Stratification of disease severity in patients with liver fibrosis, cirrhosis and portal hypertension using novel multiparametric magnetic resonance imaging techniques



Arjun Narayan Ajmer Jayaswal

Jesus College

University of Oxford

Michaelmas Term 2019

A thesis submitted for the degree of *Doctor of Philosophy*

This thesis is dedicated to the memory of my late father *Subhash Jayaswal* (1947-2018)

I miss you

Table of Contents

Acknowledgements 8

Declaration 10

Publications 11

Abbreviations 13

Index of tables 15

Index of figures 17

Thesis abstract 20

Chapter 1: Introduction 21

1.1 Chronic liver disease pathophysiology 22

1.2 Chronic liver disease epidemiology 23

1.3 Methods of stratification of disease severity and need for non-invasive tools 24

1.4 Non-invasive diagnosis of fibrosis, cirrhosis and portal hypertension 26

1.4.1 Ultrasound elastography 27

1.4.1 Magnetic resonance techniques 29

1.4.2 Magnetic resonance elastography (MRE) 32

1.4.3 Molecular MR 35

1.4.4 MR assessment of abdominal blood flow for assessment of portal hypertension 36

1.4.5 Diffusion weighted imaging 38

1.4.6 Intravoxel incoherent motion diffusion weighted imaging 40

1.4.7 Arterial spin labelling 40

1.4.8 T_1 relaxometry 41

1.4.9 $T_1\rho$ relaxometry 43

1.4.10 Dynamic contrast-enhanced studies 44

1.4.11 Texture analysis 47

1.4.12 Multiparametric MRI studies 48

1.4.13 Iron-corrected T_1 (cT_1) 49

1.5 Aims 51

Chapter 2: General Theory and Methods 53

2.1 Nuclear Magnetic Resonance theory *54*

2.2 Magnetic Resonance Imaging *62*

2.2.1 T_1 mapping 62

2.2.2 T_2* relaxation 67

2.2.3 Iron-corrected T_1 (cT_1) 67

2.2.4 Image analysis 69

2.2.5 United Kingdom Biobank (UKBB) imaging protocol 72

2.3 Magnetic Resonance Spectroscopy *73*

2.3.1 Theory 73

2.3.2 Acquisition and protocol 74

2.3.4 Analysis 75

2.4 Scanning practicalities *78*

2.5 Transient elastography (TE) *79*

2.6 Blood sampling *80*

2.7 Statistical analysis *80*

Chapter 3. Prognostic ability of liver cT_1 81

3.1 Abstract: *83*

3.2 Introduction *84*

3.3 Aims *85*

3.4 Methods *85*

3.4.1 Patient selection 85

3.4.2 MRI protocol and analysis 87

3.4.3 Histological assessment 87

3.4.4 Transient elastography 88

3.4.5 Clinical follow-up 88

3.4.6 Statistical analysis 88

3.5 Results *91*

3.5.1 Cohort characteristics 91

3.5.2 Survival analysis by biomarker 97

3.6 Discussion: *110*

Chapter 4: Changes in liver cT_1 after direct acting antiviral treatment for chronic hepatitis C virus 115

4.1 Abstract: *117*

4.2 Introduction *118*

4.3 Aims *120*

4.4 Methods *120*

4.4.1 Patient selection 120

4.4.2 Patient follow-up 121

4.4.3 MR protocol 121

4.4.4 MR analysis 122

4.4.5 Blood sampling 122

4.4.6 Transient Elastography 122

4.4.7 Statistical analysis 122

4.5 Results *124*

4.5.1 Patient characteristics 124

4.5.2 Liver cT_1 129

4.5.3 Liver fat 130

4.5.4 Liver Iron 131

4.5.5 Transient Elastography 131

4.5.6 Blood tests and biochemistry 132

4.5.7 Patient follow-up 132

4.5.8 Patients with cirrhosis vs patients without cirrhosis 133

4.5.9 Effects of excess alcohol consumption 133

4.6 Discussion *140*

Chapter 5. T_1 and cT_1 of abdominal organs for the assessment of portal hypertension 147

5.1 Abstract *149*

5.2 Introduction *150*

5.2 Aims *155*

5.3 OCMR cohort *157*

5.3.1 Introduction 157

5.3.2 Aims 157

5.3.3 Methods 157

5.3.4 Results 160

5.3.5 Discussion 172

5.4 Singapore General Hospital cohort *177*

5.4.1 Introduction 177

5.4.2 Aims 177

5.4.3 Methods 178
5.4.4 Results 182
5.4.5 Discussion 193

Chapter 6. T_1 and cT_1 of abdominal organs for the assessment of portal hypertension in a paediatric population of autoimmune liver disease 198

The work in this Chapter has been accepted as a paper in: 199

6.1 Abstract *200*

6.2 Introduction *201*

6.3 Aims *202*

6.4 Methods *203*
Patient selection 203
MR protocol 204
Statistical analysis 204

6.5 Results *205*
Spleen T_1/cT_1 in patients with and without PH 206
Spleen T_1/$cT_1$for classification of PH 206
Spleen T_1/cT_1 for classification of varices 206
Liver and spleen cT_1 by aetiology 207
Logistic regression of liver and spleen MR 207
Spleen T_2* values 208

6.6 Discussion *211*

Chapter 7. T_1 and cT_1 of abdominal organs' summary for assessment of portal hypertension 216

6.1 Summary *217*

Chapter 8. Spleen T_1 and cT_1 in healthy volunteers and the UK Biobank imaging study 224

8.1 Abstract: *225*

8.2 Introduction *226*

8.3 Aims *228*

8.4 Spleen T_1 and cT_1 in the UKBB *229*
8.4.1 Methods 229
8.4.2 Results 231

8.5 Spleen T_1 and cT_1 in healthy volunteers in OCMR *242*

8.5.1 Methods 242
8.5.2 Results 243
8.6 Comparison of healthy groups with patients at risk of PH 246
8.6.1 Aims 246
8.6.2 Methods 247
8.6.3 Results 248
8.7 Discussion 255
8.7.1 Main findings 255
8.7.2 UKBB and OCMR healthy volunteers' exploratory analysis 255
8.7.3 Healthy datasets comparison to patients with CLD 261
Chapter 9: Conclusions and suggested further work 266
9.1 Summary of main findings 267
9.2 Liver cT_1 prediction of clinical outcomes 267
9.3 Liver cT_1 for monitoring of disease progression/regression 268
9.4 Spleen T_1 and cT_1 for assessment of portal hypertension 269
9.5 Advancement in MR techniques 271
9.6 Practical Utility 272
References 273

Acknowledgements

Firstly, thanks to those who encouraged and supported my decision to completely change career direction, leave my job and start this DPhil. Thanks also to my former work colleagues and friends who gave me a great send-off before coming here to Oxford.

I would like to thank my academic supervisors Prof Stefan Neubauer and Prof Ellie Barnes for their guidance throughout this DPhil. I would also like to thank Dr Matt Kelly (my industry supervisor) and Dr Michael Pavlides, who have been my first points of contact at Perspectum Diagnostics and OCMR respectively. They both have given me a huge amount of help and a lot of their time – especially in reading through first drafts of various pieces of work.

From OCMR I would like to thank the following people:

Liver MRI group: Ferenc Mozes, Christina Levick, Emmanuel Selvaraj, Dragana Savic, Liam Young, Liz Tunnicliffe, Charles Hill and Gabriela Belsley.

Radiographers and nurses: Joana Leal, Jane Francis, Fiona McMillan, Raj Soundarajan, Cassandra Baker, Rebecca Mills, Carla Cascone and Roser Farras-Arraya.

Ethics: Hanan Lamlum, Katja Pfafferot and Philippa Kemp.

IT: Lukasz Owsiak, Vadim Benelov, Mike MacDonald and Andrew Rudgewick-Brown.

Patient coordination/reception: Karen Whatley and Ella May.

Miscellaneous: Kathryn Lacey, Aaron Hess, Adrienne Siu, Andrew Apps, Sven Jaeschke, William Watson, Matthew Burrage, Moritz Hundertmark, Marco Spartera, Marzia Rigolli, Jane Ellis, Laura Bourne, Alina Florea, Konrad Werys and Mayooran Shanmuganathan for their help, conversation and sense of camaraderie.

From Oxford University Hospitals (OUH) hepatology: Justyna Szczurkowska, Ana Vieira, Jane Phillips and Vicki Wharton.

From Perspectum Diagnostics: Andy McKay, Andrea Dennis, John McGonigle, Marija Mavar and Rajarshi Banerjee.

I have met some incredible people during my time in Oxford and made/continued some true friendships. I would like to thank the following for their friendship and moral support: Alex Jones, Rob Burleigh & Kate Beare, Constantin Reinprecht, Andrew Lindsay, Matthew Drane, Ceren Yalaz, Sorana Mighiu, Karan Nagpal, Stefan Nekovar, Thomas Chui, Jack Feltham, Priyav Shah, Maria Jakobsdottir, Santiago Paredes, Peter Spring, Alistair Leitch, Rose Crosby, Laurence de Lussy Kubisa, Alex Archer, Bradley Kyne and many others I have forgotten.

