with a recent study reporting a 25% drop-out rate[161]. Implementation of robust non-invasive surrogate end points may therefore lead to significant improvements in patient retention and the rapid evaluation of novel therapeutics. MR could also be used in clinical practice for the risk stratification of patients. Potentially, a LIF cut-off of 1.4 could be used to identify patients with significant NAFLD/NASH who may need follow up in specialist clinics or prioritisation for drug treatments or intensive lifestyle interventions. As multi-parametric MRI is completely non-invasive, it is ideally suited for monitoring disease progression and regression after therapeutic intervention.

Multi-parametric MR had a significantly higher success rate (95%) compared to TE (59%, p<0.0001), whilst there was considerable overlap in the 95% confidence intervals for the diagnostic accuracy in the detection of fibrosis, indicating no significant differences between the two techniques (Table 12). TE is solely used for the assessment of fibrosis, so its diagnostic accuracy in the assessment of activity, NASH and overall disease severity was not examined in this study. An advantage of MR technology is that this can detect the early hallmark features of NASH (inflammation and ballooning) before fibrosis is established.

The prospective design of the study, where the majority of the MR data were acquired and analysed before liver biopsy is a major strength of this study. The use of histology as the reference standard is always problematic in studies of diagnostic accuracy of non-invasive tests due to sampling and observer dependent variability of biopsy. Good inter-observer agreement was previously found between the pathologists taking part in this study, for the assessment of fibrosis (weighted kappa 0.51) and steatosis (weighted kappa 0.72)[114]. In addition, for the histological assessment of fibrosis, a minimum biopsy length of 25mm or at least 11 portal tracts are needed for reliable scoring[162]. However a biopsy with at least 6 portal tracts is generally considered adequate for routine diagnosis[163]. As the aim of the study was to conduct a comparison with all the histological aspects of NAFLD, biopsies were not excluded based on quality criteria if the pathologists could assess all the histological parameters of interest.

In conclusion, this study shows that multi-parametric MR is a promising tool for the evaluation of patients with NAFLD. Applicability of MR was higher than TE with no differences in the diagnostic accuracy for significant fibrosis or cirrhosis. Furthermore, multi-parametric MR provided additional information beyond fibrosis assessment with

good accuracy for the diagnosis of NASH and ballooning. Therefore, the ability to assess both the necro-inflammatory and fibrotic components of NASH in a single test is a particular strength of the MR technique that allows accurate evaluation of the overall disease severity. Further refinement and the technical development of non-invasive biomarkers that enable the quantification of the inflammatory and fibrotic components of NAFLD is likely to revolutionise this field.

6. Assessment of portal hypertension

6.1 Abstract

Background and Aims: Non-invasive assessment of portal hypertension is an area of unmet clinical need. Transient elastography which is now recommended as a routine screening test for the presence of oesophageal varices suffers with poor applicability and therefore the search for the ideal non-invasive biomarker of portal pressure still continues. The aim of this study was to evaluate liver and spleen cT_1 in the assessment of portal hypertension.

Methods: In this study, LMS was used to measure liver and spleen cT_1 in patients who had transjugular liver biopsy and hepatic vein pressure gradient (HVPG) measurement. HVPG was the reference standard.

Results: Nineteen patients (15 men; 79%) with a median age of 57 years were included. Liver disease aetiologies included non-alcoholic fatty liver disease (n=9; 47%) and viral hepatitis (n=4; 21%). There was a strong correlation between spleen cT_1 and HVPG ($r=0.66$, $R^2=0.43$, $p=0.002$) and a significant correlation between liver cT_1 and HVPG ($r=0.49$; $R^2= 0.24$; $p=0.038$).

Spleen cT_1 had excellent diagnostic accuracy for the diagnosis of any degree of portal hypertension (HVPG>5mmHg; AUROC 0.91), clinically significant portal hypertension (HVPG≥10mmHg; AUROC 0.91) and severe portal hypertension (HVPG≥12mmHg; AUROC 0.81). Liver cT_1 performed less well showing diagnostic accuracy only for clinically significant portal hypertension (AUROC 0.81). Furthermore, liver fat content was shown to have a confounding effect on liver cT_1.

Conclusion: Overall, this proof-of-principle study shows that spleen cT_1 is a promising biomarker of portal pressure that should be explored further in larger studies.

6.2 Introduction

Venous blood from the intestine and from the spleen is delivered to the liver through the portal vein. Once in the liver, blood is filtered through the hepatic sinusoids and returns to the systemic circulation through the hepatic veins and the inferior vena cava (IVC; Figure 31).

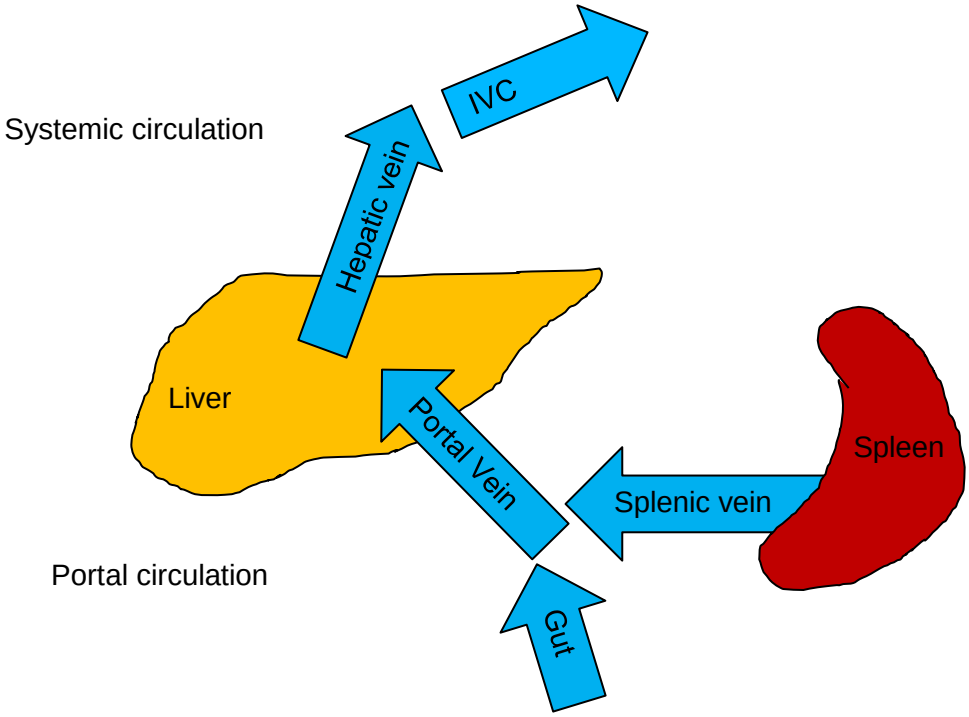


Figure 31. Diagrammatic illustration of the relationship between the liver and the portal circulation

In patients with liver disease, accumulation of liver fibrosis disrupts the normal flow of blood through the hepatic sinusoids and leads to an increase in the vascular pressure in the portal circulation (portal hypertension). Worsening of fibrosis and portal hypertension lead to the development of collateral vessels (varices) that bypass the liver and connect the portal and systemic circulations. This pathophysiological process underpins major complications of liver cirrhosis (ascites, hepatic encephalopathy and variceal bleeding)[164].

Assessment of the portal pressure is therefore an important aspect in the care of patients with liver disease as it gives an overall evaluation of the severity of disease, and allows

risk stratification of patients with cirrhosis[165-167]. Measurement of the direct portal pressure is practically very difficult as there is no direct communication between the portal and systemic circulation, although the latter is readily accessible. In routine practice, measurement of the hepatic vein pressure gradient (HVPG) give close estimates of the true portal pressure[168]. HVPG measurement is the preferred method of portal pressure assessment as it indicates the pressure gradient between the portal and systemic circulations, which is more likely to be predictive of portal hypertension complications[169]. However, this is an invasive and costly procedure that requires high levels of expertise, and is therefore not widely available[170].

Currently, the diagnosis of portal hypertension relies on the finding of ascites on clinical evaluation or on the diagnosis of oesophageal varices on endoscopy. Non- invasive assessment of portal hypertension is therefore an area of clinical need.

Measurement of liver and spleen stiffness using transient and magnetic resonance elastography have shown good association with the HVPG[50,171]. Assessment with transient elastography is now recommended by the Baveno consensus criteria for screening for the presence of oesophageal varices in patients with cirrhosis[21]. However, a recent study has found a 29% false negative rate when applying the consensus criteria[172], and the authors of this report point to many cautions in the application of transient elastography, including observer dependent variability[172-174]. There is therefore scope for improvement to the current available tests and to explore alternative approaches.

6.3 Aims

The aim of this study was to explore how multi-parametric MR could be used in the assessment of portal hypertension. Evaluation of liver cT_1 in the assessment of portal pressure would represent the natural progression in the validation of this biomarker, following on from comparison against quantitative (chapter 4) and categorical staging of

fibrosis[114]. Furthermore, the acquisition of transverse abdominal cT_1 maps through the liver mid-equatorial plane also produces quantitative images of the spleen. Therefore, the utility of spleen iron corrected T_1 was also examined, as it was hypothesised that portal hypertension would lead to splenic congestion and therefore an increase in spleen cT_1.

6.4 Patients and methods

6.4.1 Study design and patient population

The population under study were adult patients with suspected liver cirrhosis who were scheduled to have a clinically indicated trans-jugular liver biopsy and HVPG assessment. Multi-parametric MR was performed prior to the trans-jugular biopsy and HVPG measurement. The biopsy and HVPG measurement were performed by interventional radiologists who were blinded to the MR data. The median (IQR) interval between MR and HVPG measurement was 5 days (1 -15). Blood was collected at the time of the MR study for liver function tests and full blood count. Simple serum based (AST/ALT ratio[175], APRI[176], FIB-4[177]) indices that have been used previously used to assess liver fibrosis were calculated.

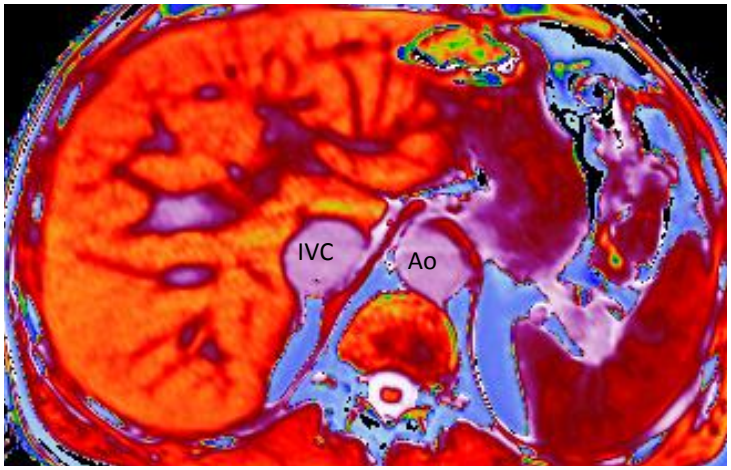
HVPG was designated as the reference standard against which the liver and spleen cT_1 were compared. The diagnostic accuracy of liver and spleen cT_1 was tested for the diagnosis of (a) any degree of portal hypertension (HVPG>5mmHg), (b) clinically significant portal hypertension (CSPH; HVPG≥10mmHg) and (c) severe portal hypertension (SPH; HVPG≥12mmHg).

Non-cirrhotic patients with no portal hypertension were included in this analysis. Patients with non-cirrhotic portal hypertension were excluded from the analysis of HVPG against liver cT_1, but not from the analysis of HVPG against spleen cT_1. Patients in whom the HVPG could not be computed were excluded.

6.4.2 Multi-parametric MR

Multi-parametric MR data were acquired as described in chapter 2. For this study data were analysed using LiverMultiScan (LMS; Perspectum Diagnostics, Oxford UK) for liver

and spleen cT_1. The choice of region of interest in the spleen followed the same principles as that of the liver and was generally easier to perform as there are no blood vessels traversing the spleen parenchyma. The IVC/ Ao ratio was measured, and liver cT_1 measurements were corrected for the effects of intra-vascular fluid status as described in chapter 3. This correction was necessary in one patient (Figure 32). Representative LMS images for the analysis of spleen cT_1 are shown in Figure 33.



Measured liver cT_1: 993ms

IVC cross sectional area: 7.16cm^2

Ao cross sectional area: 4.44cm^2 (>than the 75th percentile for healthy men, see chapter 3)

Therefore for the correction algorithm the median Ao cross sectional area from the study of healthy volunteers was used instead of the measured value (7.16 / 2.69)

Liver cT_1 adjusted for blood volume status: 909ms

Figure 32. Correction of liver cT_1 to account for the effects of intra-vascular blood volume status.

A patient with chronic hepatitis C and a failing renal transplant that was performed 20 years ago was being assessed for repeat renal transplantation. The patient was planned to start haemodialysis as a bridge to repeat renal transplantation. The clinical team wanted to exclude liver cirrhosis in view of thrombocytopenia (platelet count: 137x10^9/l), as if cirrhosis was present the patient may have benefited from a combined liver and kidney transplant. The inferior vena cava (IVC) measured 7.16cm^2 in cross section, suggesting blood volume overload, in keeping with the clinical picture of end stage renal failure. The aorta (Ao) measured 4.44cm^2 is cross section in keeping with some degree of pathological dilatation. Liver histology showed mild fibrosis (Ishak stage 2/6) and severe sinusoidal dilatation, in keeping with liver congestion. Liver cT_1 (993ms) overestimated the degree of fibrosis, but this was improved by the application of the algorithm for adjusting for blood volume status (blood volume adjusted cT_1: 909ms).

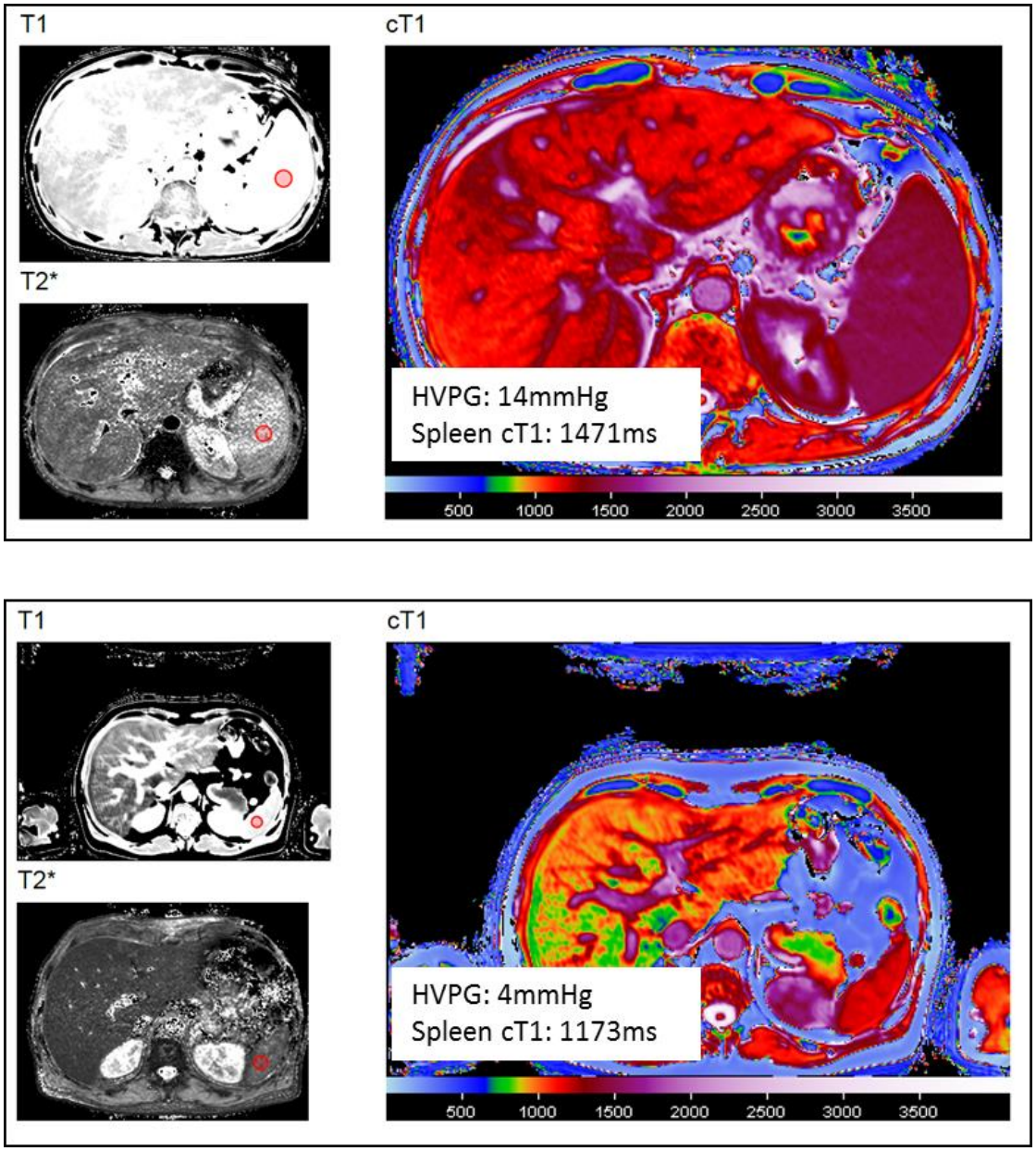


Figure 33. Measurement of spleen cT_1 using LiverMultiScan

Representative images from a patient with (top panel) and without (bottom panel) portal hypertension. Red circles on T_1 and T_2* maps represent the region of interest where the measurements were taken.