I would like to thank my collaborators Jason Chang (Singapore General Hospital) and Jonathan Dillman (Cincinnati Children's Hospital).

My mother Nina has been a constant source of encouragement and belief throughout and my parents have made many sacrifices for me – I can't thank them enough. My brother Vikram deserves a special mention – from his Nike-inspired advice on whether to undertake this DPhil: 'just do it' – to his strength in looking after my father through his illness. Thanks also to Tina, Sonny, Rohan, Ambi, my Nanima and Neelam, Chrysanthi and Kris for their general love and support.

I acknowledge funding for my DPhil from the Medical Research Council (MRC) with MRC reference MR/M015734/1.

Word count: 44,810.

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 (Non-Invasive and Comprehensive Liver Assessment (NICOLA 13/SC/0243) and Rapid Imaging Assessment of the Liver (RIAL 11/H0504/2) and OUH Foundation Trust Management Approval. The study from Singapore received approval from the SingHealth Centralized Institutional Review Board (CIRB reference number: 2017/2425) and the study from Cincinnati was HIPAA compliant.

The NICOLA study was designed by Dr Michael Pavlides (MP) but I made substantial amendments to it over the course of this DPhil. All the patients included in Chapter 3 were recruited and scanned by Dr Rajarshi Banerjee, MP or Dr Christina Levick (CL), but I collected the outcomes data and performed the statistical analysis. In Chapter 4 the first 6 patients were recruited by MP or CL, but I followed them up with repeat study visits and recruited the other 11. In Chapter 5 patients from OCMR were recruited by Dr Michael Pavlides and Dr Christina Levick, patients from Singapore General Hospital were recruited by Prof Jason Chang and patients from Cincinnati Children's hospital were recruited by the clinical team there. I performed all the imaging and statistical analysis. In Chapter 6 data was sourced from the United Kingdom Biobank resource and automated analysis of MRI scans was developed by John McGonigle and Andrea Dennis of Perspectum Diagnostics Ltd. Healthy volunteers at OCMR were recruited by MP and CL and I performed the image and statistical analysis.

Conflict of interest:

I am a shareholder in Perspectum Diagnostics Ltd.

Publications

Papers submitted:

Jayaswal A, Levick C, Selvaraj E, Dennis A, Collier J, Cobbold J, Booth J, Tunnicliffe E, Kelly M, Barnes E, Neubauer S, Banerjee R, Pavlides M. Liver cT_1 predicts clinical outcomes in patients with chronic liver disease: Further evidence from 754 years of follow-up.

Dillman J, Tkach J, Ganghi D, Singh R, *Jayaswal A*, Miethke A, Trout A. Relationship between Magnetic Resonance Imaging Spleen T_1 Relaxation and Other Radiologic and Clinical Biomarkers of Liver Fibrosis in Children and Young Adults with Autoimmune Liver Disease.

Chang J, *Jayaswal A*, Cheng L, Tay M, Gogna A, Tan H, Too C, Low A, Tan C. Multiparametric MRI of liver and spleen predicts clinically significant portal hypertension and variceal bleeding.

Jayaswal A, Levick C, Collier J, Tunnicliffe E, Kelly M, Neubauer S, Barnes E, Pavlides M. Liver cT_1 decreases following treatment by direct acting antiviral therapy in patients with chronic hepatitis C virus.

Published abstracts:

Jayaswal A, Pavlides M, Banerjee R, Kelly M, Barnes, Neubauer S, (2019). Liver cT_1 predicts clinical outcomes in patients with chronic liver disease. Journal of Hepatology, 156, S1325 - S1325.

Jayaswal A, Pavlides M, Banerjee R, Kelly M, Barnes, Neubauer S, (2018). Liver cT_1 can predict clinical outcomes in patients with chronic liver disease, equivalent to biopsy. Hepatology 68(S1), 607A–608A.

Jayaswal A, Levick C, Pavlides M, Collier J, Kelly M, Neubauer S, Barnes E, (2017). Multiparametric magnetic resonance imaging can quantitatively assess fibroinflammatory response to direct acting antiviral treatment of hepatitis C virus. Hepatology 66, 336A-337A.

McKay A, Jayaswal A, Milanesi, M, Phillips M, Kin S, Xu X, Leal J, Banerjee R. (2016). Patient reported assessment of different diagnostic imaging tests for liver disease; visual representation and ownership of results improves understanding. Hepatology 64(1) Supplement. 861A.

Abbreviations

^{1}H-MRS:	proton magnetic resonance spectroscopy
95% CI:	95% confidence interval
AIH:	autoimmune hepatitis
AILD:	autoimmune liver disease
ALP:	alkaline phosphatase
ALT:	alanine aminotransferase
APRI:	aspartate aminotransferase to platelet count ratio index
ArLD:	alcohol-related liver disease
AST:	aspartate aminotransferase
AUROC:	area under the receiver operating curve
BMI:	body mass index
bSSFP:	balanced steady state free precession
CSPH:	clinically significant portal hypertension
CLD:	chronic liver disease
DAA:	direct acting antiviral therapies
DWI:	diffusion weighted imaging
ELF:	enhanced liver fibrosis
EoT:	end of treatment
FHVP:	free hepatic venous pressure
GGT:	gamma-glutamyl transferase
HBV:	hepatitis B virus
HCC:	hepatocellular carcinoma

HCV: hepatitis C virus

HVPG: hepatic venous pressure gradient

HR: hazard ratio

IQR: interquartile range

IVIM: intravoxel incoherent motion imaging

LIF: liver inflammation and fibrosis

MR: magnetic resonance

MRE: magnetic resonance elastography

MRI: magnetic resonance imaging

MRS: magnetic resonance spectroscopy

NAFLD: non-alcoholic fatty liver disease

NASH: non-alcoholic steatohepatitis

NCPH: non-cirrhotic portal hypertension

NSBB: non-selective beta-blockers

PBC: primary biliary cholangitis

PH: portal hypertension

PSC: primary sclerosing cholangitis

RF: radiofrequency

ROI: region of interest

SD: standard deviation

shMOLLI: shortened modified Look Locker inversion

SVR: sustained virological response

TE: transient elastography

WHVP: wedged hepatic venous pressure

Index of tables

Table 1: Description of patient cohort....86
Table 2: Baseline patient characteristics....93
Table 3: Table of clinical events....95
Table 4: Survival analysis results using binary cut-offs....97
Table 5: Liver event-free survival assessed by grouped liver cT_1 cut-offs....99
Table 6: Liver event-free survival assessed by grouped liver cT_1 cut-offs in patients with NAFLD, ArLD and VH....101
Table 7: Liver event-free survival by Ishak fibrosis stage....103
Table 8: Liver event-free survival by Ishak fibrosis stage in patients with NAFLD/ArLD/VH.104
Table 9: Liver event-free survival assessed by grouped TE cut-offs....105
Table 10: Liver event-free survival assessed by grouped TE cut-offs in patients with NAFLD/ArLD/VH....106
Table 11: Additional analysis assessing influence of failure/unreliability rates on MRI and TE results....107
Table 12: Individual baseline patient characteristics and self-reported alcohol consumption at each study visit....125
Table 13: Simple blood test results at baseline, 24 weeks EoT and 48 weeks EoT....126
Table 14: Non-invasive liver tests at baseline, 24 weeks EoT and 48 weeks EoT....127
Table 15*:* Magnetic resonance parameters at baseline, 24 weeks EoT and 48 weeks EoT ..128
Table 16: Value of HVPG measurements in prognosis....153
Table 17: OCMR baseline characteristics. Characteristics provided for the whole cohort, those with HVPG < 10 mmHg and HVPG $\geq$ 10 mmHg....163
Table 18: Spleen cT_1 and T_1 values for different severities of PH (OCMR)....164
Table 19: Diagnostic accuracy of spleen cT_1/T_1 for severity of PH and varices....166
Table 20: SGH baseline characteristics. Characteristics provided for all patients, those with HVPG ≥ 10 mmHg and < 10 mmHg,....184
Table 21: Prediction of severity of portal hypertension based on clinically used HVPG thresholds. Biomarkers with highest consistent AUCs highlighted in bold to be assessed for multivariate analysis....186

Table 22: Logistic regressions of best performing tests on univariate analysis for prediction of severity of PH. 187
Table 23: Prediction of future episodes of variceal bleeding by biomarker 188
Table 24: Prediction of clinically significant portal hypertension (CSPH) by biomarker including only patients on beta-blocker treatment. 190
Table 25: Summary of diagnostic and prognostic accuracy of non-invasive tests excluding patients with NCPH. 192
Table 26: Baseline characteristics (CCH).....205
Table 27: Summary of diagnostic accuracy of abdominal T_1/cT_1 for PH 217
Table 28: Correlations between spleen T_1/cT_1 and simple blood tests from UKBB 237
Table 29: Correlations between spleen T_1/cT_1 and continuous variables from UKBB 237
Table 30: Spleen MR-derived metrics for UKBB healthy volunteer subgroup 241
Table 31: Participant characteristics of UKBB healthy volunteer subgroup.....241
Table 32: Characteristics of healthy volunteers (OCMR).....243
Table 33: Significant differences in spleen cT_1 by dichotomous variables and disease severity. 258
Table 34: Ranges of spleen cT_1, T_1 and liver cT_1 values 262

Index of figures

Figure 1: Protons in the absence of an external magnetic field 55

Figure 2: Protons in the presence of an external magnetic field 55

Figure 3: Spectral density functions of different water pools 60

Figure 4: Simplified shMOLLI pulse sequence diagram 63

Figure 5: Frequency dependence of the bSSFP signal with TR = 2.3 ms and TE = 1.15 ms. 64

Figure 6: Balanced steady-state free precession sequence, with a symmetric readout. 65

Figure 7: Simulated off resonance dependence of the bSSFP signal 66

Figure 8: Example of a LiverMultiScan^{TM} report and associated metrics 72

Figure 9: Placement of voxels for proton spectroscopy (MRS) to quantify liver fat fraction. 74

Figure 10: Spectroscopy tool output detailing proton density fat fraction (PDFF) 77

Figure 11: Flow diagram of study protocol. 91

Figure 12: Kaplan-Meier plots for clinical outcomes stratified by liver cT_1. 99

Figure 13: Kaplan-Meier plots to show liver event-free survival stratified by liver cT_1 in patients with NAFLD, ArLD and viral hepatitis (VH) 100

Figure 14: Kaplan-Meier curves for liver event-free survival stratified by histological fibrosis stage 102

Figure 15: Kaplan-Meier curves for liver event-free survival stratified by histological fibrosis stage in patients with NAFLD/ArLD/VH. 103

Figure 16: Kaplan-Meier curves for liver event-free survival stratified by TE score 105

Figure 17: Kaplan-Meier curves for liver event-free survival stratified by TE score in patients with NAFLD/ArLD/VH 106

Figure 18: Flow diagram for direct comparison between liver cT_1 and transient elastography (TE). 107

Figure 19: Example of abdominal iron-corrected T_1 (cT_1) colour maps 109

Figure 20: Study flow diagram 124

Figure 21: Response to DAA treatment from baseline to 24 weeks EoT, from baseline to 48 weeks EoT and from 24 weeks EoT to 48 weeks EoT 136

Figure 22: Abdominal cT_1 images and segmentation masks pre and post treatment.....137

Figure 23: Liver cT_1 plotted against paired TE measurements 138

Figure 24: Comparison of liver cT_1 between patients with and without cirrhosis at baseline, 24 weeks EoT and 48 weeks EoT.....139

Figure 25: Study flow diagram (OCMR). 161

Figure 26: Stratification of severity of portal hypertension by spleen T_1/cT_1 in patients with chronic liver disease.....168

Figure 27: Correlations between HVPG and spleen T_1/cT_1 169

Figure 28: Receiver operating curves (ROC) to show classification of patients with HVPG ≥ 10 mmHg 170

Figure 29: Interobserver variability analysis 171

Figure 30: Study flow diagram (SGH) 182

Figure 31: Correlation between hepatic venous pressure gradient (HVPG) measurements and abdominal cT_1 185

Figure 32: Boxplots showing liver and spleen cT_1 by disease aetiology 207

Figure 33: Example abdominal cT_1 images of paediatric patients with PH and without PH 208

Figure 34: Boxplots to show differences in spleen cT_1 and T_1 in patients with and without radiological evidence of portal hypertension.....209

Figure 35: Receiver operating curve (ROC) for the assessment of portal hypertension (PH) by spleen cT_1.....210

Figure 36: Flow diagram describing selection of UKBB scans suitable for analysis. 233

Figure 37: Histograms of spleen T_1 and cT_1 distributions in the UK Biobank. 235

Figure 38: Correlation between spleen T_1 and spleen cT_1 in the UKBB.....236

Figure 39: Difference in spleen T_1/cT_1 in UKBB stratified by sex 239

Figure 40: Boxplots showing differences in spleen T_1/cT_1 by disease severity at 3T.....248

Figure 41: Boxplots showing differences in spleen T_1/cT_1 by disease severity at 3T by sex.....249

Figure 42: Boxplots showing differences in spleen T_1/cT_1 by disease severity at 1.5T.....250

Figure 43: Boxplots showing differences in spleen T_1/cT_1 by disease severity at 1.5T by sex.....251

Figure 44: Boxplots showing increasing spleen cT_1 with increasing disease severity 252

Figure 45: Boxplot showing the difference in paediatric spleen cT_1 from CCH to adult values. 253

Figure 46: Boxplots showing spleen cT_1 across all centres categorised by severity.....254

Thesis abstract

Stratification of disease severity in patients with liver fibrosis, cirrhosis and portal hypertension using novel multiparametric magnetic resonance imaging techniques

Arjun Narayan Ajmer Jayaswal

Jesus College

Thesis submitted for DPhil in Michaelmas Term 2019

Chronic liver diseases (CLDs) affect millions worldwide and are a major cause of mortality. Recently, multiparametric magnetic resonance (MR) methods have been proposed as a non-invasive, quantitative biomarker for assessment of CLD.

CLD traverses a spectrum of severity, and the aim of this thesis was to apply MR methods across this spectrum for a quantitative assessment of disease severity. The primary variable of interest was the iron and field strength corrected T_1 (cT_1), generated by Liver*MultiScan*™ software from T_1 and T_2* mapping techniques.

Liver cT_1 was found to predict liver-related clinical outcomes in patients with CLD ($p<0.0001$), including in patients with compensated disease only. Liver cT_1 performed especially well in patients with non-alcoholic fatty liver disease (NAFLD), alcohol-related liver disease (ArLD) and viral hepatitis (VH) with a cut-off of liver cT_1>825ms having a negative predictive value (NPV) of 100%.

Liver cT_1 quantitatively assessed changes in fibroinflammation in 15 patients with chronic hepatitis C virus (HCV) undergoing treatment by direct acting antiviral (DAA) therapy. Liver cT_1 showed significant decreases from baseline to 24 weeks after end of treatment (EoT) (876 vs 806 ms, p=0.002), from baseline to 48 weeks EoT (876 vs 788 ms, p=0.0002) and from 24 weeks EoT to 48 weeks EoT (806 vs 788ms, p=0.016).

Spleen cT_1 had good diagnostic accuracy for presence of portal hypertension in both an adult cohort with CLD and a paediatric cohort with autoimmune liver disease (area under the receiver operating curve (AUC) of 0.85 and 0.81 respectively). In another adult cohort, spleen cT_1 was predictive of future episodes of variceal bleeding (AUC: 0.74), especially in those on non-selective beta-blocker treatments (AUC: 0.97). A reference range for spleen T_1 and cT_1 was established using data from the United Kingdom Biobank population health study and healthy volunteers from our centre. Higher spleen T_1 /cT_1 was observed in women vs men and in children vs adults.

This thesis provided further evidence that multiparametric MR methods combining T_1 and T_2* relaxometry can be used across the spectrum of CLD for stratification of disease severity.

Chapter 1: Introduction

1.1 Chronic liver disease pathophysiology

The liver plays multiple essential roles including, but not limited to: production of proteins and clotting factors, breakdown of toxins, excretion of waste products, fat and carbohydrate metabolism, bile production and others e.g. immune function. The liver weighs approximately one and a half kilograms, is a red/brown colour and anatomically is placed below the lungs, to the right of the abdomen. The liver's blood supply is delivered by dual sources; the portal vein (from the digestive system) and the hepatic artery (from the heart). The liver is the largest and one of the most remarkable internal organs; it has the ability to regenerate.

Chronic liver disease (CLD), regardless of the underlying aetiology – be it viral hepatitis, alcohol misuse or fatty accumulation – is characterised into distinct stages. When there is an insult to the liver, the liver becomes inflamed, and hepatocytes (liver cells) are damaged/killed as the immune cells try to deal with the underlying injurious agent.

Constant inflammation, healing and further inflammation over a period of time causes scar tissue to start forming. This is called liver fibrosis, which begins replacing the healthy tissue. The remaining healthy tissue must work harder to perform the liver's functions. Liver fibrosis is characterised by deposition of collagen in the extracellular space. Over time (15-20 years) as this scar tissue increases, becomes interlinked, and crosses over, it develops into cirrhosis and the liver obtains a rough and coarse texture. Patients with cirrhosis are at higher risk of developing hepatocellular carcinoma (HCC), but can remain asymptomatic for a long period of time (defined as having compensated advanced CLD, or cACLD).

Cirrhosis can impair blood flow from the portal circulation to the liver and can lead to portal hypertension (PH), which is associated with development of varices, clinical decompensation (development of ascites, variceal bleeding, hepatic encephalopathy (HE)), HCC and death.