6.4.3 Hepatic vein pressure gradient and transjugular liver biopsy

Patients attended after an overnight fast and having stopped any antiplatelet or anticoagulant agents for at least 72 hours. The procedure was carried out under conscious sedation with continuous electrocardiographic and pulse oximetry monitoring. With the patient placed in the supine position, the right internal jugular vein was catheterised under local anaesthetic, using a trans-jugular liver access set (Cook Medical). A straight vascular catheter controlled fluoroscopically was then introduced and was used to measure the wedged hepatic vein pressure (WHVP) and the free hepatic vein pressure (FHVP). The WHVP was obtained by advancing the catheter until a wedged position was achieved, which was checked radiologically. The FHVP was measured 2-4 cm from the confluence of the hepatic vein with the inferior vena cava. For the pressure measurements, the straight catheter was connected to a digital pressure recording device (Datex Ohmeda S/5, GE Healthcare), through a pressure transducer (Namic, Navilyst Medical, Massachusetts, USA). The pressure reading was allowed to stabilise for 30-40 seconds before being recorded.

The hepatic vein pressure gradient (HVPG) was defined as the difference between the WHVP and the FHVP (HVPG=WHVP-FHVP).

On completion of the pressure studies, 3-4 trans-jugular liver biopsies were obtained using a 19G trans-jugular biopsy needle (Cook Medical). Liver biopsies were processed according to the local clinical routine as described in chapter 4.

6.4.4 Statistical analysis

Descriptive statistics were used to summarise subject characteristics. Associations between continuous variables were tested using the Pearson’s coefficient. The Mann-

Whitney test was used to test for differences between cirrhotic and non-cirrhotic patients. Receiver operating curve analysis was used for the assessment of diagnostic accuracy. Cut-offs to maximise sensitivity and specificity were recorded.

Having identified liver fat content (chapter 4) as a confounder in the assessment of liver fibrosis by liver cT_1, multiple linear regression analysis was performed to examine whether inclusion of liver fat content into multi-variate models improved the prediction of HVPG.

The level of statistical significance was set at $p>0.05$. The analysis was done using GraphPad Prism (v6.04). The multiple linear regression analysis was done using SPSS (v22; Armonk, NY, IBM Corp).

6.5 Results

6.5.1. Cohort characteristics

Twenty five patients were recruited, and 19 were included in the final analysis. One patient developed severe alcoholic hepatitis after the MR study. The patient had to be admitted to hospital and missed the planned outpatient biopsy and HVPG assessment. Five further patients that only had measurements of the WHVP were also excluded, as the HVPG could not be calculated (Figure 34). One patient, who was correctly identified by the multi-parametric assessment as having non-cirrhotic portal hypertension (Figure 35), was included in the analysis of spleen cT_1 vs HVPG, but excluded from the analysis of liver cT_1 vs HVPG

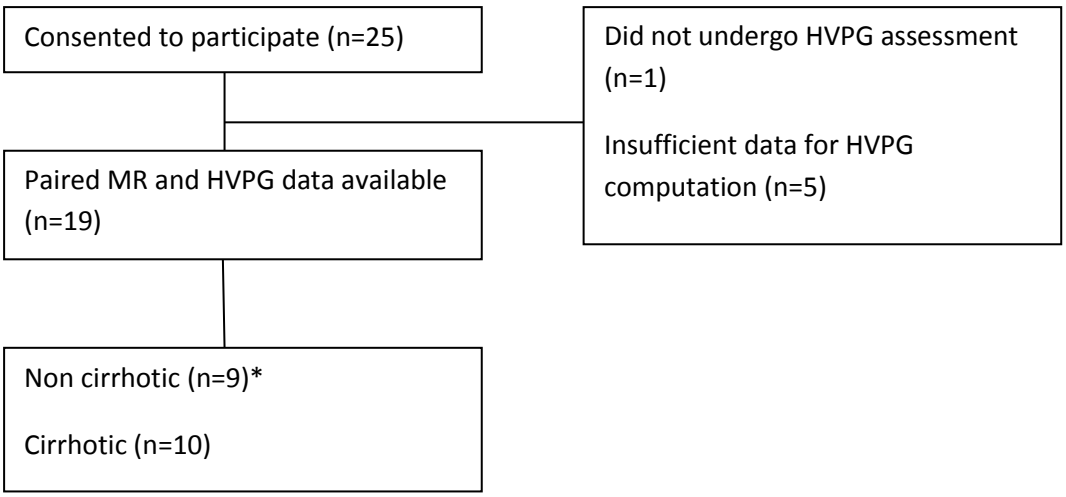


Figure 34. Study flow diagram

*One patient with non-cirrhotic portal hypertension was excluded in the analysis of liver cT_1 vs HVPG.

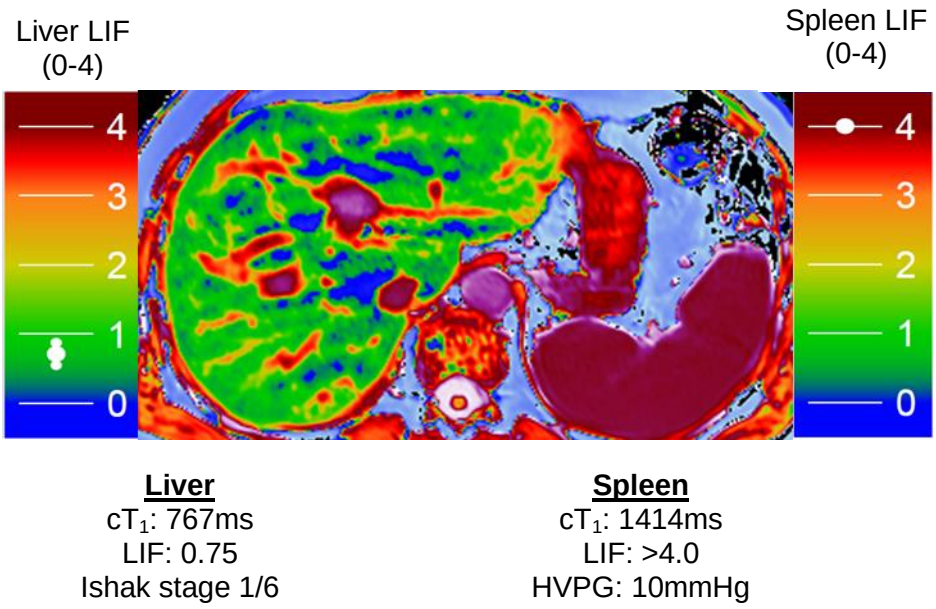


Figure 35. Multi-parametric MR for the assessment of non-cirrhotic portal hypertension

One patient in the cohort of this study had non-cirrhotic portal hypertension secondary to long term azathioprine use for Crohn's disease. The multi-parametric assessment showed a low liver cT_1 of 767ms and LIF of 0.75 which would indicate no significant liver fibrosis. This was confirmed on liver biopsy (Ishak stage 1/6). However, the patient had a high spleen cT_1 (1414ms), indicating portal hypertension, which was confirmed by HVPG measurement (10mmHg).

The majority of the patients were male (n=15, 80%), and the main liver disease aetiologies were NAFLD (n=9, 47%), alcoholic steatohepatitis (n=5; 26%) and viral hepatitis (n=4, 21%). The median age was 57 years old and median BMI was 30.0kg/m^2. The median (IQR) biopsy length was 14mm (9 – 22), containing a median (IQR) of 7 (6 - 10) portal tracts. Ten patients had liver cirrhosis (Ishak 5-6), one patient had moderate fibrosis (Ishak 3) and 8 patients had mild fibrosis (Ishak 1-2).

Compared to non-cirrhotics, patients with cirrhosis had a higher HVPG (12.8 vs 4.0mmHg; p=0.005) and spleen cT_1 (1378 vs 1192ms; p=0.007) and a strong trend

towards a higher liver cT_1 (1022 vs 868ms; p=0.053). No other laboratory tests or simple serum based non-invasive index could differentiate between the two groups (Table 14).

Table 14. Clinical, magnetic resonance and laboratory characteristics

	All (n=19)	Cirrhotic (n=10)	Non-cirrhotic (n=9)	p
Age (years)	57 (47 – 68)	54 (47 – 68)	63 (50.3 – 70)	ns
BMI (kg/m^2)	30.0 (24.2 – 32.3)	28.9 (24.6 – 32.1)	30.1 (23.9 – 34.4)	ns
HVPG (mmHg)	9.0 (4.0 – 14.0)	12.8 (8.3 – 16.3)	4.0 (2.0 – 7.0)	**0.005**
MR data				
Liver cT_1 (ms)	909 (866 – 1114)	1022 (876 – 1202)	868 (813 – 998)	0.053
Spleen cT_1 (ms)	1319 (1181 – 1414)	1378 (1319 – 1443)	1192 (1167 – 1312)	**0.007**
Liver fat (%)	4.9 (1.3 – 13.7)	5.0 (1.6 – 14.5)	2.4 (1.1 – 19.2)	ns
Liver T_2* (ms)	19 (13.5 – 21.9)	18.5 (13.4 – 21.6)	19.0 (13.0 – 22.2)	ns
Laboratory data				
Bilirubin (µmol/l)	12 (10 – 22)	17 (10 – 43)	11 (9 – 16)	ns
ALT (IU/l)	44 (22 – 71)	65 (22 – 75)	30 (22 – 78)	ns
ALP (IU/l)	188 (147 – 236)	219 (146 – 330)	170 (133- 217)	ns
Albumin (g/l)	41 (37 – 46)	40 (34 – 46)	41 (38 – 46)	ns
AST (IU/l)	39 (32 – 76)	69 (37 – 78)	35 (27 – 59)	ns
GGT (IU/l)	93 (40 – 277)	246 (66 – 289)	53 (22 – 139)	ns
Platelets (x10^9/l)	133 (112 – 216)	126 (100 – 202)	137 (117 – 220)	ns
PT (s)	14.4 (13.1 – 16.6)	14.9 (12.3 – 17.2)	14.3 (13.1 – 15.4)	ns
AST/ALT ratio	1.09 (0.84 – 1.65)	1.10 (0.82 – 1.99)	0.97 (0.82 – 1.44)	ns
APRI	0.72 (0.53 – 1.21)	0.89 (0.59 – 1.48)	0.68 (0.39 – 1.03)	ns
FIB-4	3.17 (1.84 – 4.01)	3.36 (1.94 – 5.06)	2.41 (1.70 – 3.64)	ns

All data given as median (IQR). P-values for the comparison between cirrhotic and non-cirrhotic patients (Mann Whitney test)

Abbreviations: **BMI**: body mass index, **HVPG:** hepatic vein pressure gradient, **ALT**: alanine aminotransferase, **ALP**: alkaline phosphatase, **AST**: aspartate aminotransferase, **GGT**: gamma glutamyl transferase, **PT:** prothrombin time

6.5.2 Assessment of HVPG as a continuous variable

There were significant correlations between HVPG and liver cT_1 (r=0.49; p=0.038; Figure 36) and spleen cT_1 (r=0.66; p=0.002; Figure 36). The gradient of the linear regression lines for the prediction of HVPG from liver and spleen cT_1 was similar (11.9 vs 12.3,

p=0.95) and there was a strong correlation between liver and spleen cT_1 (r=0.69; p=0.001, data not shown). There was a significant association between liver cT_1 and liver fat content (r=0.54, p=0.02; data not shown) and inclusion of liver fat measures improved the linear regression models with liver cT_1 (Table 15). There was no association between liver fat and spleen cT_1 and inclusion of liver fat into multi-variate models did not improve their performance (Table 16).

There was no association between HVPG and APRI (r=0.03, p=0.90) or FIB-4 (r=18; p=0.50). There was a strong trend toward an association between AST/ALT ratio and HVPG (r=0.47; p=0.069)

There was one outlying patient, the data from whom are marked with an asterisk on Figure 36. The HVPG measurement in this patient was 17.5mmHg, which was out of keeping with the rest of the clinical picture. His liver biopsy showed only mild fibrosis (Ishak 2), and there were no other radiological features of portal hypertension (no collateral vessels and no splenomegaly on clinically indicated CT scan). Furthermore, liver synthetic function was normal (albumin 40g/l, prothrombin time 14.4s) and platelet count 216 $x10^9$/l. It is therefore possible that the HVPG measurement in this case was not accurate. Exclusion of this data point from the analysis would have resulted in a much stronger correlations with HVPG (for spleen cT_1: r=0.85; p<0.001, for liver cT_1 r=0.61; p=0.009).

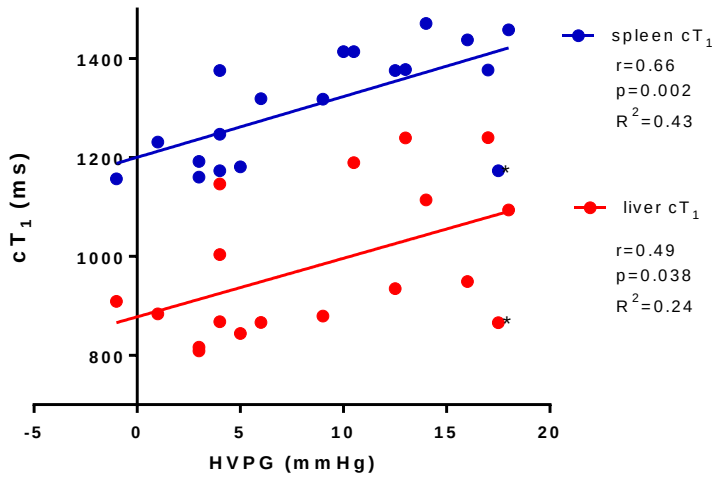


Figure 36. Relationship between the hepatic vein pressure gradient (HVPG) and spleen and liver cT_1.

Table 15. Multiple linear regression models of liver cT_1 and liver fat for prediction of HVPG

Variable	B	P	R^2
Model 1			**0.24**
Constant	-11.43	0.221	
Liver cT_1	0.021	0.038	
Model 2			**0.40**
constant	-19.44	0.052	
Liver cT_1	0.031	0.007	
Liver fat	-0.264	0.067	

Table 16. Multiple linear regression models of spleen cT_1 and liver fat for prediction of HVPG

Variable	B	P	R^2
Model 1			**0.43**
Constant	-37.65	0.009	
Spleen cT_1	0.035	0.002	
Model 2			**0.44**
Constant	-37.32	0.012	
Spleen cT_1	0.036	0.003	
Liver fat	-0.043	0.68	

6.5.3 Diagnostic accuracy against HVPG severity categories

Liver cT_1 showed limited utility in the categorical assessment of portal hypertension with significant diagnostic accuracy only for clinically significant portal hypertension (AUROC 0.83, p=0.03). In contrast, spleen cT_1 had significant diagnostic accuracy for all the severity levels examined (Table 17).