1.2 Chronic liver disease epidemiology

CLDs affect an estimated 1.5 billion people worldwide[1], with mortality rates of 2 million per year[2]. The estimated worldwide prevalence is 25% for non-alcoholic fatty liver disease (NAFLD)[3], 4.5% for chronic viral hepatitis (VH) B and C[4] and 5.1% for alcohol use disorders (AUD)[5], which can lead to alcohol related liver disease (ArLD). CLDs have previously been reported to induce cirrhosis in approximately 633,000 patients per year in the United States[6]. Once cirrhosis is established, each year approximately 5-7% of patients develop decompensation[7]. Cirrhosis is the 11th most common cause of mortality worldwide[8] and liver cancer is reported as the 4th most common cause of cancer-related mortality[9], with 5-year survival of only 18%[10].

New efficacious treatments have become available for viral hepatitis C (HCV), and although large numbers have been treated, there are likely to be many more undiagnosed or still requiring treatment. There are currently no approved pharmacological treatments for NAFLD's more advanced embodiment non-alcoholic steatohepatitis (NASH) or for alcohol-related liver disease (ArLD). Lifestyle changes are shown to be extremely effective in normalising liver fat[11], and abstinence from consumption of alcohol is a consistent predictor of survival in ArLD[12,13] but adherence to both these lifestyle adjustments is low. As a result, major public health initiatives such as minimum unit pricing (MUP) of alcohol in Scotland and the sugar tax in the UK have been introduced to tackle the increase in mortality from alcohol and contribution to the obesity (and subsequently NAFLD/NASH) epidemic from excess sugar.

In the UK, the burden of CLDs from both the human and financial standpoints have been described multiple times. In a seminal report from the Lancet Commission on live disease in 2014[14] it was estimated that since the 1970s standardised liver mortality had quadrupled,

even as death from other causes had decreased. Financially, the full societal costs of alcohol misuse and obesity have been estimated at between £27-52 billion and £27-46 billion respectively[15,16]. Liver disease disproportionally affects the poorest and most deprived in society and is a major cause of mortality in younger patients, with patients in the most deprived quintile of the country dying on average 9 years earlier than the least deprived quintile[17]. This compounds socioeconomic inequality. Liver disease is also associated with reduced productivity[18] and is therefore detrimental to the overall economy. CLDs have been shown to be associated with higher unemployment, health care expenditure and impairment in varying aspects of quality of life in the US[18].

It is clear that liver disease is now a major healthcare issue that requires urgent attention. In this introduction, I aim to briefly describe current diagnostic tools (invasive and non-invasive) for disease stratification and some of the unmet clinical needs in liver disease. I will then provide a more in-depth background of how magnetic resonance imaging (MRI) techniques in particular can be of use in the assessment and risk stratification of liver disease.

1.3 Methods of stratification of disease severity and need for non-invasive tools

Extent and severity of hepatic fibrosis is considered a marker of current disease stage and hepatic inflammation is thought to be the underlying driver of fibrosis progression. Liver biopsy, which examines approximately 1/50,000th of the liver parenchyma, is still the clinical 'gold-standard' method of assessing these parameters (in addition to liver fat and iron deposition), despite it being affected by sampling variation, interobserver variability and high costs[19,20]. It also has associated risks of complications and requires patients to remain hospitalised for several hours. Clearly, non-invasive methods to quantify liver fibrosis, inflammation, fat and iron would be of clinical utility. In addition, there are multiple

histological scores used to assess fibrosis (Ishak, METAVIR, NASH CRN) for different disease aetiologies and differences in fibrosis stage are marked by changes in tissue architecture, not necessarily amount of fibrosis, thus rendering it only a *semi*-quantitative tool. These drawbacks cause some to consider biopsy as neither a true gold-standard for diagnostic assessment nor an accurate comparator for development of non-invasive diagnostic tools[21]. Quantitative histology techniques have been described (e.g. collagen proportionate area (CPA) analysis), but they still are affected by sampling error. A more useful metric against which to compare potential non-invasive tools would be prognostic ability to predict clinical outcomes.

The ability to track changes in liver disease status in response to therapy is an essential facet of a diagnostic tool in both clinical care and drug development.

The requirement for histological endpoints in pharmacological therapeutics trials for liver diseases hinders both recruitment and retention rates and requires larger cohorts due to inherent reporting variability of liver biopsy. There is therefore an unmet need for non-invasive diagnostic tools that can quantitatively assess changes in liver disease in both the clinical setting and in trials for drug development, especially in NASH and antifibrotic therapy trials.

Liver biopsy is also unsuitable for routine follow-up of patients in clinical care. In the clinical setting of treatment for chronic HCV, until recently the focus was on stratification of disease severity to enable those with most advanced liver disease to be put forward for treatment to eradicate the virus. Now the unmet need is shifting towards assessing the effects of these treatments on liver fibrosis/inflammation and identifying those who, despite viral eradication, are at heightened risk of HCC development post-treatment. A non-invasive tool that can both

track disease change and stratify those at future risk of HCC development would be useful in this context.

PH is a serious complication of advanced liver disease. The current gold-standard measure of severity of PH is the hepatic venous pressure gradient (HVPG), which is invasive, expensive and requires considerable operator expertise[22]. The HVPG is an excellent diagnostic and prognostic tool[23–26], and if a non-invasive surrogate biomarker could be found that can perform similarly, it would be of great clinical utility. Development of varices and subsequent variceal bleeding are common symptoms of clinically significant portal hypertension (CSPH), and currently detection of varices is performed by endoscopy. Conducting research to identify a non-invasive screening tool to avoid unnecessary endoscopies has been recommended by the Baveno VI recommendations report[27].

1.4 Non-invasive diagnosis of fibrosis, cirrhosis and portal hypertension

Blood serum-based markers are used extensively in both primary, secondary and tertiary care settings. Simple blood tests and scores derived from them are cheap, repeatable and have well-established, standardised cut-offs for disease evaluation. They can be affected however, by extrahepatic factors, exercise levels[28] and other variables, therefore lacking specificity to liver disease. Guidelines[29] suggest liver function tests (LFTs) may be normal even in advanced liver disease and are not infrequently abnormal in the absence of liver disease. Therefore, the interpretation of blood tests in isolation may be ineffective in diagnosis or exclusion of liver disease. Certain serum-based composite markers of fibrosis such as the fibrosis-4 score (FIB-4), aspartate aminotransferase / alanine aminotransferase (AST/ALT) ratio, and the AST to platelet count ratio index, (APRI) have all shown to have good diagnostic accuracy for advanced fibrosis/cirrhosis and higher FIB-4 and APRI values are associated with development

of clinical outcomes[30,31], but mainly in studies assessing patients with hepatitis C. Additionally, serum tests have been shown to fluctuate rapidly post-treatment[32], indicating that their relationship with fibrosis are likely being affected by inflammation. Other more advanced tests based on blood serum parameters, such as the enhanced liver fibrosis (ELF™) test and FibroTest™ are considerably more expensive but are able to diagnose severity of fibrosis and have shown prognostic ability[31,33–37] as well, but not all local laboratories perform the analysis required.

1.4.1 Ultrasound elastography

Elastography measures the stiffness of an organ by assessing the speed of shear wave propagation through it. Transient elastography (TE) is used widely now in clinical practice and forms part of published NICE guidelines on management of certain liver diseases[38,39]. In TE mechanical waves are generated and evaluated using an ultrasound probe placed on the patient's right side. It measures tissue stiffness over approximately a 1 cm diameter by 4 cm length region of tissue. It has the advantage of being quick to perform, with instantaneous results. TE has been shown to correlate with liver fibrosis stage in multiple studies outlined below. Cut-offs however are not standardised across disease aetiology and TE can be affected by inflammation, steatosis, congestion and operator expertise. Additionally, it has shown to have significant technical failure rates and/or rate of unreliability, especially in NAFLD/NASH patients[40,41]. Some of these limitations have been addressed with the addition of an XL probe that allows the shear waves to travel further into the tissue. Other methods of elastography measurement include two-dimensional shear wave elastography (2D-SWE), point shear wave elastography (pSWE)[42], acoustic radiation force impulse (ARFI) imaging[43]. For these techniques, either the shear wave speed (ms^{-1}) or Young's modulus (kPa) is reported.

Liver fibrosis severity

TE has shown good correlation with stage of liver fibrosis[44,45] and good accuracy for presence of significant fibrosis but has a wide range of cut-offs reported in studies, often with different cut-offs for different diseases.

Recently a large individual patient data analysis (collecting data from 16,032 individuals from 26 cohorts established healthy ranges for TE, including in healthy nonobese individuals (mean stiffness 4.68 kPa (95% CI 4.64-4.73) and identified factors that modified TE values including diabetes, hypertension, dyslipidaemia and steatosis and that affected the measurement from the M and XL probes differently[46]. Additionally, in this study patients TE values assessed with the XL probe were higher than TE values assessed by the M probe.

Portal hypertension

Both liver and spleen stiffness as assessed by various forms of elastography have been evaluated in many studies for the assessment of portal hypertension and its symptoms/complications. The methods include pSWE, ARFI and TE. Liver stiffness' ability to predict clinically significant PH and its correlation with HVPG values are reviewed in a paper by Berzigotti[47]. Additionally, a meta-analysis of 910 patients from 11 studies also showed that a liver TE cut-off of 13.6 kPa could exclude CSPH effectively and 22 kPa could confirm CSPH effectively[48].

Spleen stiffness as measured by TE was first shown to correlate well with HVPG measurements by Colecchia et al[49] in 141 patients with compensated HCV-related cirrhosis and to diagnose presence of varices by Stefanescu et al[50] in 174 patients of mixed aetiology. Studies by Sharma[51], Zykus[52] and Tseng et al[53] also independently showed significant correlations between spleen stiffness as measured by TE and HVPG. Most of these studies

were in mixed cohorts of patients with liver cirrhosis. However, when including other studies assessing spleen stiffness' ability to diagnose CSPH, reported AUROCs have varied between 0.640 and 0.966[47].

One systematic review and meta-analysis showed spleen stiffness to have summary correlation coefficient of 0.72 (95% CI: 0.63-0.8) and summary AUC of 0.84 (0.72-0.92). Spleen stiffness has also been used as an adjunct marker to liver stiffness for the detection of varices as mentioned above[54]. A drawback has been that spleen stiffness has been technically feasible in a smaller proportion of patients than liver stiffness.

Originally, spleen stiffness values only reached 75 kPa, the same upper limit for liver stiffness. Recently a new spleen specific probe has been developed[55] and described in a paper by Stefanescu et al.[56] to increase the spared endoscopy rate when used in combination with the Baveno VI criteria (for screening of patients with varices) to 26.5%.

1.4.1 Magnetic resonance techniques

The phenomenon of nuclear magnetic resonance (NMR) and the subsequent techniques of spectroscopy (MRS) and imaging (MRI) are excellent methods to probe for biological, chemical and physical information *in vivo*. This is because they do not ionise matter and are non-invasive.

In comparison to other diagnostic imaging tools such as x-ray (developed by Roentgen in 1895[57]), ultrasound (used medically by Ludwig in 1949[58]) and CT, MRI is a relatively modern technique. Original research by Rabi showed that molecules sent through a magnetic field could emit electromagnetic waves in the radiofrequency range at specific frequencies. Bloch[59]

and Purcell[60] extended this work in 1946, and Damadian[61] in 1971 showed that NMR could detect cancerous tissue. Lauterbur[62] created the first MR image in 1973, and subsequently images from the pelvis up to the brain were acquired in the 1980s[63–68].

MR techniques are increasingly being used in the assessment of liver disease, both for qualitative (i.e. tumour detection) and quantitative (quantitation of liver fat, iron or fibrosis/inflammation) tasks and some of these techniques are reviewed here.

Liver fat

Fatty accumulation can often occur in CLD and is histologically defined as an abnormal deposition of fat (also termed steatosis) in more than 5% of hepatocytes[69]. Fatty liver is most commonly found in patients with ArLD and NAFLD. MRS can accurately quantify percentage of liver fat and is widely accepted as the non-invasive gold standard. MRS has been extensively validated against biopsy with excellent agreement[70]. Furthermore, it provides a continuous scale, rather than a semi quantitative one, making it more suitable for longitudinal observation.

MRS has limitations in that it only samples a small part of the liver (as shown in Chapter 2), is not always available on clinical scanners and requires significant operator expertise. MRI PDFF (proton density fat fraction) is another method of quantifying liver fat and is based on a simple technique of acquiring in and out of phase water and fat images of the whole liver and generating PDFF through a simple 'add and subtract' scheme. MRI PDFF also agrees strongly with MRS using both DIXON[71] and multispectral MRI techniques[72] in adults and in children[73,74].

Both MRI and MRS have been shown to be more accurate in quantifying liver fat fraction than pathologists' assessment of histology[75,76]. MR quantification of liver fat is now commonly used in research into fatty liver, as clinical trial endpoints, and in longitudinal studies assessing

effects of lifestyle modifications or weight loss treatment on liver fat[77,78]. MRI-PDFF has also highlighted the prevalence of elevated liver fat (using 5.21% liver fat, 5% of all volunteers had steatosis and 33% of those with BMI > 30kg/m^2 had steatosis) in the United Kingdom Biobank (UKBB) population health study[79].

Liver iron content

Detecting excess amounts of liver iron content (LIC) is important as it is toxic to hepatocytes. Elevated levels of LIC are most commonly found in patients with hereditary haemochromatosis (HH) but are also often found in patients with ArLD and NAFLD. Iron deposition can be identified histologically using Prussian blue staining and graded on a semiquantitative scoring system but is not used for routine diagnosis of iron overload.

MRI parameters R_2 and R_2* (reciprocals of T_2 and T_2*, described in more detail in Chapter 2) have been shown to correlate extremely well with LIC[80–84] and are considered the non-invasive gold standard for LIC quantification. Commercially available quantification of LIC by R_2 methods is provided by FerriScan® (Resonance Health Ltd) for example, and guidelines for management of patients with HH recommend using MR techniques[85].

MRI T_2* quantification of LIC has shown the prevalence of elevated LIC (> 1.8 mg/g) to be approximately 5% in a sample of the United Kingdom Biobank (UKBB) population health study by McKay et al[86]. Liver T_2* has also been shown to correlate with levels of serum ferritin in patients with NAFLD[87]. T_2* values of the liver have also been described in healthy volunteers, again correlating with levels of serum ferritin[88]. Other less widely utilised methods involve magnetic susceptibility imaging[89] and some groups have also attempted to used diffusion weighted imaging (DWI) to assess LIC in beta-thalassemia patients[90].

1.4.2 Magnetic resonance elastography (MRE)

MRE is a method of investigating the biomechanical properties of tissue. Mechanical excitations generate longitudinal waves which, when they reach organ tissue, are changed into shear waves of the transverse kind. These shear waves are imaged using phase contrast MRI. Post-processing converts the images into elastograms that quantify tissue stiffness. MRE can be incorporated into a standard MRI examination with the addition of certain hardware and software. MRE measurements of the liver are now well established as a marker of liver fibrosis and measurements of the liver and spleen have been investigated for assessment of PH.

MRE for assessment of liver fibrosis and cirrhosis

MRE has been extensively assessed against histological liver fibrosis[91] and has been shown to have good diagnostic accuracy for differentiating between stages of liver fibrosis[92].

In a systematic review and meta-analysis carried out at our centre (personal communication from the LITMUS: Liver Investigation: Testing Marker Utility in Steatohepatitis) MRE performed excellently for diagnosis of significant fibrosis, advanced fibrosis and cirrhosis.

In addition to some single centre studies[93–95], MRE has been shown in a systematic review and IPD analysis[40] (3 cohorts included, 2 from USA, 1 from Japan) to have higher diagnostic accuracy than TE in detection of each stage of fibrosis in patients with biopsy proven NAFLD. In this study the AUCs for stage 1, 2, 3 and 4 for TE and MRE were 0.82 vs 0.87, 0.87 vs 0.92, 0.84 vs 0.93 and 0.84 vs 0.94 respectively. This study excluded failed TE results (81/318, 25%) and unavailable MRE results (3/318) from the analysis. Another study by Caussy et al[96] found that BMI was significantly associated with discordance between MRE and TE in diagnosis of fibrosis stage and Wagner et al showed BMI to be independently associated with increased

technical failure, indicating a confounding effect of factors associated with BMI e.g. liver fat. It is known that low T_2* (high LIC) can degrade the MRE signal, resulting in failure. Ghoz et al[97] showed that in a study of 95 patients with iron overload, low T_2* (<20 ms) was able to predict non-diagnostic MRE test results with AUC 0.95 and that patients with non-diagnostic MRE had median T_2* 6 ms vs 20 ms for those with acceptable MRE scans. This technique may not be as applicable in the presence of iron overload and/or obesity. However recent adjustments to acquisition parameters can improve technical performance in the presence of iron overload[98].

MRE for assessment of PH

A study showed liver stiffness as assessed by MRE to be correlated with HVPG measurements ($r = 0.64$, $p = 0.01$) in 15 adults with advanced liver disease caused by HCV, HCV/HIV or HIV/NASH. In this study only 4/15 had HVPG $\geq$ 10 mmHg[99].

In a pilot study by Guo et al[100], 10 patients with cirrhosis had MRE before and after transjugular intrahepatic portosystemic shunt (TIPS) procedure, with HVPG also being measured. The magnitude of the viscoelastic modulus decreased significantly after surgery in both liver and spleen, but only the spleen modulus was correlated with relative change of HVPG.

In a paediatric cohort of 22 patients with biliary atresia having undergone Kasai operations, splenic MRE values (but not liver MRE values) were significantly different in those with and without clinical diagnosis of PH, with/without presence of varices, and could predict presence of gastroesophageal varices with AUC 0.844. However, a low APRI cut-off (0.65) was able to predict varices with the same AUC[101].