Table 17. Accuracy of liver and spleen cT_1 for the diagnosis of HVPG severity categories

	AUROC	p-value	Cut-off	Sensitivity (%)	Specificity (%)
Liver cT_1					
HVPG >5mmHg	0.75	0.076	-	-	-
HVPG ≥10mmHg	0.83	0.016	922	88	80
HVPG ≥12mmHg	0.77	0.06	-	-	-
Spleen cT_1					
HVPG >5mmHg	0.91	0.003	1283	91	88
HVPG ≥10mmHg	0.91	0.003	1348	89	90
HVPG ≥12mmHg	0.81	0.03	1348	86	75

Abbreviations: **AUROC**: area under the receiver operating characteristic curve, **HVPG**: hepatic vein pressure gradient

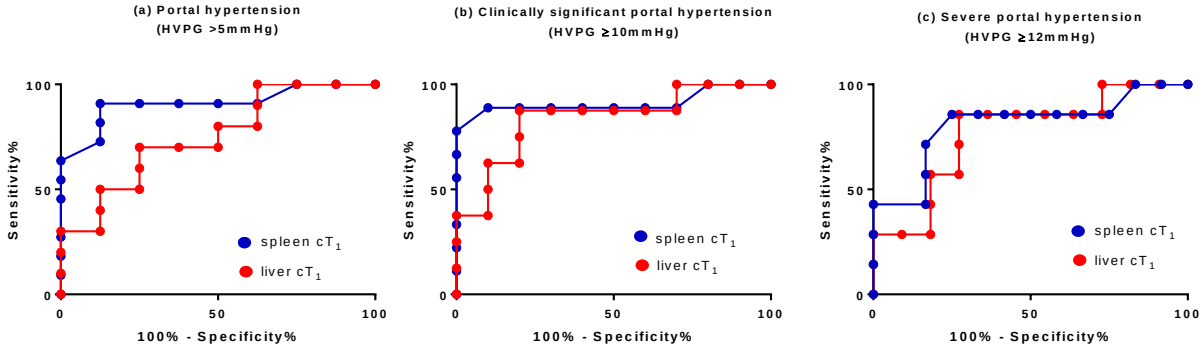


Figure 37. Receiver operating characteristic curves for the accuracy of liver and spleen cT_1 in the diagnosis of (a) any degree of portal hypertension, (b) clinically significant portal hypertension and (c) severe portal hypertension

6.6 Discussion

This study is the first to establish spleen cT_1 as a non-invasive biomarker of portal hypertension. There was a strong overall association with the HVPG (r=0.66, p<0.00, Figure 36) and excellent accuracy for the diagnosis of portal hypertension and clinically significant portal hypertension (AUROC 0.91 for both; Table 17 and Figure 37).

The performance of spleen cT_1 was superior to that of liver cT_1 and this is most likely explained by the presence of confounding factors in the liver. The effects of iron and liver congestion are measured and accounted for by the iron correction[114] and blood volume adjustment algorithms (chapter 3). However, liver fat still appears to exert a significant effect, to such an extent that liver cT_1 has a numerically stronger association with liver fat content (r=0.54, p=0.02; data not shown) than with HVPG (r=0.49, p=0.038; Figure 36). Furthermore, inclusion of fat and liver cT_1 in multi-variate models, improves the accuracy for the prediction of HVPG (R^2=0.24 in the univariate model of liver cT_1; R^2=0.40 in the model including liver cT_1 and liver fat; Table 15).

The main confounder in the spleen is iron, and this is addressed by application of the iron correction algorithm as in the liver[114]. This algorithm was developed based on liver data therefore may not be optimal for application in the spleen. Despite this, the results are encouraging and any further adjustments to the spleen iron correction algorithm are likely to result in improved performance.

Despite the limitation in liver cT_1 measurements, the results suggest a direct relationship between liver fibrosis and portal hypertension. Firstly, there is a strong association between liver and spleen cT_1 (r=0.69, p=0.001; data not shown). Furthermore, it appears that liver and spleen cT_1 change by the same proportion for each unit of HVPG change as demonstrated in Figure 36.

Spleen parameters are routinely used to assess for the presence of portal hypertension. Clinical examination and imaging can be used to assess for splenomegaly, but both these approaches lack sensitivity[178]. Thrombocytopaenia resulting from portal hypertension related splenomegaly and hypersplenism is also routinely assessed in clinical practice and can be useful particularly when incorporated into prediction indices with other non-invasive tests[176,179]

Magnetic resonance imaging techniques of the spleen that have shown some promise include diffusion weighted imaging[180], susceptibility weighted imaging[181,182], and dynamic contrast enhanced MR[183].Liver and spleen stiffness are however the most extensively studied biomarkers of portal hypertension, particularly in patients with viral hepatitis. In a study using transient elastography in patients with HCV, HVPG was closely associated to spleen stiffness (R^2=0.78, p<0.001) and liver stiffness (R^2=0.70, P<0.001)[171]. In another large study where 76% of the patients had viral hepatitis, liver stiffness had an AUROC of 0.88 for the diagnosis of CSPH[179].

However, applications of these techniques to mixed patient cohorts have produced less impressive results. In one study, where 44% of patients had alcoholic liver disease, 30% had viral hepatitis and 26% had cryptogenic cirrhosis, spleen stiffness had only a modest association with HVPG (r=0.433, p=0.001) while liver stiffness did not correlate at all (r=0.178, p=0.2)[184]. In identifying patients with CSPH, the AUROC of liver stiffness and spleen stiffness were respectively 0.79 and 0.72 when stiffness was measured using shear wave elastography, and 0.78 and 0.63 when measured using transient elastography[185].

Liver and spleen stiffness measured using MR elastography have also shown excellent association with HVPG in animal models[50,51]. However, in a patient study, MRE resulted

only in modest association between HVPG and spleen (r=0.53, p0.004) and liver (r=0.44, p=0.05) visco-elastic parameters[186].

The preferred method for evaluating the HVPG is to measure the WHVP and the FHVP using a balloon catheter, where WHVP is measured when the flow in the hepatic vein is occluded by inflating the balloon and the FHVP is taken when the balloon is deflated[169].The use of straight catheters, is associated with more variability in the measurements, and is less accurate[187]. In one study, use of a straight catheter to measure HVPG in two different hepatic veins in the same patient, resulted in a variation from 4 to 34mmHg in the HVPG reading in 61% of the cases[188]. The use of the straight catheter technique for the evaluation of HVPG in this study is therefore a limitation. As it is highlighted in the results section, the HVPG assessment was out of keeping with the rest of the clinical information in at least one patient and this seems to have an adverse effect on the diagnostic accuracy of liver and spleen cT_1.

This study recruited patients who were suspected but not known to have liver cirrhosis. As a result 47% of patients were found to be non-cirrhotic, while, the cirrhotic patients included were at the mild end of the disease spectrum. This high false negative rate of clinical suspicion to diagnose cirrhosis (47% in this study) further highlights the clinical need for better non-invasive tools for liver disease assessment. These demographics may be regarded as a limitation of the study, since portal pressure assessment is only indicated in patients with cirrhosis. However, this was a pilot study that aimed primarily to investigate the association between liver and spleen cT_1 and HVPG, something that was achieved. These results would therefore justify larger studies including only patients with cirrhosis in order to address how this technique could be used in clinical practice.

7. Prediction of clinical outcomes

7.1 Abstract

Background and Aims: Having previously demonstrated that multi-parametric magnetic resonance (MR) imaging can quantify hepatic fibrosis and portal hypertension, the aim of this study was to determine if MR could be used to predict negative clinical outcomes in liver disease patients.

Methods: Patients with chronic liver disease (n=112) were recruited for MR imaging. Data on the development of liver related clinical events were collected by case note review. The median follow up was 27months. MR data were analysed blinded for the Liver Inflammation and Fibrosis score (LIF; <1, 1–1.99, 2–2.99, and ≥3 representing normal, mild, moderate and severe liver disease, respectively), T2* for liver iron content and liver fat content. Baseline liver biopsy was performed in 102 patients.

Results: Liver disease aetiologies included non-alcoholic fatty liver disease (35%) and chronic viral hepatitis (30%). Histologically, fibrosis was mild in 54 (48%), moderate in 17 (15%), and severe in 31 (28%) patients. Overall mortality was 5%. Ten patients (11%) developed at least one liver related clinical event. The negative predictive value of LIF <2 was 100%. Two patients with LIF 2–2.99 and eight with LIF ≥3 had a clinical event. Patients with LIF ≥3 had a higher cumulative risk for developing clinical events, compared to those with LIF <1 (p=0.02) and LIF 1–1.99 (p=0.03). Cox regression analysis including all 3 variables (fat, iron, LIF), resulted in an enhanced LIF predictive value.

Conclusions: Non-invasive standardised multi-parametric MR technology may be used to predict clinical outcomes in patients with chronic liver disease.

7.2 Introduction

The development of liver cirrhosis is a negative milestone in the evolution of liver fibrosis that is associated with morbidity and mortality. The establishment of fibrotic nodules marks a major disturbance in the liver microarchitecture and liver function deteriorates. As fibrosis progresses beyond cirrhosis, the patient is at risk of developing complications from portal hypertension, and poor liver function.

However, even after the development of cirrhosis, patients may remain well ("compensated") for long periods of time, but approximately 5-7% will develop complications (become "decompensated") annually. Furthermore, annual mortality rates can be as high as 57% once decompensation is established[189]. It is therefore, important for non-invasive tests to demonstrate that they can identify patients who are at risk of developing clinical end points as this will facilitate timely therapeutic interventions, liver transplantation, and stratification within clinical studies.

Whereas the assessment of non-invasive tests of liver fibrosis generally examines diagnostic accuracy against liver histology, a more robust and clinically relevant approach would involve the assessment of the ability of non-invasive technologies to predict clinically meaningful endpoints. For example, the enhanced liver fibrosis (ELF) panel reported a sensitivity of 90%, but had a specificity of only 41%, with an area under the receiver operating characteristic curve of 0.80 for the detection of severe fibrosis [190], and a subsequent 7 year follow-up study suggested that ELF score could predict clinical outcomes[27]. However, patients with any disorder associated with extra-hepatic fibrosis were excluded from these studies making it difficult to assess how this test could be applied to a general, unselected population. Liver stiffness measurements using transient or magnetic resonance elastography have also shown promise in the prediction of clinical outcomes[49,191].

However, despite these advances in non-invasive liver assessment, the drawbacks of the currently available techniques mean that they are not widely available and have not been validated for use as surrogate end points in clinical trials. Furthermore, most non-invasive scores have been validated in patients with hepatitis C, and validation mixed cohorts and patients with NAFLD in particular are lacking. For these reasons, several professional and regulatory bodies recognise the need for better non-invasive risk stratification tools[8,192,193].

7.3 Aims

The aim of this study was therefore to assess whether data obtained from this multi-parametric MR protocol could be used to predict all-cause mortality and liver-related clinical events, irrespective of stage of fibrosis or disease aetiology.

7.4 Patients and methods

7.4.1 Study design and patient population

The population under study were those scheduled to have a clinically indicated liver biopsy (n=116) and adult patients who had cirrhosis diagnosed on biopsy within 5 years of their MRI scan (n=7). Patients were included irrespective of underlying liver disease aetiology or disease stage. The only exclusion criterion was the presence of MRI contraindications. Patients were recruited from two UK centres (Oxford and Reading) between April 2011 and August 2013. All patients were followed for the development of clinical outcomes except those who were lost to follow up or had incomplete MR data (n=11; Figure 38). Baseline data were collected at the time of recruitment. Outcome data were collected through review of the individual electronic and paper patient records.

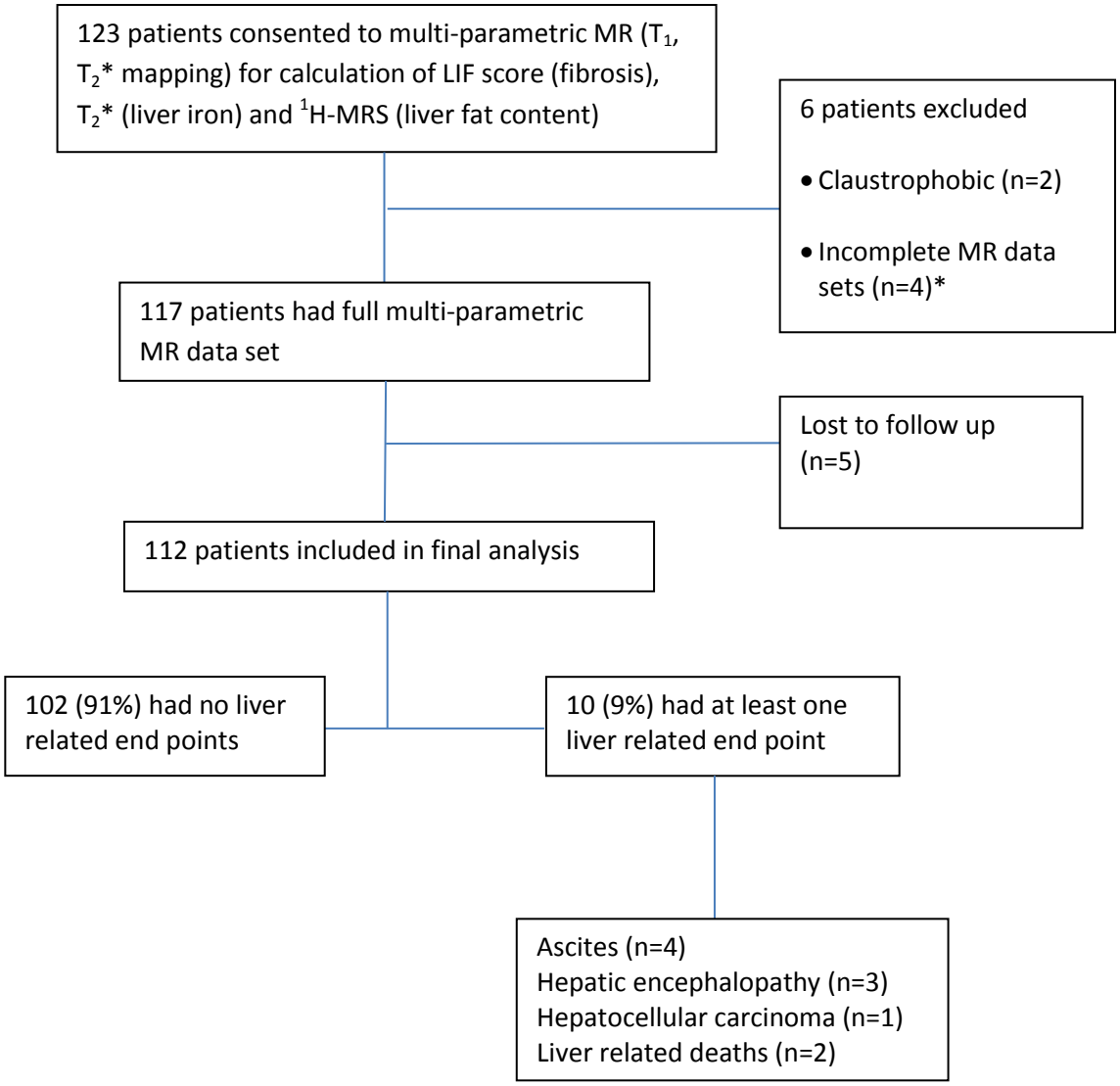


Figure 38. Study flow chart.

Liver iron concentration from T_2 maps and hence LIF calculation was not possible in 4 cases.

Abbreviations: **MR:** magnetic resonance, **^{1}H MRS**: proton (^{1}H) magnetic resonance spectroscopy, **LIF**: liver inflammation and fibrosis score

7.4.2 All-cause mortality and liver related clinical events

Both all-cause mortality and liver related clinical events were evaluated. Liver related clinical events were defined as liver related death, the development of hepatocellular carcinoma (HCC) and any new episode of hepatic decompensation (clinically evident ascites, variceal bleeding, and hepatic encephalopathy). Although more than one event per patient was possible and occurred, patients were only counted once in the analysis at the time of the first liver related clinical event. Patients who had evidence of liver-related complications at or before enrolment were only counted again if they developed a complication that was distinct from that observed before enrolment, or died. Patients were followed up until their last clinical review or until they died, but the index liver related event as defined above was used in the analysis.

7.4.3 Multi-parametric MRI

All patients had multi-parametric MRI as described in chapter 2.

Representative MR data are shown in Figure 39.

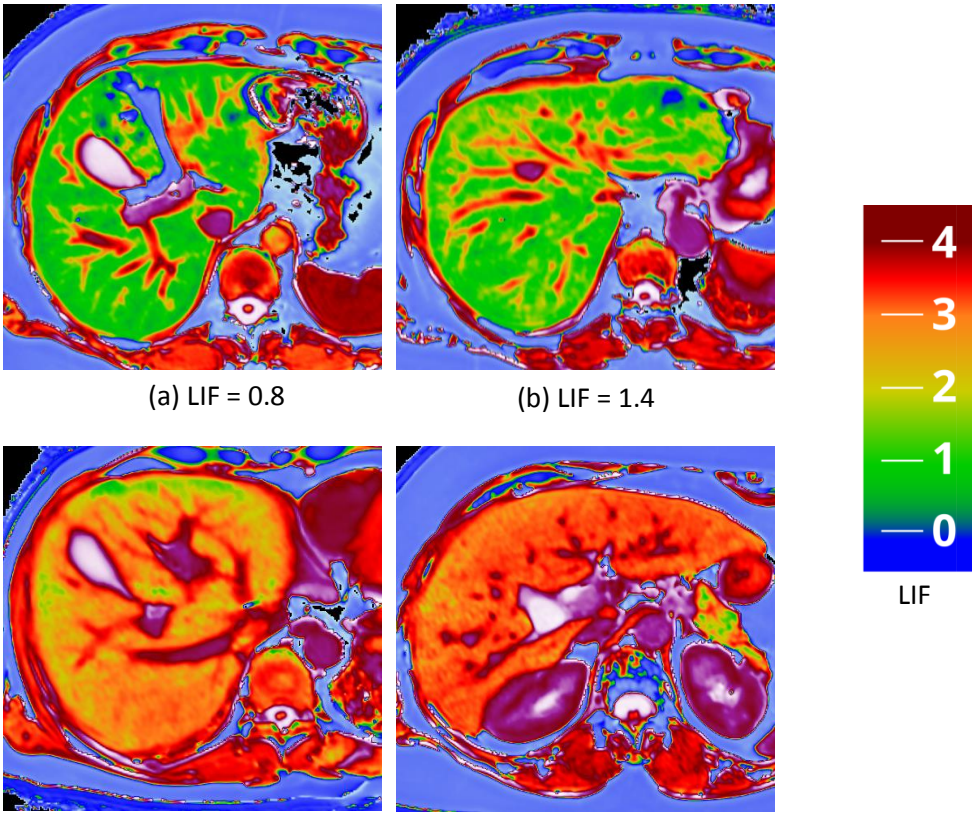


Figure 39. Examples of LiverMultiScan MR data.

Representative images from patients in each Liver Inflammation and Fibrosis (LIF) severity category, produced by analysis of the raw data using LiverMultiScan. LIF was measured in operator chosen regions of interest in the right liver lobe, in the liver parenchyma, away from vascular and biliary structures. The LIF scores measured for each image in this figure are indicated under each image. The predefined colour scale used to generate these maps is also included to the right.

Abbreviations: **LIF**: liver inflammation and fibrosis score.

7.4.4 Statistical analysis

The primary variable of interest was the LIF score measured using LMS. Patients were stratified according to LIF severity as LIF 0-0.99, 1-1.99, 2-2.99 and 3-4. Secondary variables were liver iron measured by T_2* and liver fat measured by ^{1}H-MRS. Cut-off values derived from the work of Banerjee et al[114] were used to define two severity categories for iron (High Liver Iron: T_2* ≤12.5ms, Low Liver Iron: T_2* >12.5ms) and 3 categories for fat fraction (0: fat<1.5%, 1: fat 1.5- <7.5%, 2: fat ≥7.5%). The cut-off for defining the iron severity categories represents the best threshold for identifying those with any grade of histological iron deposition and corresponds to a dry weight iron 1.7mg/g. The cut-offs for defining the liver fat categories represent the best thresholds for identifying those with no steatosis (fat fraction <1.5%) and those with steatosis in more than 66% of hepatocytes (steatosis grade 3; fat fraction ≥7.5%). The primary outcome was the index liver related event as described above (liver related death, HCC, clinically evident ascites, encephalopathy, variceal bleeding). Subgroup analysis including only patients with compensated disease at baseline was also carried out.

Descriptive statistics were used to summarise the subjects' baseline characteristics for different LIF categories. The Shapiro-Wilks test was used to test for normality of the data across the whole cohort. Analysis of variance was used to test for differences between multiple groups. Kaplan Meier analysis was used to establish the proportions of patients in each LIF, fat and iron severity category that remained free of liver related events in the follow-up period. Differences between the curves were compared using the log-rank test. Cox regression analysis was used to examine the additive effects of LIF, liver fat and liver iron (measured by T_2*) on predicting liver related events. The level of statistical significance was set at $p<0.05$. The analysis was carried out using SPSS v22 (Armonk, NY; IBM Corp) software. The GraphPad Prism software (v6.04) was only used to generate the Kaplan Meier curves for the illustration in the figures.

7.5 Results

7.5.1 Baseline characteristics

During the study period 123 patients consented to take part. Six were excluded due to claustrophobia (n=2) or incomplete MR data (n=4). Follow up data were available for 112 of the 117 patients with complete MR data sets (Figure 1). Patients were lost to follow up either because they moved out of the area, or they did not return for any further clinical follow up after their MR scan. The median (IQR) follow up period was 27 (15 – 31) months. The majority of patients were male (n=71, 79%) and the mean (±SD) age was 51 (±13.8) years. Eight (7%) patients had evidence of decompensation at or prior to enrolment, and 16 (14%) patients received treatment for their underlying disease in the follow up period (6 patients with chronic hepatitis C virus infection achieved sustained virological response, 7 patients with chronic hepatitis B achieved viral suppression on treatment and 3 patients with obesity had weight reduction surgery). Baseline clinical, histological and biochemical characteristics for the whole cohort and for the four LIF severity categories are summarised in Tables 1-3.

Table 18: Baseline demographics and MR data

	All (n=112)	LIF <1 (n=22)	LIF 1 - 1.99 (n=34)	LIF 2 – 2.99 (n=18)	LIF 3 - 4 (n=38)	p
Age; years (mean ±SD)	51 ± 13.8	47.6 ± 12	49.3 ± 16	52.8 ± 10	51.7 ± 12.4	0.5
Male; n (%)	79 (71)	15 (71)	25 (74)	17 (94)	22 (58)	0.04
BMI; kg/m^2 (median; IQR)	27.5 (24.7 – 30.7)	27.1 (24.0 – 28.6)	26 (23.4 - 28.5)	27.1 (25.2 – 32.9)	29.2 (26.3 – 34)	0.02
T2DM; n (%)	16 (14)	0	2 (6)	5 (28)	9 (24)	0.04
Excessive alcohol; n (%)	27 (24)	5 (23)	4 (12)	5 (28)	13 (34)	0.47
Decompensation at baseline; n (%)	8 (7)	0	1 (3)	0	7 (18)	0.03
Treatment for the underlying disease in the follow-up period; n (%)	16 (14)	3 (14)	7 (21)	4 (22)	2 (5)	0.21
Mean (±SD) liver fat content by ^{1}H-MRS (%)	10.3 (12.6)	3.5 (3.3)	8.2 (10.8)	7.1 (7.6)	17.6 (15.3)	0.01
Mean (±SD) T_2*(ms)	17.8 (6.6)	19.2 (5.4)	17.8 (6.4)	17.5 (7.3)	17.2 (7.1)	0.74

p values quoted for the differences between the 4 LIF severity groups.

Abbreviations: **LIF**: liver inflammation and fibrosis score, **T2DM**: type 2 diabetes mellitus , **^{1}H-MRS**: proton magnetic resonance spectroscopy

Table 19: Baseline histological data

	All (n=112)	LIF <1 (n=22)	LIF 1 - 1.99 (n=34)	LIF 2 – 2.99 (n=18)	LIF 3 - 4 (n=38)	p
Primary histological diagnosis; n (%)						
Viral hepatitis	34 (30)	5 (23)	15 (44)	6 (33)	8 (21)	0.02
NAFLD	39 (35)	5 (23)	8 (24)	9 (50)	17 (45)	
ASH	12 (1)	1 (5)	1 (3)	2 (11)	8 (21)*	
PSC/PBC/AIH	7 (6)	3 (14)	2 (6)	1 (6)	1 (3)	
Other	10 (9)	5 (23)	4 (12)	0	1 (3)	
No biopsy	10 (9)	3 (14)	4 (12)	0	3 (8)	
Histological fibrosis; Ishak stage; n (%)						
Mild; F0-2	54 (48)	18 (82)	17 (50)	9 (50)	10 (26)	<0.01
Moderate; F3-4	17 (15)	1 (5)	9 (26)	1 (6)	6 (16)	
Severe; F5-6	31 (28)	0	4 (12)	8 (44)	19 (50)	
No biopsy	10 (9)	3 (14)	4 (12)	0	3 (8)	

p values quoted for the differences between the 4 LIF severity groups.

*One patient included here had alcoholic hepatitis and cirrhosis on biopsy. The patient did not develop any liver related clinical events.

Abbreviations: **NAFLD**: non-alcoholic fatty liver disease, **ASH**: alcoholic steatohepatitis, **PSC:** primary sclerosing cholangitis, **PBC**: primary biliary cirrhosis, **AIH**: autoimmune hepatitis

Table 20. Baseline laboratory data

	All (n=112)	LIF <1 (n=22)	LIF 1 - 1.99 (n=34	LIF 2 – 2.99 (n=18)	LIF 3 - 4 (n=38)	p
Bilirubin (µmol/l)	12	10.5	10	12.5	12	0.55
	(8-17)	(6 – 14.5)	(7 – 16.5)	(8 – 24)	(8 – 18)	
ALT (IU/l)	60 (35 - 99)	57 (38 – 97)	40 (29 – 80)	70 (35 – 86)	78 (37 – 126)	0.21
ALP (IU/l)	200	191	200	205	206	0.93
	(161 - 288)	(145 – 321)	(157 – 253)	(156 – 290)	(168 – 300)	
Albumin (g/l)	45 (43 - 47)	46 (44 – 47)	45 (43 – 46)	45 (43 – 48)	44 (40 – 46)	0.13
GGT (IU/l)	78	51	47	81	100	0.006
	(36 – 153)	(28 – 131)	(26 – 59)	(48 – 243)	(65 – 230)	
AST (IU/l)	46 (31 - 76)	36 (31 - 57)	33 (26 - 59)	53 (32 – 96)	66 (41 – 112)	0.001
Platelet count ($x10^9$/l)	188	203	205	181	167	0.11
	(152 - 248)	(179 – 266)	(155 - 257)	(133 – 212)	(144 – 240)	
Prothrombin time (s)	13.8	13.3	13.6	13.4	14.5	0.001
	(13 – 14.6)	(12.9 – 14.1)	(13 – 14.2)	(12.9 – 14.7)	(13.5 – 15.6)	

Data given as median (IQR). p values quoted for the differences between the 4 LIF severity groups.

Liver biopsy was performed in 102 patients. Chronic viral hepatitis and non-alcoholic fatty liver disease (NAFLD) were the two most common histological diagnoses affecting 34 (30%) and 39 (35%) patients, respectively. Histologically, liver fibrosis was assessed using the Ishak score[120], and staged as mild (F0-2) in 54 (48%), moderate (F3-4) in 17 (15%), and severe (F5-6) in 31 (28%) patients.

When patients were stratified using the LIF score; 38 (34%) had severe disease, 18 (16%) had moderate disease, 34 (30%) had mild disease, and 22 (20%) had no liver disease. More patients with high LIF scores had severe histological fibrosis (LIF 3-4: 50%, LIF 2–2.99: 44%, LIF 1-1.99: 12%, LIF <1: 0%, p<0.001). Furthermore, patients with high LIF scores had higher BMI (LIF 3-4: 29.2kg/m^2, LIF 2–2.99: 27.1kg/m^2, LIF 1–1.99: 26kg/m^2, LIF <1: 27.1kg/m^2, p=0.02), higher liver fat content (LIF 3-4: 17.6%, LIF 2–2.99: 7.1%, LIF 1–1.99: 8.2%, LIF <1: 3.5%, p=0.01) and were more likely to have type 2 diabetes (LIF 3-4: 24%, LIF 2-2.99: 28%, LIF 1–1.99: 6%, LIF <1: 0%, p=0.04).

Seven out of the eight patients who had baseline decompensation had severe disease by LIF criteria (LIF ≥ 3) and four of these went on to develop further liver related clinical events. One patient with baseline decompensation (ascites) had mild disease by LIF criteria (LIF 1-1.99) and remained event free during follow up.

7.5.2 Follow-up data

Ten (9%) patients developed an index liver related event (4 ascites, 3 encephalopathy, 1 HCC, 2 liver related deaths), and six patients died during the follow up period. There were no deaths relating to causes other than liver disease. The clinical features of the ten patients who had at least one liver related clinical event are detailed in Table 21.

Table 21. Clinical features of patients who had liver related events in the follow up period.

LIF score	Ishak stage	CP	Diagnosis	Index event (months from scan)	Patient mortality (months from scan)
2.5	6	B	NAFLD	Encephalopathy (15)	Yes (24)
2.7	6	A	HCV	HCC (11)	Yes (19)
3.0	6	A	HCV	Encephalopathy (7)	No
3.1	6	B	HCV	Ascites (15)	No
3.5	5	A	ASH	Liver related death (3)	Yes (3)
3.5	6	A	NAFLD	Ascites (22)	No
3.5	6	A	ASH	Ascites (3)	Yes (4)
3.6	6	A	ASH	Encephalopathy (5)	No
3.8	6	B	NAFLD	Liver related death (7)	Yes (7)
3.8	No biopsy	A	ASH*	Ascites (7)	Yes (10)

*The patient took part in the imaging study but did not attend his liver biopsy. He had a history of excessive alcohol use suggestive of a diagnosis of alcoholic steatohepatitis. Abbreviations. **LIF:** liver inflammation and fibrosis score, **CP**: Child Pugh class, **NAFLD**: non-alcoholic fatty liver disease, **HCV**: hepatitis C, **ASH**: alcoholic steatohepatitis, **HCC**: hepatocellular carcinoma

There were no liver related events in patients with a LIF score of less than 2 (negative predictive value 100%). Ten out of 56 (18%) patients with a LIF ≥2 experienced at least one liver related event. Kaplan Meier analysis showed that patients with a LIF ≥3 had a significantly higher cumulative risk of developing liver associated clinical complications and death over time, compared to those with LIF <1 (p=0.02) and LIF 1-1.99 (p=0.03). Furthermore, there was a strong trend towards a higher risk of complications in those with LIF 2-2.99 compared to LIF 1-1.99 (p=0.054, Figure 40).

When patients with decompensation at or prior to baseline were excluded from the analysis, those with severe disease (LIF 3-4) were still at an increased cumulative risk of complications compared to those with mild disease (LIF 1-1.99; p=0.023), and there was

a strong trend for a difference between those with moderate disease (LIF 2-2.99) and those with mild disease (LIF 1-1.99; p=0.058, Figure 40).

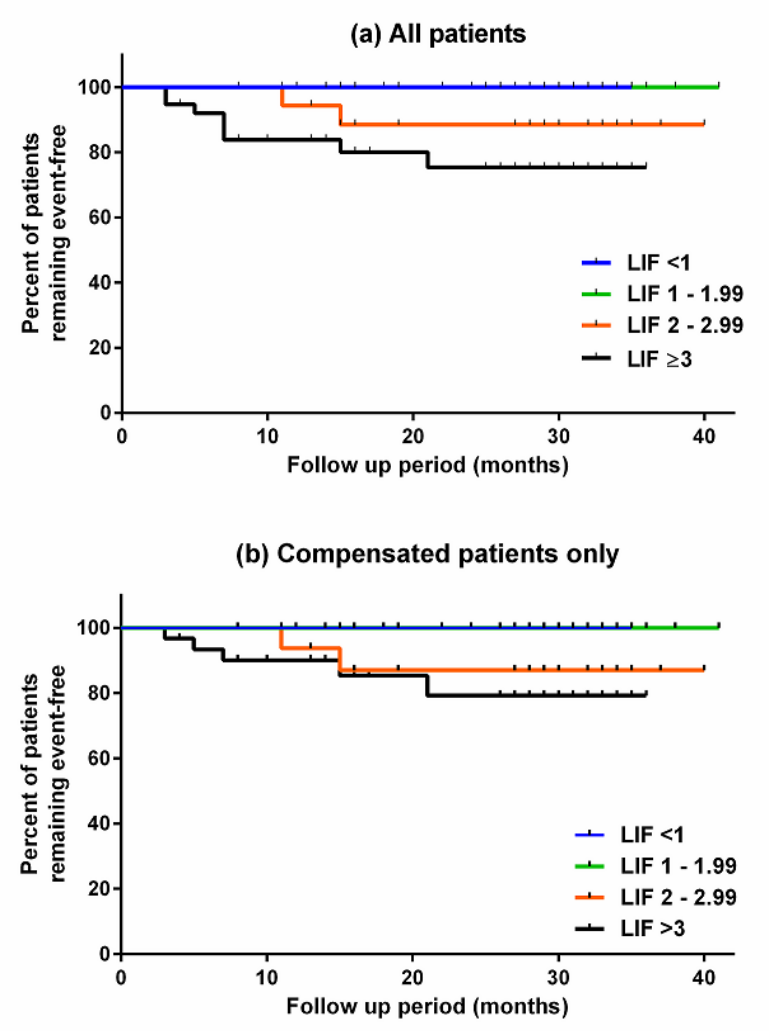


Figure 40. Kaplan Meier curves for liver related event free survival with patients stratified according to liver inflammation and fibrosis scores.