However, a study by Wagner et al.[102] showed that liver stiffness but not spleen stiffness, spleen T_1 or liver T_1 was correlated with HVPG. In this study, spleen stiffness was only successful in 27/33 patients (81%) with MRE failure defined as the absence of visualized wave propagation on the wave images and/or no pixel with a confidence index higher than 95% on the confidence map. This increased failure rate may be due to the longer distance the shear waves have to propagate through to reach the spleen than the liver. However, in separate study (assessing correlation of spleen MRE with liver fibrosis stage) it was found that all (n=109) patients had 100% technical success rate[103].

3D multifrequency MRE can decompose the viscoelastic parameters of the organ under observation into different components, namely the stiffness, elasticity and viscosity. The multifrequency approach also allows the generation of the frequency power law γ, which represent the architectural structure of the tissue. This technique was assessed in 36 patients with cirrhosis being evaluated for liver transplantation and showed that the splenic loss modulus, which represents the viscosity of the spleen, was best correlated with HVPG ($r = 0.57$, $p = 0.002$) and could identify those with severe PH (HVPG $\geq$ 12 mmHg) with AUC = 0.81 and those with high-risk varices with AUC = 0.93 but had a technical success rate of only 86%[104]. Another study by Shin et al. showed in 139 patients with liver cirrhosis screened for varices spleen stiffness by MRE to correlate with grade of oesophageal varices, but failed 17/156 (11%) of patients[105]. A similar study showed liver stiffness by MRE to have a good AUC for presence of varices AUC: 0.859 and high grade varices AUC 0.810[106].

This summary shows that there are conflicting results relating to whether elevated spleen or liver stiffness as assessed by MRE is best associated with HVPG values and clinical symptoms of PH.

1.4.3 Molecular MR

Molecular MR (MolMR) techniques have been developed to quantify collagen, which is a major protein component of the extracellular matrix accumulation that occurs in the setting of fibrosis. A molecular probe is used in a similar method to the usage of contrast agents in DCE-MRI. In this case the probe targets collagen rather than hepatocytes, and therefore, is theoretically a more accurate measure of fibrosis rather than liver function. The probe most commonly used is EP-3533.

Strong correlations between EP-3533 concentrations and Ishak fibrosis stage have been observed in rats and mice (r = 0.84 (rats), r = 0.79 (mice)) in one study[107] and liver-to-muscle contrast to noise ratio following EP-3533 injection exhibited strong correlation with hydroxyproline (r = 0.89) in mouse model of CCl_4 induced fibrosis in another study[108].

MolMR's ability to track response to anti-fibrotic therapy has been assessed in a BDL rat model in two separate studies, the first assessing response to rapamycin[109] (Farrar) and the second assessing response to antifibrotic treatment by EDP-305 in BDL rats[110].

Another study combining MolMR and MRE showed that MolMR was more sensitive to early fibrosis but MRE more so to advanced fibrosis in a rat model of diethylnitrosamine (DEN) induced fibrosis, indicating a combination of the two could lead to increased accuracy for fibrosis staging[111].

These studies have all been carried out in preclinical models only, indicating significant further development required before being clinically applicable.

1.4.4 MR assessment of abdominal blood flow for assessment of portal hypertension

Blood flow and haemodynamics in the portal circulation can become impaired in the setting of PH, and therefore MR methods of assessing these changes have been investigated to see if clinically useful information can be obtained from imaging. The most common technique for assessment of the portal circulation is Doppler ultrasound (US), which can measure flow velocity. It can only be performed in large vessels, is often unsuitable for visualising collaterals due to presence of bowel gas and has been shown to have high variability in application for the assessment of PH[112]. It therefore has limited clinical applicability.

2D phase contrast MRI (2D PC-MRI) is a more reproducible[113,114] technique for evaluation of blood flow dynamics and PH[115] but requires significant time to position imaging planes and often only visualises large vessels such as the portal vein.

McDonald et al[116] showed that blood flow in the superior abdominal aorta was significantly reduced ($p = 0.03$) in patients with PH and varices but in no other vessels following the administration of 4 weeks of nonselective beta-blockers. Palaniyappan et al[117] showed that the increased blood flow in the azygous vein was significantly correlated with HVPG using 2D PC-MRI ($r = 0.656$, $p < 0.001$), but not in patients with CSPH.

Another application of PC-MRI can yield the splenic arterial resistive index (SARI), which is a haemodynamic property that quantifies the difference between the peak blood velocity and the end-diastolic velocity in the splenic artery (SA), normalised to the peak velocity of the SA. SARI had an AUC of 0.873 for the diagnosis of CSPH in a cohort of 47 patients with cirrhosis of mixed aetiologies (ArLD 57%, HBV/HCV 34%)[118]. Whilst exhibiting a moderate correlation with HVPG in the entire cohort ($r = 0.504$), when excluding patients with splenomegaly, SARI was

strongly correlated with HVPG (r = 0.830) in comparison to liver stiffness by SWE, which only had a moderate correlation (r = 0.601).

4D flow (i.e. time-resolved 3D flow) analysis allows for retrospective analysis of blood vessels and visualisation of haemodynamics. Detailed images are obtainable, but small collateral vessels are still difficult to detect. These techniques have shown mixed interobserver agreement[119,120] and variable test-retest repeatability depending on size of vessels examined[120]. Studies have used this technique in particular to observe haemodynamic changes in the portal circulation pre and post transjugular intrahepatic portosystemic shunt (TIPS) procedures to alleviate portal pressure and have been suggested to be of use in assessing haemodynamic changes after partial hepatectomy[121].

A recent study by Motosugi et al[122] assessed the ability of 4D flow MRI to predict severity of varices in 23 patients with known liver cirrhosis scheduled for endoscopy. It showed that flow in the azygous vein > 0.1 L/min (100% sensitivity and 62% specificity) and fractional reductions in portal venous flow (100% sensitivity and 94% specificity) were able to distinguish participants with high-risk varices from those with low-risk or no varices. However, visualisation only assisted radiologists to view varices in 4/15 patients and although encouragingly the scan time was reported as only 10 minutes, post processing of images took an estimated 0.5-1 hour per patient.

2 studies have investigated 4D flow pre-and post-TIPS insertion. In the study by Stankovic et al[123], 11 patients with advanced cirrhosis were imaged 24 hrs before and 4 weeks after TIPS insertion with significant increases in blood flow quantified in the PV, hepatic artery (HA) and combined splenic and superior mesenteric arteries. In a similar study by Bannas et al[124] seven patients with PH and refractory ascites were imaged before TIPS and 2- and 12- weeks post

TIPS insertion. This study showed increases in PV, superior mesenteric vein and splenic vein, and that most of the haemodynamic changes occurred within the 2 weeks post TIPS with no significant difference in blood flow parameters between 2- and 12-weeks post-TIPS. The investigators observed TIPS-to-PV flow ratio of 0.8 and 0.9 at timepoints after TIPS in six out of seven patients who resolved ascites, but a ratio of 4.6 and 4.3 in the single patient with refractory ascites post TIPS.

The area of portal haemodynamics assessment by MR has therefore shown varied results to date with varying acquisition techniques and parameters used.

1.4.5 Diffusion weighted imaging

Diffusion weighted images (DWI) provide contrast between tissues due to the differences in thermal molecular motion of water, otherwise termed molecular diffusion. For an environment in which molecules' movement is unhindered, the relationship between displacement and diffusion coefficient (*D*) is simple and displacement rises linearly with time. Contrast this with a bound environment in which molecules can only be displaced a certain distance and this relationship will not hold. MRI techniques can track phase changes in water molecules due to their movement and determine the difference in position due to diffusion. In patients with liver fibrosis and cirrhosis, deposition of collagen is postulated to also increase the restrictiveness of the environment. Diffusion can be expressed as the apparent diffusion coefficient (ADC), with a lower value indicating a more restricted environment. DWI techniques do not require the use of a contrast agent.

DWI has been used in many preclinical studies and also used in studies attempting to assess extent of fibrosis. However, most of the studies (summarised below) have shown mixed results.

Initial studies showed conflicting results with one study showing strong ADC association with histology activity index (HAI) stage in patients with viral hepatitis ($r = -0.798$, $p < 0.0001$) and Child Pugh Score ($p < 0.0001$)[125] and another showing no correlation between ADC and inflammation or fibrosis score in patients with HCV[126]. Since then most studies have used DWI to assess differences between tumour tissue and surrounding liver parenchyma. However, in assessment of liver fibrosis the initial story remains similar, with some studies reporting good associations with stage of fibrosis and other studies reporting the opposite. Bonekamp et al[127] reported a correlation ($r = -0.54$, $p < 0.001$) between ADC and fibrosis stage, but with significant overlaps in the spread of ADC values between stage, Bakan et al reported[128] lower ADC values in patients with chronic hepatitis than in controls. On the other hand Verloh et al[129] reported reducing ADC with increasing fibrosis stage, but no difference between cirrhosis and advanced fibrosis. Bulow et al[130] in a patient population of 95 showed that ADC was not correlated with fibrosis stage ($R^2 = 0.38$, $p = 0.17$), and in another study ADC was not associated with fibrosis in 64 children with NAFLD[131]. Currently it is accepted that DWI is a good marker of cirrhosis, but still cannot consistently differentiate between stages of fibrosis.