In the entire cohort, (a) there were significant differences between those with Liver Inflammation and Fibrosis (LIF) ≥3 vs LIF <1 (p=0.02) and vs LIF 1-1.99 (p=0.003). There was a strong trend towards significance between LIF 2-2.99 vs LIF 1–1.99 (p=0.054). Including only compensated patients at baseline, (b) there was a significant difference between LIF ≥ 3 vs LIF 1–1.99 (p=0.023) and a strong trend toward significance between LIF 2–2.99 vs LIF 1–1.99 (p=0.058).

There were no differences in the risk of developing liver related clinical events when patients were stratified according to MR determined severity categories for liver fat and liver iron deposition (Figure 41).

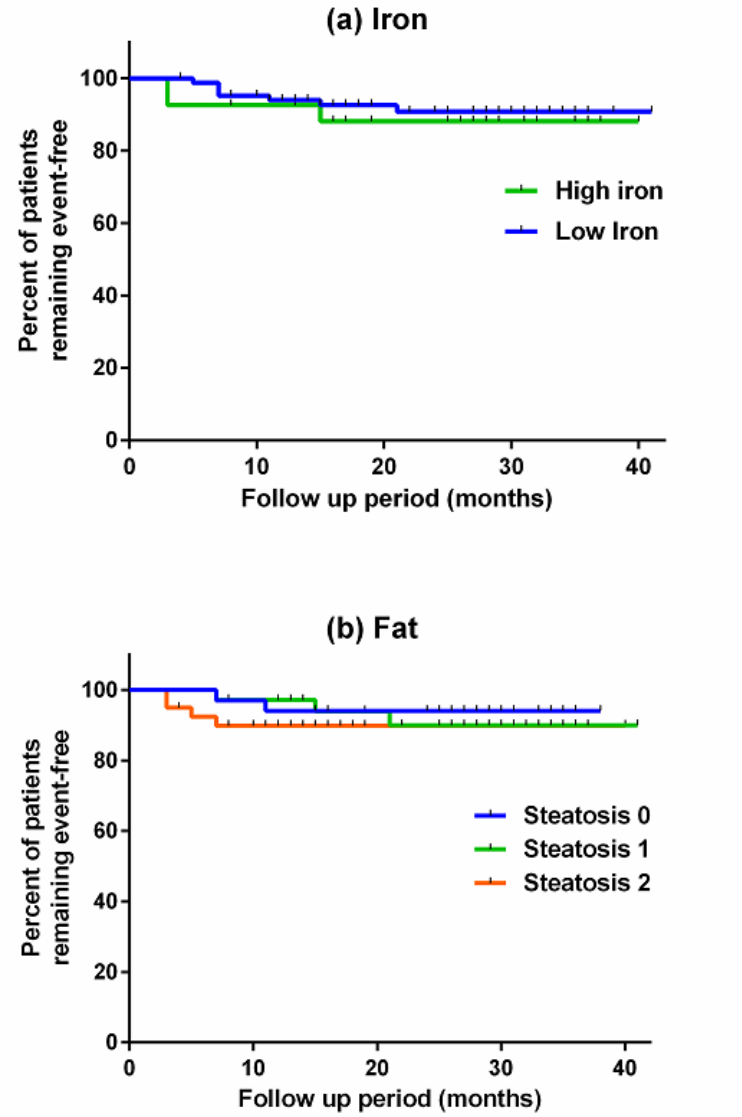


Figure 41. Kaplan Meier curves for liver related event free survival with patients stratified according to severity of liver iron and fat.

There were no significant differences between the curves for (a) iron or (b) fat. Liver fat was categorised according to the liver fat content measured by proton magnetic resonance spectroscopy as 0: < 1.5%, 1: 1.5% < fat < 7.5% and 2: ≥7.5%. Liver iron was categorised according to T2* as low iron: T2* > 12.5ms or high iron: T2* ≤ 12.5ms.

7.5.3 Cox regression analysis

LIF had a hazard ratio (HR) of 9.7 (95% CI 2.2 – 43.2; p=0.003) when it was entered as the sole variable in the model. The HR increased to 75.7 (95%CI: 7.6 – 752; p<0.001) when all 3 variables were entered in the model. Furthermore, in the model containing all 3 variables liver fat (HR 0.80; 95% CI 0.70 – 0.91; p<0.001) and T2* which is inversely related to iron load (HR 0.87; 95% CI 0.80 – 0.95; p=0.001) had a protective effect. The results of the Cox regression analysis with all 4 possible covariate combinations including LIF are shown in Table 22.

Table 22. Cox regression analysis for the prediction of liver related events for LIF, fat and T2*

Variable	HR	95% CI	p
Model 1			
LIF	9.7	2.2 – 43.2	0.003
Model 2			
LIF	30.5	3.3 – 284	0.003
Liver fat	0.95	0.9 – 1.0	0.084
Model 3			
LIF	8.2	1.9 – 35.3	0.004
T_2*	0.97	0.9 – 1.1	0.51
Model 4			
LIF	75.7	7.6 – 752	<0.001
Liver fat	0.80	0.70- 0.91	<0.001
T_2*	0.87	0.80- 0.95	0.001

Abbreviations: **LIF**: liver inflammation and fibrosis score, **HR:** hazard ratio, **95% CI**: 95% confidence interval

7.6 Discussion

This study shows that the Liver Inflammation and Fibrosis (LIF) score, a newly established MR score for assessment of liver fibrosis and inflammation, strongly predicts clinical outcomes in patients with chronic liver disease of mixed aetiologies. No patients with LIF <2 developed liver related clinical events, in contrast to almost 1 in 5 (18%) of those with LIF ≥2. Furthermore, quantification of liver fat content and liver iron using a multi-parametric MRI technique added to the predictive value of LIF.

This is the first study to demonstrate the value of a liver T_1 mapping technique in predicting hard clinical end points. It follows from previous work showing that multi-parametric MR had good diagnostic accuracy compared to histology, particularly for the diagnosis of early disease, with an area under the receiver operating characteristic curve of 0.94 for the identification of those with any degree of fibrosis[114]. Other studies of similar T_1 mapping techniques have also shown promise in differentiating patients with cirrhosis from normal controls[70,71,88].

This study included patients at all stages of fibrosis and those with known cirrhosis. Patients with cirrhosis were specifically included, since the presence of cirrhosis on a liver biopsy is simply quantified categorically into present/absent. The advantage of multi-parametric MR is that this is able to quantify fibrosis along a numeric continuum, and so stratify patients within the category of cirrhosis into those at high risk of decompensation.

MR-quantified liver fat content and liver iron did not provide any useful prognostic information when these parameters were considered in isolation. However, in the multivariate analysis using the Cox regression model, inclusion of fat and iron, added to the predictive value of LIF increasing the HR from 9.7 to 75.7 (Table 3). Furthermore, in the model including all 3 covariates, liver fat and T2* had a protective effect.

T2* in inversely related to iron load, therefore the protective effect seen with decreasing liver iron (increasing T2*) would be in keeping with previous observations that increased iron is associated with worse prognosis [119] and more aggressive NAFLD phenotypes[194].

Previous studies have demonstrated that liver fat decreases as fibrosis progresses[195], and also that liver fat is a risk factor for fibrosis progression[153,196]. The observation here that liver fat is protective, may seem at odds with the published literature. However, the most likely explanation for this finding is that patients with high liver fat would be more likely to have early stages of fibrosis and were therefore at lower risk developing clinical outcome within the timeframe of this study

One other important observation from this study was that patients with higher LIF scores had higher BMI, were more likely to have type 2 diabetes and had higher liver fat content. Diabetes, obesity and liver steatosis have all been associated with more severe liver fibrosis [197-199]. In contrast to what might be expected, in this study 26% of patients with severe disease assessed by MR (LIF ≥3) and 50% of patients with moderate disease (LIF 2-2.99) had only mild fibrosis using histological staging (Ishak 0-2). The reason for the lack of concordance, between histology and MR parameters in these patients, is not clear. However, whilst this may reflect MR limitations, due to confounding from liver fat, it is also plausible that histology under-staged the fibrosis severity or that these patients had an inflammatory burden that is not adequately assessed by histology. Further work and longer follow up is needed to establish whether patients with mild histological fibrosis but high LIF (≥2) scores are at increased risk progression and decompensation compared to patients with mild histological fibrosis and low LIF (<2) scores.

This was a small proof of principle study that included patients across a spectrum of liver disease severity and aetiologies. One limitation was the relatively short follow up duration. However, in spite of this, it demonstrated a utility for multi-parametric MR in defining

prognosis and the development of clinical endpoints. These results therefore support the further evaluation of multi-parametric MR in specific patient populations over longer follow up periods. Since the majority of patients included in this study were ones that had been referred for liver biopsy by their clinical team, those with clinically evident cirrhosis, and those with no clinical suspicion for significant liver disease may be under-represented.

Due to the relatively short follow up, only patients with liver cirrhosis developed complications. The inclusion of all patients in the main analysis irrespective of the presence of decompensation at baseline may have biased the results in the overall cohort. However, when patients with clinically evident liver decompensation were excluded from the analysis, statistically significant differences still remained between the LIF severity groups. Furthermore, seven out of the ten patients that developed complications were classified as Child Pugh class A cirrhosis at baseline, suggesting that the LIF score may provide valuable additional independent prognostic data in patients with compensated liver disease.

In conclusion, this study provides the first evidence that rapid, non-invasive multi-parametric MRI technology that quantifies liver inflammation and fibrosis in a standardised test (LIF score), may provide valuable prognostic information in patients with chronic liver disease. This finding extends the validation of this technique against histology and HVPG. Overall, these data support the further evaluation and development of multi-parametric MR as a technology that can be used to stage liver disease, identify patients who will progress to cirrhosis and develop complications of liver disease. This would enable timely clinical interventions like transplantation and drug therapy, and facilitate clinical trials as a robust surrogate endpoint.

8 Conclusion and suggested further work

8.1 Summary of main findings

The work presented in this thesis supports the continued evaluation of multi-parametric MR in the assessment of patients with liver disease. A major advantage that has emerged from this work is the ability of the technique to assess the whole spectrum of liver disease (Figure 42), something that has been identified as a priority[200] in the field. Liver cT_1 has shown close correlation with morphometrically quantified fibrosis. In NAFLD, liver cT_1/ LIF had good diagnostic accuracy for NASH, which is the earliest manifestation of aggressive disease and for fibrosis. Furthermore, applicability was superior to transient elastography, without any statistically significant differences in diagnostic performance. Incorporation of spleen cT_1 allows for the severity assessment of patients at the more advanced stages of disease while liver cT_1 / LIF and quantification of liver fat content and liver iron can be incorporated into multivariate models to provide strong predictive power for the development of adverse clinical outcomes.

This work has identified two further liver T_1 confounders in addition to liver iron and inflammation (Figure 43). The study in healthy volunteers showed that the major physiological determinant of liver T_1 is liver congestion, which can be quantified by the IVC/Ao ratio. Furthermore, liver fat content was shown to have a confounding effect in the use of liver cT_1 as a biomarker of fibrosis and portal pressure. Both the confounding effects can be removed from the measurements through the application of the blood volume adjustment algorithm described in chapter 3, or though the incorporation of liver fat measures into multi-variate models (Table 23).

Even though the work to date has shown that quantification of confounders and adjustments to account for them, lead to improved accuracy of liver T_1 / LIF, validation of

this approach in independent cohorts would be required and should be the focus of future studies.

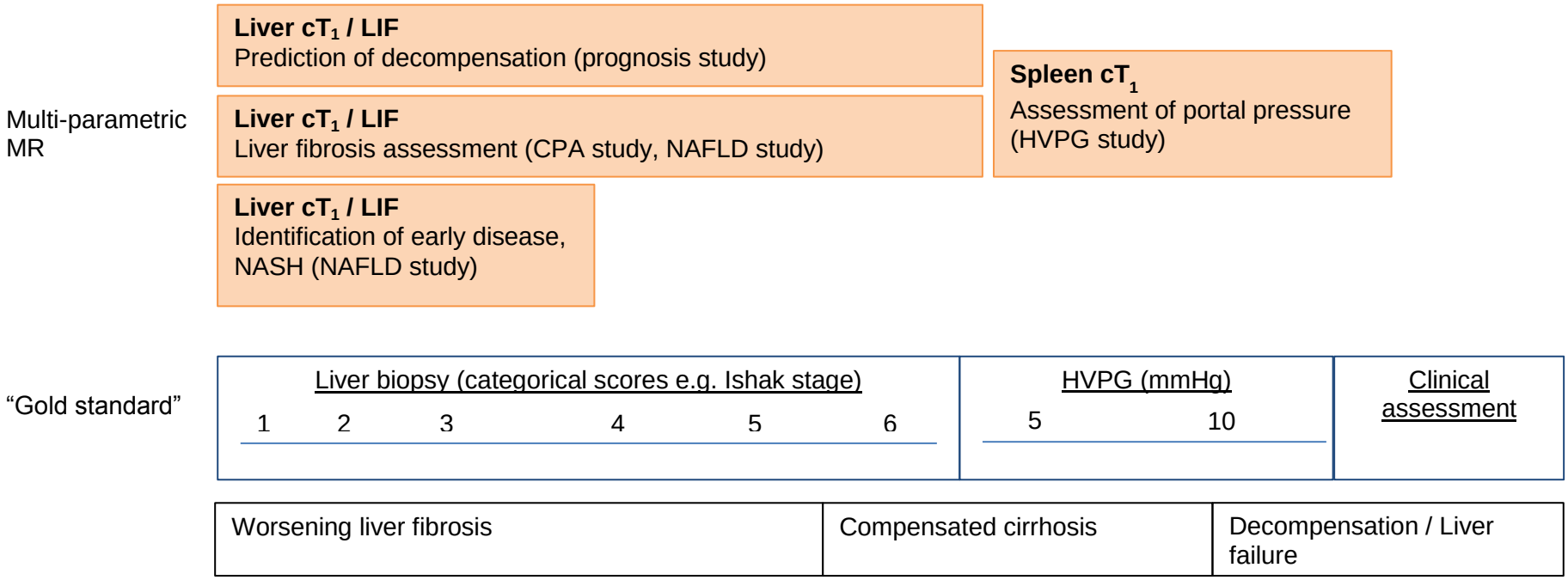


Figure 42. Diagrammatic summary for potential applications of multi-parametric magnetic resonance in the evaluation of disease severity at different stages of chronic liver disease, based on the results of studies in this thesis

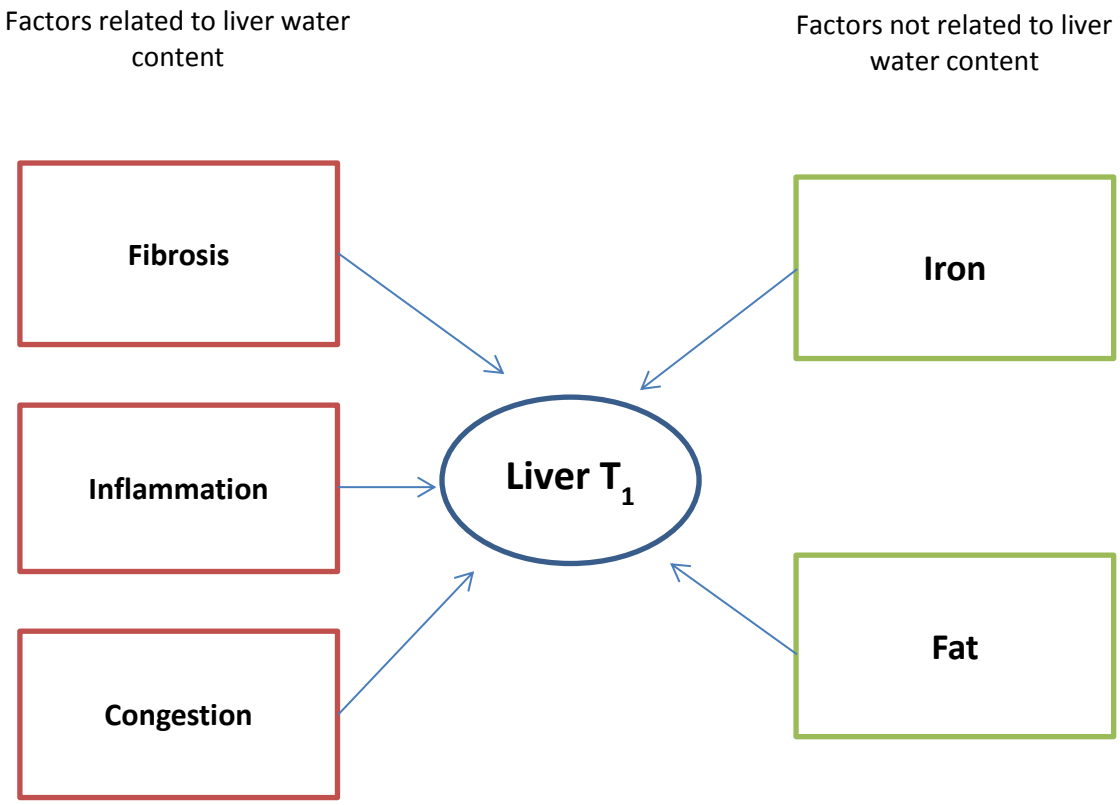


Figure 43. Factors that influence liver T_1 measurements

Table 23. Strategies to improve the performance of liver T_1 as a biomarker of fibrosis and inflammation

Confounding factor	Quantification of confounder	T_1 adjustment
Congestion	IVC/Ao ratio	Blood volume adjusted liver T_1
Iron	T_2*	Iron corrected T_1
Fat	Liver fat content from ^{1}H-MRS	Multi-variate analysis

Abbreviations: **IVC/Ao**: Inferior vena cava / aorta cross sectional area, **^{1}H-MRS**: proton magnetic resonance spectroscopy

8.2 Spleen cT_1

The study investigating the application of multi-parametric MR in the investigation of portal hypertension has identified spleen cT_1 as a biomarker of portal pressure. Despite the small number of patients in the study, it was evident that spleen cT_1 had a stronger association with HVPG rather than liver cT_1, most likely due to the confounding effect of liver fat content on liver cT_1. This is not surprising for two reasons: (a) it would be expected that the spleen provides a more accurate reflection of portal pressure, and (b) the only potential confounder in the spleen is liver iron, which is removed by the application of the iron correction algorithm.

Consideration of spleen cT_1 together with the established parameters of the Oxford multi-parametric MR protocol (liver cT_1/ LIF, liver fat content and liver iron) has the potential to improve the overall performance of the technique. As seen in the HVPG study, a combination of low liver cT_1 / LIF with a high spleen cT_1 correctly classified a patient with non- cirrhotic portal hypertension.

The relative lack of confounders in the spleen, make spleen cT_1 an attractive biomarker to explore further. A study using DCE MRI showed that spleen parameters could correctly classify fibrosis in an animal model[183]. It is therefore plausible that portal haemodynamic changes occur before cirrhosis is established[50], which could be quantified using sensitive MR techniques like spleen cT_1, and this should be addressed in future studies.

8.3 Comparison with other non-invasive tests

Unlike serum biomarkers, multi-parametric MR is liver specific and with a higher applicability than transient elastography. Furthermore, it can be performed without the need for additional hardware, unlike MRE (Table 24). In the liver inflammation and fibrosis

score, the cut-offs for fibrosis stages are standardised and defined, something that is lacking from all other non-invasive biomarkers of fibrosis. In the recent MRE meta-analysis for example, the best cut-offs for liver fibrosis classification based on the pooled data are different from the cut-offs reported in the individual studies[43], something that is also a major criticism for transient elastography[201]. Further studies validating the LIF cut-offs are therefore required.

The above features of multi-parametric MR in addition to the results presented here would support this technique as a robust biomarker of liver fibrosis. Further validation is required in all areas and future studies should focus on cost effectiveness analysis of various diagnostic pathways in order to determine the place of this technique in clinical practice.

Table 24. Comparison of multi-parametric liver MR with other non-invasive biomarkers of liver fibrosis

	Serum biomarkers	Transient elastography	MR Elastography	Multi-parametric MR
Applicability	High[202]	Limited[158,203]	High	High, similar to MRE
Reproducibility	High[204]	Conflicting reports[174,205]	High[206,207]	High
Liver specific	No	Yes (but blind choice of region of interest)	Yes	Yes
Can detect regional variability	No	No	Yes	Yes
Quick to perform	Yes, but delay to get the results	Yes	No	Yes
Confounders	Yes[208]	Yes [209,210]	Yes [211,212]	Yes, but measures can be adjusted to account for the confounders
Accuracy for cross sectional assessment	Yes, but mainly validated in HCV	Yes[201], but mainly validated in HCV	Yes[43]	Yes[114]
Well defined cut-offs	No	No[201]	No[43]	Yes; further validation required
Accuracy for longitudinal assessment	Poor[28]	Poor[213]	No data	No data
Prognostic value	Yes[27]	Yes[191,214]	Yes[49], further validation required	Yes, further validation required

8.4 Multi-parametric MR as a surrogate end point

Iron corrected T_1 and the liver inflammation and fibrosis score (LIF) are two of the newest biomarkers of liver fibrosis and therefore lack the validation and experience that is available for other non-invasive tests like transient elastography that was first described in 2003[22]. However, despite this extensive experience in non-invasive liver fibrosis assessment, none of the currently available tests have been validated as surrogate end point in clinical trials.

It is well documented that this is an area of urgent need as the lack of reliable surrogate endpoints is impacting on drug development for liver disease. Currently, clinical trials generally rely on histological assessment of fibrosis requiring multiple biopsies over short time frames. This is frequently unacceptable to patients so limiting recruitment, whilst concurrently the sampling and observer dependent variability of liver biopsy will increase the number of patients required in a study to achieve adequate statistical power to meet the study end points.

Many studies have shown that liver stiffness and serum biomarkers decrease[215,216] after therapeutic intervention in hepatitis C. However, these observations would be a lot more useful if they were accompanied by outcome data to demonstrate what magnitude of change in the non-invasive score is associated with improved clinical outcomes. It is not therefore surprising that when non –invasive tests were compared to histology before and after therapeutic intervention they were found to be of limited value[213]. As a result current non-invasive biomarkers are not recommended for monitoring therapeutic interventions, or for the diagnosis of cirrhosis regression[20]. Studies to quantify the decrement of liver cT_1 / LIF that is associated with improved clinical outcomes is therefore likely to distinguish this test from other non-invasive tests, and would be needed for any tests considered as a

surrogate end point in clinical trials. This field should be a major focus of future studies for all for all non-invasive biomarkers as development of robust surrogate end points would revolutionise therapeutics development in liver disease.

9. References

1. Wong VW, Chu WC, Wong GL, et al. Prevalence of non-alcoholic fatty liver disease and advanced fibrosis in Hong Kong Chinese: a population study using proton-magnetic resonance spectroscopy and transient elastography. *Gut* 2012; **61**(3): 409-15.

2. Browning JD, Szczepaniak LS, Dobbins R, et al. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. *Hepatology* 2004; **40**(6): 1387-95.

3. Szczepaniak LS, Nurenberg P, Leonard D, et al. Magnetic resonance spectroscopy to measure hepatic triglyceride content: prevalence of hepatic steatosis in the general population. *Am J Physiol Endocrinol Metab* 2005; **288**(2): E462-8.

4. Hanafiah KM, Groeger J, Flaxman AD, Wiersma ST. Global epidemiology of hepatitis C virus infection: New estimates of age-specific antibody to HCV seroprevalence. *Hepatology* 2013; **57**(4): 1333-42.

5. Schweitzer A, Horn J, Mikolajczyk RT, Krause G, Ott JJ. Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. *Lancet* 2015.

6. Blachier M, Leleu H, Peck-Radosavljevic M, Valla DC, Roudot-Thoraval F. The burden of liver disease in Europe: a review of available epidemiological data. *J Hepatol* 2013; **58**(3): 593-608.

7. GLOBOCAN2012: Estimated cancer indidence, mortality and prevalence worldwide in 2012. http://globocan.iarc.fr/Pages/fact_sheets_population.aspx (accessed 27TH aUG 2015 2015).

8. 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*; **384**(9958): 1953-97.

9. Effiong Kufre OA, Pring Andy. Deaths from liver disease. Implications for end of life care in the UK: National End of Life Care Intelligence Network, 2012.

10. Davies DS. CMO Annual Report 2011: On the state of the Public's Health, 2012.

11. Trautwein C, Friedman SL, Schuppan D, Pinzani M. Hepatic fibrosis: Concept to treatment. *Journal of hepatology* 2015; **62**(1 Suppl): 24.

12. Webster DP, Klenerman P, Dusheiko GM. Hepatitis C. *The Lancet*; **385**(9973): 1124-35.

13. Bravo AA, Sheth SG, Chopra S. Liver Biopsy. *New England Journal of Medicine* 2001; **344**(7): 495-500.

14. Afdhal N, Zeuzem S, Kwo P, et al. Ledipasvir and sofosbuvir for untreated HCV genotype 1 infection. *N Engl J Med* 2014; **370**(20): 1889-98.

15. Afdhal N, Reddy KR, Nelson DR, et al. Ledipasvir and sofosbuvir for previously treated HCV genotype 1 infection. *N Engl J Med* 2014; **370**(16): 1483-93.

16. Kowdley KV, Gordon SC, Reddy KR, et al. Ledipasvir and sofosbuvir for 8 or 12 weeks for chronic HCV without cirrhosis. *N Engl J Med* 2014; **370**(20): 1879-88.

17. Tsochatzis EA, Crossan C, Longworth L, et al. Cost-effectiveness of noninvasive liver fibrosis tests for treatment decisions in patients with chronic hepatitis C. *Hepatology* 2014; **60**(3): 832-43.

18. Sanyal AJ, Chalasani N, Kowdley KV, et al. Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. *N Engl J Med* 2010; **362**(18): 1675-85.

19. Neuschwander-Tetri BA, Loomba R, Sanyal AJ, et al. Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. *Lancet* 2014.

20. European Association for the Study of the LiverAsociacion Latinoamericana para el Estudio del H. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. *Journal of Hepatology* 2015; **63**(1): 237-64.

21. de Franchis R, Baveno VIF. Expanding consensus in portal hypertension: Report of the Baveno VI Consensus Workshop: Stratifying risk and individualizing care for portal hypertension. *J Hepatol* 2015; **63**(3): 743-52.

22. Sandrin L, Fourquet B, Hasquenoph JM, et al. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. *Ultrasound Med Biol* 2003; **29**(12): 1705-13.

23. Nightingale K, Soo MS, Nightingale R, Trahey G. Acoustic radiation force impulse imaging: in vivo demonstration of clinical feasibility. *Ultrasound Med Biol* 2002; **28**(2): 227-35.

24. Bercoff J, Tanter M, Fink M. Supersonic shear imaging: a new technique for soft tissue elasticity mapping. *IEEE Trans Ultrason Ferroelectr Freq Control* 2004; **51**(4): 396-409.

25. Chen S, Sanchez W, Callstrom MR, et al. Assessment of liver viscoelasticity by using shear waves induced by ultrasound radiation force. *Radiology* 2013; **266**(3): 964-70.

26. McPherson S, Stewart SF, Henderson E, Burt AD, Day CP. Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. *Gut* 2010; **59**(9): 1265-9.

27. 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**(9): 1245-51.

28. McHutchison J, Goodman Z, Patel K, et al. Farglitazar Lacks Antifibrotic Activity in Patients With Chronic Hepatitis C Infection. *Gastroenterology* 2010; **138**(4): 1365-73.e2.

29. Fontana RJ, Dienstag JL, Bonkovsky HL, et al. Serum fibrosis markers are associated with liver disease progression in non-responder patients with chronic hepatitis C. *Gut* 2010; **59**(10): 1401-9.

30. Raptis DA, Fischer MA, Graf R, et al. MRI: the new reference standard in quantifying hepatic steatosis? *Gut* 2012; **61**(1): 117-27.

31. Roldan-Valadez E, Favila R, Martinez-Lopez M, Uribe M, Rios C, Mendez-Sanchez N. In vivo 3T spectroscopic quantification of liver fat content in nonalcoholic fatty liver disease: Correlation with biochemical method and morphometry. *J Hepatol* 2010; **53**(4): 732-7.

32. Le TA, Chen J, Changchien C, et al. Effect of colesevelam on liver fat quantified by magnetic resonance in nonalcoholic steatohepatitis: A randomized controlled trial. *Hepatology* 2012; **56**(3): 922-32.

33. Loomba R, Sirlin CB, Ang B, et al. Ezetimibe for the treatment of nonalcoholic steatohepatitis: assessment by novel magnetic resonance imaging and magnetic resonance elastography in a randomized trial (MOZART trial). *Hepatology* 2015; **61**(4): 1239-50.

34. St Pierre TG, Clark PR, Chua-anusorn W, et al. Noninvasive measurement and imaging of liver iron concentrations using proton magnetic resonance. *Blood* 2005; **105**(2): 855-61.

35. Gandon Y, Olivie D, Guyader D, et al. Non-invasive assessment of hepatic iron stores by MRI. *Lancet* 2004; **363**(9406): 357-62.

36. EASL clinical practice guidelines for HFE hemochromatosis. *J Hepatol* 2010; **53**(1): 3-22.

37. Angelucci E, Barosi G, Camaschella C, et al. Italian Society of Hematology practice guidelines for the management of iron overload in thalassemia major and related disorders. *Haematologica* 2008; **93**(5): 741-52.

38. Brittenham GM. Iron-Chelating Therapy for Transfusional Iron Overload. *New England Journal of Medicine* 2011; **364**(2): 146-56.

39. Huwart L, Sempoux C, Vicaut E, et al. Magnetic resonance elastography for the noninvasive staging of liver fibrosis. *Gastroenterology* 2008; **135**(1): 32-40.

40. Venkatesh SK, Wang G, Lim SG, Wee A. Magnetic resonance elastography for the detection and staging of liver fibrosis in chronic hepatitis B. *Eur Radiol* 2014; **24**(1): 70-8.

41. Ichikawa S, Motosugi U, Ichikawa T, et al. Magnetic resonance elastography for staging liver fibrosis in chronic hepatitis C. *Magn Reson Med Sci* 2012; **11**(4): 291-7.

42. Venkatesh SK, Xu S, Tai D, Yu H, Wee A. Correlation of MR elastography with morphometric quantification of liver fibrosis (Fibro-C-Index) in chronic hepatitis B.

Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 2014; **72**(4): 1123-9.

43. 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**(3): 440-51 e6.

44. Su LN, Guo SL, Li BX, Yang P. Diagnostic value of magnetic resonance elastography for detecting and staging of hepatic fibrosis: a meta-analysis. *Clinical radiology* 2014; **69**(12): e545-52.

45. 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* 2014; **60**(6): 1920-8.

46. Kim D, Kim WR, Talwalkar JA, Kim HJ, Ehman RL. Advanced Fibrosis in Nonalcoholic Fatty Liver Disease: Noninvasive Assessment with MR Elastography. *Radiology* 2013; **268**(2): 411-9.

47. Salameh N, Larrat B, Abarca-Quinones J, et al. Early detection of steatohepatitis in fatty rat liver by using MR elastography. *Radiology* 2009; **253**(1): 90-7.

48. Chen J, Talwalkar JA, Yin M, Glaser KJ, Sanderson SO, Ehman RL. Early detection of nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease by using MR elastography. *Radiology* 2011; **259**(3): 749-56.

49. Asrani SK, Talwalkar JA, Kamath PS, et al. Role of magnetic resonance elastography in compensated and decompensated liver disease. *J Hepatol* 2014; **60**(5): 934-9.

50. Nedredal GI, Yin M, McKenzie T, et al. Portal hypertension correlates with splenic stiffness as measured with MR elastography. *Journal of magnetic resonance imaging : JMRI* 2011; **34**(1): 79-87.

51. Huang SY, Abdelsalam ME, Harmoush S, et al. Evaluation of liver fibrosis and hepatic venous pressure gradient with MR elastography in a novel swine model of cirrhosis. *Journal of magnetic resonance imaging : JMRI* 2014; **39**(3): 590-7.

52. Sun HY, Lee JM, Han JK, Choi BI. Usefulness of MR elastography for predicting esophageal varices in cirrhotic patients. *Journal of magnetic resonance imaging : JMRI* 2014; **39**(3): 559-66.

53. Shin SU, Lee JM, Yu MH, et al. Prediction of Esophageal Varices in Patients with Cirrhosis: Usefulness of Three-dimensional MR Elastography with Echo-planar Imaging Technique. *Radiology* 2014: 130916.

54. Dyvorne HA, Jajamovich GH, Besa C, Cooper N, Taouli B. Simultaneous measurement of hepatic and splenic stiffness using MR elastography: preliminary experience. *Abdom Imaging* 2015; **40**(4): 803-9.

55. Bakan AA, Inci E, Bakan S, Gokturk S, Cimilli T. Utility of diffusion-weighted imaging in the evaluation of liver fibrosis. *Eur Radiol* 2012; **22**(3): 682-7.

56. Wang QB, Zhu H, Liu HL, Zhang B. Performance of magnetic resonance elastography and diffusion-weighted imaging for the staging of hepatic fibrosis: A meta-analysis. *Hepatology* 2012; **56**(1): 239-47.