Some of the reason for this may be that DWI is affected by confounders such as steatosis[130,132,133], iron[130,134] and perfusion changes in the liver. Technically, DWI is also challenging, with sequences being extremely sensitive to motion.

In a preclinical study, although fibrosis with decreasing ADC values, the relationship was not as strong as that of MRE and fibrosis stage[135]. DWI has also been found to be suboptimal when compared to MRE in other studies in humans[136,137].

1.4.6 Intravoxel incoherent motion diffusion weighted imaging

Intravoxel incoherent motion (IVIM) diffusion imaging is another MRI method that attempts to separate the signal from microvascular perfusion from diffusion to quantify liver fibrosis[138]. This is achieved by performing multiple MRI acquisitions similar to those used in DWI but using a range of certain scan parameters. Multiple parameters are obtained from IVIM protocols that relate to diffusion, perfusion-related diffusion and perfusion fraction. Some of these parameters have been shown in some studies to negatively correlate with increasing fibrosis stage[139–141] and liver function scores[142]. However, these studies have also shown large overlaps in these perfusion/diffusion values between fibrosis stage, making IVIM as yet unsuitable for accurate diagnosis of fibrosis stage as evaluated in a review by Li et al[143]. Another study has shown IVIM to be inferior to MRE for predicting significant and advanced fibrosis[144]. Additionally, there are major variations between research groups in the MR scan protocol parameters and post-processing protocols used. IVIM may be a promising field but requires significant improvements in these areas.

1.4.7 Arterial spin labelling

Other studies have aimed to use arterial spin labelling (ASL) to quantify perfusion parameters of the liver parenchyma, as both cirrhosis and HCC can cause regional and global perfusion changes. This technique magnetically 'tags' a group or 'slab' of protons that are entering the liver via either the hepatic artery or portal vein. These water protons flow into the area of interest, where they modify the magnetic equilibrium of that tissue (by about 1-2%). Subtraction of this image from a normal image acquired prior to labelling yields a perfusion-weighted image. This technique has been used in the brain extensively but is more challenging in the liver due to movement and also the dual input of arterial and venous blood.

Studies have shown feasibility of the technique in mice[145], differences between controls and mice with hepatitis ($p = 0.008$)[146] and normal values in the human liver[147,148]. Studies have examined the ability of ASL to dynamically assess changes in hepatic perfusion in pre- and post- prandial conditions, showing a significant increase in perfusion both healthy volunteers[149] and in patients with compensated cirrhosis[150].

1.4.8 T_1 relaxometry

T_1 relaxation is a characteristic MRI property of atoms. Hydrogen is the most abundant atom in the human body and is present in almost all tissue as a constituent of water, fat, cartilage etc. Hydrogen has different T_1 relaxation times depending on its surroundings, described in further detail in Chapter 2. This results in different tissues having different characteristic T_1 values. This can be utilised in generating contrast in traditional MR images to assist clinicians in observing abnormalities (e.g. torn cartilage in the knee) but can also be used quantitatively to extract numerical T_1 values from images (or 'maps') of various tissues.

Quantitative T_1 mapping was first utilised in the brain and is currently extensively used in assessment of cardiac disease, with many current techniques based on the Look-Locker method proposed in 1970[151]. Liver T_1 mapping came to the fore slightly later, due in part to initially modest results in diagnosis of fibrosis stage, but also due to the discovery of confounding factors that affect measurements of T_1.

Studies in 1981[152], 1982[153] and 1994[154] showed that liver T_1 measurements were higher in patients with either cirrhosis or chronic active hepatitis (CAH) (now believed to be hepatitis C, which was not discovered until 1989) when compared to healthy controls. Thomsen et al[155] also showed longer T_1 in cirrhosis but T_1 did not show an association with histological fibrosis ($n = 10$). The ability to diagnose anything other than cirrhosis vs healthy controls in both

humans and in animal studies did not advance until fairly recently. In 1987, the Clinical NMR group[156] not only observed a tendency of longer T_1 times in patients with cirrhosis and hepatitis, but also in patients with inflammation and no cirrhosis indicating inflammation as a confounding variable on liver T_1. Shorter T_1 was observed in patients with high LIC in this study.

In 2012, conflicting results from two studies showed T_1 to be shorter[157] and longer[158] in patients with cirrhosis. In addition, 2 studies in rat models of fibrosis induced by thioacetamide and carbon tetrachloride found no significant differences between controls and treated rats[159,160]. An important finding by Chamuleau et al[161] elucidated another confounding variable that could contribute to longer T_1 values. The investigators observed increases in T_1 in rats administered with CCl_4 that were due to drug induced oedema rather that fibrosis.

A study by Cassinotto et al[162] then showed that liver T_1 mapping had a higher diagnostic accuracy than liver and spleen DWI for the diagnosis of cirrhosis, CP B/C cirrhosis and large varices ($p < 0.01$) in a prospective cohort study of 40 healthy volunteers and 89 patients with varying levels of cirrhosis. A study investigating fibrosis induced in 60 rats by bile duct ligation (BDL) and CCl_4 measured T_1 values 4 weeks after treatment and showed that T_1 was increased in the BDL rats but not in the CCl_4-administered rats[163].

Hoad et al[164] provided further evidence that presence of high LIC shortened T_1 relaxation times and that presence of hepatic inflammation in patients without advanced fibrosis lengthened T_1 times. Recently another study recommended that T_1 mapping imaging slices should be acquired in parts of the liver not adjacent to the lung, as liver T_1 measurements were found be significantly shorter by an average of 33 ms nearer the lung ($p < 0.001$)[165]. The

lower values were postulated to be due to air-induced susceptibility effects, suggesting another confounding variable or at least a positioning constraint on the acquisition technique.

1.4.9 $T_1\rho$ relaxometry

T_1 relaxation time in the rotating frame, 'spin-lock', or $T_1\rho$ is a specific form of T_1 mapping. In this technique, after longitudinal magnetisation is tipped into the transverse plane, a second pulse is applied, effectively locking the magnetisation into the transverse plane without phase decay. The $T_1\rho$ time is the time taken for this magnetisation to decay and is dependent on both T_1 and T_2 of the tissues, certain aspects of which can be amplified by acquisition parameters. The benefits of $T_1\rho$ are that like T_1 mapping it does not require a contrast agent

Two studies have shown $T_1\rho$'s ability to differentiate between Child Pugh (CP) class of liver cirrhosis[1], with values for CP-A, CP-B and CP-C being 45.4, 50.0 and 54 ms respectively at 1.5T[166] and values for CP-A and CP-B/C being 47.07 and 51.09 ms respectively at 3T[167]. Both these studies were also able to differentiate between healthy volunteers and any stage of cirrhosis with healthy volunteer values of and 39.37 at 3T. Another study has also examined healthy and fibrotic livers at 1.5T, showing the coefficient of variation (CoV) of $T_1\rho$ values between two repetitions in the same day session to be 0.83 ± 0.8 % and on different days to be 5.4 ± 2.7 %[168]. In this same study $T_1\rho$ correlated with METAVIR fibrosis stage ($r = 0.99$, $p < 0.05$), albeit in only 5 patients, with some showing significant inflammation which could be a confounding factor.

In a study investigating concurrent $T_1\rho$, T_2* and PDFF in 18 healthy volunteers, patients with fatty liver and patients with fibrosis apiece it was shown that mean $T_1\rho$ values were not

[1] Child-Pugh (CP) class A – compensated cirrhosis, CP class B – decompensated, CP class C – end stage.

significantly different (p = 0.999) between the healthy volunteers and those with fatty liver, indicating that it is unaffected by presence of steatosis[169]. In addition, $T_1\rho$ and PDFF values were different between the groups, but T_2* was not, but within the normal range so effect of iron overload on $T_1\rho$ could not be studied. $T_1\rho$ was able to predict liver fibrosis with AUC 0.855 with sensitivity 78% and specificity 100%. In this study no patients with both fibrosis and fatty liver were included, limiting the investigation of possible effects of fat on $T_1\rho$.

Confounding effects of inflammation on fibrosis estimation are also observed in $T_1\rho$ mapping as shown in a rat study[170]. This study showed significant increases in $T_1\rho$ in a CCl_4 administered rats but was shown that this increase could not solely be attributed to fibrosis as there was significant inflammatory presence but not evidence of fibrosis on histology in early assessment post CCl_4 administration despite an increase in $T_1\rho$. However, the $T_1\rho$ values did not return to normal after CCl_4 was removed, and this coincided with residual fibrosis in the rats. Another study of 52 rats with CCl_4 induced fibrosis showed that $T_1\rho$ had a similar AUC to the IVIM-derived true diffusion coefficients (D_{true}) for across all fibrosis stages (0.92), which was higher than T_2 and the ADC values alone (0.86 and 0.82 respectively)[171].

$T_1\rho$ is susceptible to B_0 and B_1 magnetic field inhomogeneities and is therefore an interesting emerging technique requiring further validation and assessment of confounding variables.