57. Luciani A, Vignaud A, Cavet M, et al. Liver cirrhosis: intravoxel incoherent motion MR imaging--pilot study. *Radiology* 2008; **249**(3): 891-9.

58. Leitao HS, Doblas S, d'Assignies G, et al. Fat deposition decreases diffusion parameters at MRI: a study in phantoms and patients with liver steatosis. *Eur Radiol* 2013; **23**(2): 461-7.

59. Cassinotto C, Feldis M, Vergniol J, et al. MR relaxometry in chronic liver diseases: Comparison of T1 mapping, T2 mapping, and diffusion-weighted imaging for assessing cirrhosis diagnosis and severity. *Eur J Radiol* 2015; **84**(8): 1459-65.

60. Lewin M, Poujol-Robert A, Boelle PY, et al. Diffusion-weighted magnetic resonance imaging for the assessment of fibrosis in chronic hepatitis C. *Hepatology* 2007; **46**(3): 658-65.

61. Bloch F, Hansen WW, Packard M. The Nuclear Induction Experiment. *Physical Review* 1946; **70**(7-8): 474-85.

62. Purcell EM, Torrey HC, Pound RV. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Physical Review* 1946; **69**(1-2): 37-8.

63. Mansfield P, Maudsley A. Medical imaging by NMR. *The British journal of radiology* 1977; **50**(591): 188-94.

64. Mansfield P, Pykett I, Morris P, Coupland R. Human whole body line-scan imaging by NMR. *The British journal of radiology* 1978; **51**(611): 921-2.

65. Smith FW, Mallard JR, Reid A, Hutchison JM. Nuclear magnetic resonance tomographic imaging in liver disease. *Lancet* 1981; **1**(8227): 963-6.

66. Doyle F, Pennock J, Banks L, et al. Nuclear magnetic resonance imaging of the liver: initial experience. *American Journal of Roentgenology* 1982; **138**(2): 193-200.

67. Magnetic resonance imaging of parenchymal liver disease: a comparison with ultrasound, radionuclide scintigraphy and X-ray computed tomography. The Clinical NMR Group. *Clinical radiology* 1987; **38**(5): 495-502.

68. Thomsen C, Christoffersen P, Henriksen O, Juhl E. Prolonged T1 in patients with liver cirrhosis: an in vivo MRI study. *Magnetic resonance imaging* 1990; **8**(5): 599-604.

69. Keevil SF, Alstead EM, Dolke G, Brooks AP, Armstrong P, Farthing MJ. 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**(803): 1083-7.

70. Heye T, Yang SR, Bock M, et al. MR relaxometry of the liver: significant elevation of T1 relaxation time in patients with liver cirrhosis. *Eur Radiol* 2012; **22**(6): 1224-32.

71. Kim KA, Park M-S, Kim I-S, et al. Quantitative evaluation of liver cirrhosis using T1 relaxation time with 3 tesla MRI before and after oxygen inhalation. *Journal of Magnetic Resonance Imaging* 2012; **36**(2): 405-10.

72. 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**(6): 706-14.

73. Goldberg HI, Moss AA, Stark DD, McKerrow J, Engelstad B, Brito A. Hepatic cirrhosis: magnetic resonance imaging. *Radiology* 1984; **153**(3): 737-9.

74. Chamuleau RAFM, De Nie JHNCI, Moerland MA, Van der Lende OR, Smidt J. Is the magnetic resonance imaging proton spin-lattice relaxation time a reliable noninvasive parameter of developing liver fibrosis? *Hepatology* 1988; **8**(2): 217-21.

75. Aisen AM, Doi K, Swanson SD. Detection of liver fibrosis with magnetic cross-relaxation. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 1994; **31**(5): 551-6.

76. William MCR, Michael V, Robert T, et al. Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology* 2004; **127**(6): 1704-13.

77. Wang Y-XJX, Yuan J, Chu ES, et al. T1rho MR imaging is sensitive to evaluate liver fibrosis: an experimental study in a rat biliary duct ligation model. *Radiology* 2011; **259**(3): 712-9.

78. Deng M, Zhao F, Yuan J, Ahuja AT, Wang YXJ. Liver T1ρ MRI measurement in healthy human subjects at 3 T: a preliminary study with a two-dimensional fast-field echo sequence. *The British journal of radiology* 2012.

79. Thomas A, Florian S, Vito C, et al. Evaluation of fibrotic liver disease with whole-liver T1ρ MR imaging: a feasibility study at 1.5 T. *Radiology* 2013.

80. Rauscher I, Eiber M, Ganter C, et al. Evaluation of T1ρ as a potential MR biomarker for liver cirrhosis: comparison of healthy control subjects and patients with liver cirrhosis. *European journal of radiology* 2014; **83**(6): 900-4.

81. Yukihisa T, Akihiro N, Yoshiki A, et al. T1ρ Relaxation of the liver: A potential biomarker of liver function. *Journal of Magnetic Resonance Imaging* 2015.

82. Zhao X, Huang M, Zhu Q, Wang T, Liu Q. The relationship between liver function and liver parenchymal contrast enhancement on Gd-BOPTA-enhanced MR imaging in the hepatocyte phase. *Magnetic resonance imaging* 2015; **33**(6): 768-73.

83. Xie S, Sun Y, Wang L, Yang Z, Luo J, Wang W. Assessment of liver function and liver fibrosis with dynamic Gd-EOB-DTPA-enhanced MRI. *Academic radiology* 2015; **22**(4): 460-6.

84. Michael H, Mona S, Niklas V, et al. Volume-assisted estimation of liver function based on Gd-EOB-DTPA–enhanced MR relaxometry. *European Radiology* 2015.

85. Pugh RNH, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. *British Journal of Surgery* 1973; **60**(8): 646-9.

86. Kamath PS, Wiesner RH, Malinchoc M, et al. A model to predict survival in patients with end-stage liver disease. *Hepatology* 2001; **33**(2): 464-70.

87. Haimerl M, Verloh N, Zeman F, et al. Assessment of clinical signs of liver cirrhosis using T1 mapping on Gd-EOB-DTPA-enhanced 3T MRI. *PloS one* 2013; **8**(12): e85658.

88. Haimerl M, Verloh N, Fellner C, et al. MRI-based estimation of liver function: Gd-EOB-DTPA-enhanced T1 relaxometry of 3T vs. the MELD score. *Scientific reports* 2014; **4**: 5621.

89. Verloh N, Haimerl M, Zeman F, et al. Assessing liver function by liver enhancement during the hepatobiliary phase with Gd-EOB-DTPA-enhanced MRI at 3 Tesla. *Eur Radiol* 2014; **24**(5): 1013-9.

90. Nilsson H, Blomqvist L, Douglas L, et al. Gd-EOB-DTPA-enhanced MRI for the assessment of liver function and volume in liver cirrhosis. *The British journal of radiology* 2013; **86**(1026): 20120653.

91. Kanki A, Tamada T, Higaki A, et al. Hepatic parenchymal enhancement at Gd-EOB-DTPA-enhanced MR imaging: correlation with morphological grading of severity in cirrhosis and chronic hepatitis. *Magnetic resonance imaging* 2012; **30**(3): 356-60.

92. Nilsson H, Blomqvist L, Douglas L, Nordell A, Jonas E. Assessment of liver function in primary biliary cirrhosis using Gd-EOB-DTPA-enhanced liver MRI. *HPB (Oxford)* 2010; **12**(8): 567-76.

93. Nilsson H, Blomqvist L, Douglas L, et al. Dynamic gadoxetate-enhanced MRI for the assessment of total and segmental liver function and volume in primary sclerosing cholangitis. *Journal of magnetic resonance imaging : JMRI* 2014; **39**(4): 879-86.

94. Ding Y, Rao SX, Meng T, Chen C, Li R, Zeng MS. Usefulness of T1 mapping on Gd-EOB-DTPA-enhanced MR imaging in assessment of non-alcoholic fatty liver disease. *Eur Radiol* 2014; **24**(4): 959-66.

95. Bastati N, Feier D, Wibmer A, et al. Noninvasive Differentiation of Simple Steatosis and Steatohepatitis by Using Gadoxetic Acid-enhanced MR Imaging in Patients with Nonalcoholic Fatty Liver Disease: A Proof-of-Concept Study. *Radiology* 2014; **271**(3): 739-47.

96. Ding Y, Rao SXX, Zhu T, Chen CZZ, Li RCC, Zeng MSS. Liver fibrosis staging using T1 mapping on gadoxetic acid-enhanced MRI compared with DW imaging. *Clinical radiology* 2015.

97. Feier D, Balassy C, Bastati N, Stift J, Badea R, Ba-Ssalamah A. Liver fibrosis: histopathologic and biochemical influences on diagnostic efficacy of hepatobiliary contrast-enhanced MR imaging in staging. *Radiology* 2013; **269**(2): 460-8.

98. Masahiro O, Takamichi M, Norihisa Y, et al. Comparison between T1 relaxation time of Gd-EOB-DTPA-enhanced MRI and liver stiffness measurement of ultrasound elastography in the evaluation of cirrhotic liver. *Journal of Magnetic Resonance Imaging* 2015.

99. Nishie A, Asayama Y, Ishigami K, et al. MR prediction of liver fibrosis using a liver-specific contrast agent: Superparamagnetic iron oxide versus Gd-EOB-DTPA. *Journal of Magnetic Resonance Imaging* 2012; **36**(3): 664-71.

100. Nishie A, Ushijima Y, Tajima T, et al. Quantitative analysis of liver function using superparamagnetic iron oxide- and Gd-EOB-DTPA-enhanced MRI: Comparison with

Technetium-99m galactosyl serum albumin scintigraphy. *European Journal of Radiology* 2012; **81**(6): 1100-4.

101. Polasek M, Fuchs BC, Uppal R, et al. Molecular MR imaging of liver fibrosis: a feasibility study using rat and mouse models. *J Hepatol* 2012; **57**(3): 549-55.

102. Fuchs BC, Wang H, Yang Y, et al. Molecular MRI of collagen to diagnose and stage liver fibrosis. *J Hepatol* 2013; **59**(5): 992-8.

103. Farrar CT, DePeralta DK, Day H, et al. 3D molecular MR imaging of liver fibrosis and response to rapamycin therapy in a bile duct ligation rat model. *J Hepatol* 2015.

104. Iwamoto DV, Calderwood DA. Regulation of integrin-mediated adhesions. *Current opinion in cell biology* 2015; **36**: 41-7.

105. Zhou X, Murphy FR, Gehdu N, Zhang J, Iredale JP, Benyon RC. Engagement of alphavbeta3 integrin regulates proliferation and apoptosis of hepatic stellate cells. *The Journal of biological chemistry* 2004; **279**(23): 23996-4006.

106. Neubauer K, Knittel T, Armbrust T, Ramadori G. Accumulation and cellular localization of fibrinogen/fibrin during short-term and long-term rat liver injury. *Gastroenterology* 1995; **108**(4): 1124-35.

107. Chow AM, Tan M, Gao DS, et al. Molecular MRI of liver fibrosis by a peptide-targeted contrast agent in an experimental mouse model. *Investigative radiology* 2013; **48**(1): 46-54.

108. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology 2005 Jun;41(6):1313-21*.

109. Pellicoro A, Aucott RL, Ramachandran P, et al. Elastin accumulation is regulated at the level of degradation by macrophage metalloelastase (MMP 12) during experimental liver fibrosis. *Hepatology* 2012; **55**(6): 1965-75.

110. Feier D, Balassy C, Bastati N, Fragner R, Wrba F, Ba-Ssalamah A. The diagnostic efficacy of quantitative liver MR imaging with diffusion-weighted, SWI, and hepato-specific contrast-enhanced sequences in staging liver fibrosis-a multiparametric approach. *European radiology* 2015.

111. House MJ, Bangma SJ, Thomas M, et al. Texture-based classification of liver fibrosis using MRI. *Journal of magnetic resonance imaging : JMRI* 2015; **41**(2): 322-8.

112. Wu Z, Matsui O, Kitao A, et al. Hepatitis C related chronic liver cirrhosis: feasibility of texture analysis of MR images for classification of fibrosis stage and necroinflammatory activity grade. *PloS one* 2015; **10**(3): e0118297.

113. Tunnicliffe E, Robson M, Banerjee R. Medical Imaging. Application number PCT/GB2014/050815; 2014.

114. Banerjee R, Pavlides M, Tunnicliffe EM, et al. Multiparametric magnetic resonance for the non-invasive diagnosis of liver disease. *J Hepatol* 2014; **60**(1): 69-77.

115. 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**(4): 1236-44.

116. Piechnik SK, Ferreira VM, Dall'Armellina E, et al. Shortened Modified Look-Locker Inversion recovery (ShMOLLI) for clinical myocardial T1-mapping at 1.5 and 3 T within a 9 heartbeat breathhold. *J Cardiovasc Magn Reson* 2010; **12**: 69.

117. Ferreira VM, Piechnik SK, Dall'Armellina E, et al. Native T1-mapping detects the location, extent and patterns of acute myocarditis without the need for gadolinium contrast agents. *J Cardiovasc Magn Reson* 2014; **16**: 36.

118. Ntusi NA, Piechnik SK, Francis JM, et al. Diffuse Myocardial Fibrosis and Inflammation in Rheumatoid Arthritis: Insights From CMR T1 Mapping. *JACC Cardiovascular imaging* 2015; **8**(5): 526-36.

119. Ganne-Carrie N, Christidis C, Chastang C, et al. Liver iron is predictive of death in alcoholic cirrhosis: a multivariate study of 229 consecutive patients with alcoholic and/or hepatitis C virus cirrhosis: a prospective follow up study. *Gut* 2000; **46**(2): 277-82.

120. Ishak K, Baptista A, Bianchi L, et al. Histological grading and staging of chronic hepatitis. *J Hepatol* 1995; **22**(6): 696-9.

121. DeWitt LD, Kistler JP, Miller DC, Richardson EP, Jr., Buonanno FS. NMR-neuropathologic correlation in stroke. *Stroke* 1987; **18**(2): 342-51.

122. Conlon P, Trimble MR, Rogers D, Callicott C. Magnetic resonance imaging in epilepsy: a controlled study. *Epilepsy Res* 1988; **2**(1): 37-43.

123. Stikov N, Keenan KE, Pauly JM, Smith RL, Dougherty RF, Gold GE. Cross-relaxation imaging of human articular cartilage. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 2011; **66**(3): 725-34.

124. de Miguel MH, Yeung HN, Goyal M, et al. Evaluation of quantitative magnetic resonance imaging as a noninvasive technique for measuring renal scarring in a rabbit model of antiglomerular basement membrane disease. *J Am Soc Nephrol* 1994; **4**(11): 1861-8.

125. Pitkin J, Rees MC, Gray S, et al. Management of premature menopause. *Menopause Int* 2007; **13**(1): 44-5.

126. Jauregui J, Nelson D, Choo E, et al. The BUDDY (Bedside Ultrasound to Detect Dehydration in Youth) study. *Crit Ultrasound J* 2014; **6**(1): 15.

127. Hanneman K, Thavendiranathan P, Nguyen ET, et al. Use of Cardiac Magnetic Resonance Imaging Based Measurements of Inferior Vena Cava Cross-Sectional Area in the Diagnosis of Pericardial Constriction. *Can Assoc Radiol J* 2015.

128. Wolak A, Gransar H, Thomson LE, et al. Aortic size assessment by noncontrast cardiac computed tomography: normal limits by age, gender, and body surface area. *JACC Cardiovascular imaging* 2008; **1**(2): 200-9.

129. Chapman AB, Zamudio S, Woodmansee W, et al. Systemic and renal hemodynamic changes in the luteal phase of the menstrual cycle mimic early pregnancy. *Am J Physiol* 1997; **273**(5 Pt 2): F777-82.

130. Stachenfeld NS, Silva C, Keefe DL, Kokoszka CA, Nadel ER. Effects of oral contraceptives on body fluid regulation. *J Appl Physiol (1985)* 1999; **87**(3): 1016-25.

131. Taylor JW. The timing of menstruation-related symptoms assessed by a daily symptom rating scale. *Acta Psychiatr Scand* 1979; **60**(1): 87-105.

132. Meaden PM, Hartlage SA, Cook-Karr J. Timing and severity of symptoms associated with the menstrual cycle in a community-based sample in the Midwestern United States. *Psychiatry Res* 2005; **134**(1): 27-36.

133. White CP, Hitchcock CL, Vigna YM, Prior JC. Fluid Retention over the Menstrual Cycle: 1-Year Data from the Prospective Ovulation Cohort. *Obstet Gynecol Int* 2011; **2011**: 138451.

134. Richards MA, Webb JA, Jewell SE, Gregory WM, Reznek RH. In-vivo measurement of spin lattice relaxation time (T1) of liver in healthy volunteers: the effects of age, sex and oral contraceptive usage. *The British journal of radiology* 1988; **61**(721): 34-7.

135. Piechnik SK, Ferreira VM, Lewandowski AJ, et al. Normal variation of magnetic resonance T1 relaxation times in the human population at 1.5 T using ShMOLLI. *J Cardiovasc Magn Reson*; **15**: 13.

136. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology* 1996; **24**(2): 289-93.

137. Goodman ZD, Becker RL, Jr., Pockros PJ, Afdhal NH. Progression of fibrosis in advanced chronic hepatitis C: evaluation by morphometric image analysis. *Hepatology* 2007; **45**(4): 886-94.

138. Huang Y, de Boer WB, Adams LA, Macquillan G, Bulsara MK, Jeffrey GP. Image analysis of liver biopsy samples measures fibrosis and predicts clinical outcome. *J Hepatol* 2014.

139. Tsochatzis E, Bruno S, Isgro G, et al. Collagen proportionate area is superior to other histological methods for sub-classifying cirrhosis and determining prognosis. *J Hepatol* 2014; **60**(5): 948-54.

140. Scheuer PJ, Williams R, Muir AR. Hepatic pathology in relatives of patients with haemochromatosis. *The Journal of pathology and bacteriology* 1962; **84**: 53-64.

141. Hall AR, Tsochatzis E, Morris R, Burroughs AK, Dhillon AP. Sample size requirement for digital image analysis of collagen proportionate area in cirrhotic livers. *Histopathology* 2013; **62**(3): 421-30.

142. Isgro G, Calvaruso V, Andreana L, et al. The relationship between transient elastography and histological collagen proportionate area for assessing fibrosis in chronic viral hepatitis. *J Gastroenterol* 2013; **48**(8): 921-9.

143. 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**(7): 1227-32.

144. Ryan MC, Wilson AM, Slavin J, Best JD, Jenkins AJ, Desmond PV. Associations Between Liver Histology and Severity of the Metabolic Syndrome in Subjects With Nonalcoholic Fatty Liver Disease. *Diabetes Care* 2005; **28**(5): 1222-4.

145. Loomba R, Abraham M, Unalp A, et al. Association between diabetes, family history of diabetes, and risk of nonalcoholic steatohepatitis and fibrosis. *Hepatology* 2012; **56**(3): 943-51.

146. Williams CD, Stengel J, Asike MI, et al. Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. *Gastroenterology* 2011; **140**(1): 124-31.

147. Hyysalo J, Mannisto VT, Zhou Y, et al. A population-based study on the prevalence of NASH using scores validated against liver histology. *J Hepatol* 2014; **60**(4): 839-46.

148. Wong RJ, Aguilar M, Cheung R, et al. Nonalcoholic steatohepatitis is the second leading etiology of liver disease among adults awaiting liver transplantation in the United States. *Gastroenterology* 2015; **148**(3): 547-55.

149. Ekstedt M, Franzen LE, Mathiesen UL, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. *Hepatology* 2006; **44**(4): 865-73.

150. 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* 2014.

151. Argo CK, Northup PG, Al-Osaimi AM, Caldwell SH. Systematic review of risk factors for fibrosis progression in non-alcoholic steatohepatitis. *J Hepatol* 2009; **51**(2): 371-9.

152. 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**(1): 113-21.

153. Pais R, Charlotte F, Fedchuk L, et al. A systematic review of follow-up biopsies reveals disease progression in patients with non-alcoholic fatty liver. *J Hepatol*; **59**(3): 550-6.

154. McPherson S, Hardy T, Henderson E, Burt AD, Day CP, Anstee QM. Evidence of NAFLD progression from steatosis to fibrosing-steatohepatitis using paired biopsies: implications for prognosis and clinical management. *J Hepatol* 2015; **62**(5): 1148-55.

155. Rawlins SR, Mullen CM, Simon HM, et al. Wedge and needle liver biopsies show discordant histopathology in morbidly obese patients undergoing Roux-en-Y gastric bypass surgery. *Gastroenterology report* 2013; **1**(1): 51-7.

156. Kwok R, Tse YK, Wong GL, et al. Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease--the role of transient elastography and plasma cytokeratin-18 fragments. *Aliment Pharmacol Ther* 2014; **39**(3): 254-69.

157. Cusi K, Chang Z, Harrison S, et al. Limited value of plasma cytokeratin-18 as a biomarker for NASH and fibrosis in patients with non-alcoholic fatty liver disease. *J Hepatol* 2014; **60**(1): 167-74.

158. Castera L, Foucher J, Bernard PH, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. *Hepatology* 2010; **51**(3): 828-35.

159. Wong VW, Vergniol J, Wong GL, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. *The American journal of gastroenterology* 2012; **107**(12): 1862-71.

160. Bedossa P. Utility and appropriateness of the fatty liver inhibition of progression (FLIP) algorithm and steatosis, activity, and fibrosis (SAF) score in the evaluation of biopsies of nonalcoholic fatty liver disease. *Hepatology* 2014; **60**(2): 565-75.

161. Sanyal AJ, Abdelmalek MF, Suzuki A, Cummings OW, Chojkier M. No significant effects of ethyl-eicosapentanoic acid on histologic features of nonalcoholic steatohepatitis in a phase 2 trial. *Gastroenterology* 2014; **147**(2): 377-84.e1.

162. Pierre B, Delphine D, Valerie P. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 2003; **38**(6): 1449-57.

163. 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 UK, 2014.

164. Bosch J, Garcia-Pagan JC. Complications of cirrhosis. I. Portal hypertension. *J Hepatol* 2000; **32**(1 Suppl): 141-56.

165. Patch D, Armonis A, Sabin C, et al. Single portal pressure measurement predicts survival in cirrhotic patients with recent bleeding. *Gut* 1999; **44**(2): 264-9.

166. Ripoll C, Groszmann R, Garcia-Tsao G, et al. Hepatic venous pressure gradient predicts clinical decompensation in patients with compensated cirrhosis. *Gastroenterology* 2007; **133**(2): 481-8.

167. Vorobioff J, Groszmann RJ, Picabea E, et al. Prognostic value of hepatic venous pressure gradient measurements in alcoholic cirrhosis: a 10-year prospective study. *Gastroenterology* 1996; **111**(3): 701-9.

168. Escorsell À, Bru C, Gilabert R, Moitinho E, García-Pagán JC, Bosch J. Wedged hepatic venous pressure adequately reflects portal pressure in hepatitis C virus–related cirrhosis. *Hepatology* 1999; **30**(6): 1393-7.

169. Bosch J, Abraldes JG, Berzigotti A, Garcia-Pagan JC. The clinical use of HVPG measurements in chronic liver disease. *Nat Rev Gastroenterol Hepatol* 2009; **6**(10): 573-82.

170. Thalheimer U, Bellis L, Puoti C, Burroughs AK. Should we routinely measure portal pressure in patients with cirrhosis, using hepatic venous pressure gradient (HVPG) as a guide for prophylaxis and therapy of bleeding and rebleeding? No. *Eur J Intern Med* 2011; **22**(1): 5-7.

171. Colecchia A, Montrone L, Scaioli E, et al. Measurement of Spleen Stiffness to Evaluate Portal Hypertension and the Presence of Esophageal Varices in Patients With HCV-Related Cirrhosis. *Gastroenterology* 2012; **143**(3): 646-54.

172. Perazzo H, Fernandes FF, Castro Filho EC, Perez RM. Points to be considered when using transient elastography for diagnosis of portal hypertension according to the Baveno's VI consensus. *J Hepatol* 2015.

173. Perazzo H, Fernandes FF, Gomes A, Terra C, Perez RM, Figueiredo FAF. Interobserver variability in transient elastography analysis of patients with chronic hepatitis C. *Liver International* 2015; **35**(5): 1533-9.

174. Nascimbeni F, Lebray P, Fedchuk L, et al. Significant variations in elastometry measurements made within short-term in patients with chronic liver diseases. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association* 2015; **13**(4): 763-71 e1-6.

175. Williams AL, Hoofnagle JH. Ratio of serum aspartate to alanine aminotransferase in chronic hepatitis. Relationship to cirrhosis. *Gastroenterology* 1988; **95**(3): 734-9.

176. 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**(2): 518-26.

177. Vallet-Pichard A, Mallet V, Nalpas B, et al. FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. *Hepatology* 2007; **46**(1): 32-6.

178. Gibson PR, Gibson RN, Ditchfield MR, Donlan JD. Splenomegaly—an insensitive sign of portal hypertension. *Australian and New Zealand Journal of Medicine* 1990; **20**(6): 771-4.

179. Berzigotti A, Seijo S, Arena U, et al. Elastography, Spleen Size, and Platelet Count Identify Portal Hypertension in Patients With Compensated Cirrhosis. *Gastroenterology* 2013; **144**(1): 102-11.e1.

180. Klasen J, Lanzman RS, Wittsack HJ, et al. Diffusion-weighted imaging (DWI) of the spleen in patients with liver cirrhosis and portal hypertension. *Magnetic resonance imaging* 2013; **31**(7): 1092-6.

181. Zhang J, Tao R, You Z, et al. Gamna-Gandy bodies of the spleen detected with susceptibility weighted imaging: maybe a new potential non-invasive marker of esophageal varices. *PloS one* 2013; **8**(1): e55626.

182. Ouyang HQ, Gong ZJ, Zha YF, Liu CS, Yang ZH. Splenic siderotic nodules in patients with liver cirrhosis. *Exp Ther Med* 2013; **6**(2): 445-50.

183. Zhou L, Chen TW, Zhang XM, et al. Spleen dynamic contrast-enhanced magnetic resonance imaging as a new method for staging liver fibrosis in a piglet model. *PloS one* 2013; **8**(12): e83697.

184. Sharma P, Kirnake V, Tyagi P, et al. Spleen Stiffness in Patients With Cirrhosis in Predicting Esophageal Varices. *The American journal of gastroenterology* 2013; **108**(7): 1101-7.

185. Elkrief L, Rautou PE, Ronot M, et al. Prospective comparison of spleen and liver stiffness by using shear-wave and transient elastography for detection of portal hypertension in cirrhosis. *Radiology* 2015; **275**(2): 589-98.

186. Ronot M, Lambert S, Elkrief L, et al. Assessment of portal hypertension and high-risk oesophageal varices with liver and spleen three-dimensional multifrequency MR elastography in liver cirrhosis. *European Radiology* 2014; **24**(6): 1394-402.

187. Zipprich A, Winkler M, Seufferlein T, Dollinger MM. Comparison of balloon vs. straight catheter for the measurement of portal hypertension. *Aliment Pharmacol Ther* 2010; **32**(11-12): 1351-6.

188. Keiding S, Vilstrup H. Intrahepatic Heterogeneity of Hepatic Venous Pressure Gradient in Human Cirrhosis. *Scandinavian journal of gastroenterology* 2002; **37**(8): 960-4.

189. D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. *J Hepatol* 2006; **44**(1): 217-31.

190. Rosenberg WM, Voelker M, Thiel R, et al. Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology 2004 Dec;127(6):1704-13*.

191. Robic MA, Procopet B, Métivier S, et al. Liver stiffness accurately predicts portal hypertension related complications in patients with chronic liver disease: A prospective study. *Journal of Hepatology* 2011; **55**(5): 1017-24.

192. Torok N, Dranoff JA, Schuppan D, Friedman SL. Strategies and Endpoints of Antifibrotic Drug Trials. *Hepatology* 2015.

193. Mahajan R, Gupta K. Food and drug administration's critical path initiative and innovations in drug development paradigm: Challenges, progress, and controversies. *Journal of pharmacy & bioallied sciences* 2010; **2**(4): 307-13.

194. Valenti L, Fracanzani AL, Bugianesi E, et al. HFE Genotype, Parenchymal Iron Accumulation, and Liver Fibrosis in Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology* 2010; **138**(3): 905-12.

195. Lok AS, Everhart JE, Chung RT, et al. Evolution of hepatic steatosis in patients with advanced hepatitis C: results from the hepatitis C antiviral long-term treatment against cirrhosis (HALT-C) trial. *Hepatology* 2009; **49**(6): 1828-37.

196. Castera L, Hezode C, Roudot-Thoraval F, et al. Worsening of steatosis is an independent factor of fibrosis progression in untreated patients with chronic hepatitis C and paired liver biopsies. *Gut* 2003; **52**(2): 288-92.

197. Huang YW, Yang SS, Fu SC, et al. Increased risk of cirrhosis and its decompensation in chronic hepatitis C patients with new-onset diabetes: a nationwide cohort study. *Hepatology* 2014; **60**(3): 807-14.

198. Friedenberg F, Pungpapong S, Zaeri N, Braitman LE. The impact of diabetes and obesity on liver histology in patients with hepatitis C. *Diabetes, Obesity and Metabolism* 2003; **5**(3): 150-5.

199. Wyatt J, Baker H, Prasad P, Gong YY, Millson C. Steatosis and fibrosis in patients with chronic hepatitis C. *Journal of Clinical Pathology* 2004; **57**(4): 402-6.

200. Tsochatzis EA, Manousou P, Dhillon AP, Burroughs AK. Regression of fibrosis: the need for quantitative methods of assessment. *J Hepatol* 2012.

201. Tsochatzis EA, Gurusamy KS, Ntaoula S, Cholongitas E, Davidson BR, Burroughs AK. Elastography for the diagnosis of severity of fibrosis in chronic liver disease: a meta-analysis of diagnostic accuracy. *J Hepatol* 2011; **54**(4): 650-9.

202. Poynard T, Munteanu M, Deckmyn O, et al. Applicability and precautions of use of liver injury biomarker FibroTest. A reappraisal at 7 years of age. *BMC gastroenterology* 2011; **11**: 39.

203. 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**(1): 199-208.

204. Imbert-Bismut F, Messous D, Thibault V, et al. Intra-laboratory analytical variability of biochemical markers of fibrosis (Fibrotest) and activity (Actitest) and reference ranges in healthy blood donors. *Clinical chemistry and laboratory medicine : CCLM / FESCC* 2004; **42**(3): 323-33.

205. Fraquelli M, Rigamonti C, Casazza G, et al. Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. *Gut* 2007; **56**(7): 968-73.

206. Venkatesh S, Yin M, Takahashi N, Glockner J, Talwalkar J, Ehman R. Non-invasive detection of liver fibrosis: MR imaging features vs. MR elastography. *Abdominal Imaging* 2015; **40**(4): 766-75.

207. Shi Y, Guo Q, Xia F, Sun J, Gao Y. Short- and midterm repeatability of magnetic resonance elastography in healthy volunteers at 3.0 T. *Magnetic resonance imaging* 2014; **32**(6): 665-70.

208. Volpi N, Schiller J, Stern R, Soltes L. Role, metabolism, chemical modifications and applications of hyaluronan. *Current medicinal chemistry* 2009; **16**(14): 1718-45.

209. Coco B, Oliveri F, Maina AM, et al. Transient elastography: a new surrogate marker of liver fibrosis influenced by major changes of transaminases. *J Viral Hepat* 2007; **14**(5): 360-9.

210. 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**(1): 65-72.

211. Ichikawa S, Motosugi U, Nakazawa T, et al. Hepatitis activity should be considered a confounder of liver stiffness measured with MR elastography. *Journal of magnetic resonance imaging : JMRI* 2015; **41**(5): 1203-8.

212. Shi Y, Guo Q, Xia F, et al. MR elastography for the assessment of hepatic fibrosis in patients with chronic hepatitis B infection: does histologic necroinflammation influence the measurement of hepatic stiffness? *Radiology* 2014; **273**(1): 88-98.

213. D'Ambrosio R, Aghemo A, Fraquelli M, et al. The diagnostic accuracy of Fibroscan® for cirrhosis is influenced by liver morphometry in HCV patients with a sustained virological response. *Journal of Hepatology* 2013; **59**(2): 251-6.

214. Vergniol J, Foucher J, Terrebonne E, et al. Noninvasive tests for fibrosis and liver stiffness predict 5-year outcomes of patients with chronic hepatitis C. *Gastroenterology* 2011; **140**(7): 1970-9, 9.e1-3.

215. Poynard T, Imbert-Bismut F, Ratziu V, et al. Biochemical markers of liver fibrosis in patients infected by hepatitis C virus: longitudinal validation in a randomized trial. *Journal of Viral Hepatitis* 2002; **9**(2): 128-33.

216. Ogawa E, Furusyo N, Toyoda K, Takeoka H, Maeda S, Hayashi J. The longitudinal quantitative assessment by transient elastography of chronic hepatitis C patients treated with pegylated interferon alpha-2b and ribavirin. *Antiviral research* 2009; **83**(2): 127-34.