1.4.10 Dynamic contrast-enhanced studies

The injection into the bloodstream of a contrast agent, usually a chelated metal (e.g. gadoxetate, Gd-EOB-DTPA, can lower the measured tissue T_1. Certain contrast agents are taken up by hepatocytes and excreted into bile. Images are continuously acquired starting before injection and at certain times after injection. Viewing the 'wash in' or 'wash out' phases can yield improvements in detection of tumours and lesions. Since the contrast agent

is taken up by the hepatocytes specifically, DCE-MRI can assess liver function as well as severity of fibrosis/cirrhosis. Those patients with poorer liver function will retain more contrast agent for a longer period of time as their hepatocyte count may be smaller, or their hepatocytes may not be functioning as well as normal. Indices associated with DCE-MRI include the ratio of signal intensity (SI) to a reference SI such as that of spleen or muscle, and rate of uptake (k_1) which is the rate of agent uptake by the hepatocytes. Some contrast agents used can be unsuitable for those with renal insufficiency as the agent may be excreted via the kidneys.

Liver fibrosis severity

Juluru et al[172] showed in a prospective case-control study of 22 patients with various stages of fibrosis that hepatocellular uptake rate (k_1) was negatively correlated with fibrosis stage ($r = -0.55$, $p = 0.01$) and had AUC 0.84 for distinguishing severe (F3-4) from non-severe (F0-2) fibrosis. Noren et al[173] showed that liver to spleen contrast ratios (LSC) at 10 and 20 minutes post injection and k_1 could distinguish severe from mild fibrosis (AUCS 0.71, 0.80 and 0.78 respectively) in a cohort of 38 patients with varying aetiologies of CLD. Hako et[174] in a prospective cohort of 19 healthy individuals and 59 with chronic HCV/HBV showed that a combination of a composite score including liver surface changes, inhomogeneities and LSC ratio (MRI Triple test) after 20 minutes post injection could differentiate between mild and severe fibrosis (F1-2 vs 3-4) and correlated with fibrosis stage ($r = 0.86$, $p < 0.01$).

Liver function

Ning et al[175] showed that relative uptake rate was associated with bilirubin ($r = -0.52$, $p = 0.015$), pre-albumin ($r = 0.58$, $p = 0.0015$) and prothrombin time (PT) ($r = -0.51$, $p = 0.026$), indicating its ability to assess liver function. Another study showed that post contrast T_1

reduction rate was associated with indocyanine green (ICG) clearance rate (r = 0.86, p < 0.001)[176].

Preclinical studies

A multicentre preclinical trial aimed to assess the suitability of gadoxetate DCE-MRI to assess damage to the liver by drug induced liver injury (DILI) in 22 rats[177]. DCE-MRI parameters k_1 and k_2 (uptake to and efflux from hepatocytes) were significantly impaired with consistent effect size across all 4 centres by a clinical dose of rifampicin. Another study aimed to assess the correlation between fibrosis induced by thioacetamide and DCE-MRI parameters in 24 rats split equally into 3 groups (high dose, low dose and control)[178]. The initial AUC at 15 mins (iAUC-15), iAUC-3, relative liver enhancement at 3 minutes (RLE-3) and maximum enhancement E_{max} correlated best with fibrosis (r = −0.81, −0.78, −0.80, and −0.81, respectively), and had best diagnostic accuracy (100%) for any degree of fibrosis. Zhou et al[179] were able to show that DCE-MRI parameters were able to distinguish between all stages of fibrosis severity in a piglet model of induced fibrosis over a 22-week longitudinal study.

Direct comparison with other techniques

Although DCE-MRI has in general shown good performance for diagnosis of fibrosis and cirrhosis, a recent study by Keller et al[180] prospectively comparing DWI and DCE-MRI for the diagnosis and staging of hepatic fibrosis in 47 patients with primary sclerosing cholangitis (PSC) showed that DWI (ADC) had a higher diagnostic accuracy for detection of fibrosis F0/1 from F2/3 and F4 than DCE-MRI k_1. However, in this study fibrosis stage was determined by TE as a reference standard, not histology.

In a study by Dyvorne et al[137], 60 patients with CLD were prospectively assessed using IVIM, DWI, DCE-MRI, MRE and TE in a head-to-head study to determine the best predictor of

fibrosis. DCE-MRI was not available in 9 patients due to renal insufficiency, whereas MRE failed in 4, DWI in 3 (both mainly due to iron deposition) and TE failed in 2/48 (4.1%). DCE-MRI mean transit time (MMT) was moderately negatively correlated (r = -0.37, p with fibrosis stage but not inflammation grade. However, MRE was the only variable that correlated with collagen content and performed better than DCE-MRI parameters for diagnosis of fibrosis stage. This highlights the pitfalls of the contrast agent as well as indicating that DCE-MRI may not be as useful as a marker of fibrosis than MRE. Another study by Wagner et al[102] examined the ability of a DCE-MRI derived parameter, the liver upslope for the prediction of PH and CSPH. Liver upslope was shown to be no better than liver stiffness as assessed by MRE for either PH or CSPH.

DCE-MRI remains a useful technique, and is used for tumour detection already, but still requires further research into its diagnostic ability for fibrosis stage.

1.4.11 Texture analysis

Another interesting but as yet unestablished MR technique is texture analysis (TA) of MRI scans to diagnose fibrosis. TA involves statistical quantification of organ MR signal heterogeneity, the underlying theory being that as fibrosis advances, these statistical markers of signal heterogeneity also change. Some of the statistical parameters assessed are mean value, standard deviation (SD), skewness, entropy and kurtosis of the image signal, usually taken from an operator-defined region of interest. A recent study showed that SD and entropy of a T_1 weighted image was able to identify severity of fibrosis with AUC of 0.755 and 0.769 (significant fibrosis) and 0.746 and 0.754 (advanced fibrosis) respectively[181]. These results indicate only fair diagnostic accuracy, and so more work is needed in this field.

MR fingerprinting (MRF) is another technique that rapidly acquires T_1, T_2 and B_1 (magnetic field) maps in one breath hold. In MRF acquisition, the imaging settings are allowed to vary as if by random. The signals obtained are then matched by an algorithm to a template or pre-defined 'dictionary' of signal evolutions from all possible scenarios. Applying this technique in the abdomen is challenging due to respiratory motion, but one study has shown its early applicability in obtaining healthy volunteer T_1 and T_2 values from liver and spleen (among other tissues) that are similar to the literature in a 19 second breath hold[182]. MRF was also able to distinguish T_1 and T_2 differences in liver metastases from the surrounding liver parenchyma in 6 patients, indicating multiple uses in clinical practice.

1.4.12 Multiparametric MRI studies

With the field of liver MRI growing and the techniques therein advancing significantly, several studies have investigated the use of combinations of techniques to best diagnose disease or stratify patients by severity of risk.

More recent studies investigating liver (and spleen T_1) have yielded important information. Hoad et al[164] presented a decision tree model using liver T_1, T_2* and the ELF test to assess which patients needed to be referred for biopsy. In this study all patients with high iron (T_2* < 22.6 ms) were referred for biopsy for assessment of iron overload and the remainder assessed by firstly the ELF test and then liver T_1. The aim of this study was to accurately identify those with either fibrosis or inflammation and exclude those at low risk. Feier et al[183] took another multiparametric approach including ADC from DWI, signal intensity (SI) from susceptibility weighted imaging (SWI) and contrast enhanced T_1 mapping. This study showed that when combined, relative enhancement (RE), liver-to-muscle ratio (LMR) and ADC measurements could distinguish between all METAVIR fibrosis stages with excellent

diagnostic accuracy (≥ 90% for all). In a cohort of 60 patients with compensated cirrhosis (devoid of high liver iron content (LIC)), liver T_1 was shown to be predictive of liver-related clinical outcomes, and showed increasing liver T_1 ($p < 0.001$) across healthy volunteers (n = 40), patients with compensated cirrhosis and decompensated cirrhosis (n = 7)[184]. These changes in liver T_1 were accompanied by increases in liver volume ($p = 0.006$) and reductions in liver ($p <0.001$) and splenic perfusion ($p < 0.001$). Another study by Palaniyappan et al[117] examined the association of liver and spleen T_1 among other parameters with severity of PH in 30 patients undergoing HVPG measurements (47% had CSPH (HVPG ≥ 10mmHg). The investigators showed that liver T_1 correlated well with HVPG ($r = 0.835$, $p < 0.001$) along with splenic artery, superior mesenteric artery velocity and azygous vein flow, and that spleen T_1 correlated with HVPG across the whole patient group ($r = 0.40$, $p = 0.028$), but not in patients with CSPH. Decreases in liver T_1, T_2 and T_2* along with an increase in liver tissue perfusion have also been observed in the short-term (median follow-up 22 days) after treatment for chronic HCV in a study of 17 patients imaged pre and post direct acting antiviral (DAA) therapy by Bradley et al[185].

1.4.13 Iron-corrected T_1 (cT_1)

One such multiparametric approach is to use T_1 mapping, T_2* mapping and MRS/MRI for a holistic liver assessment of liver fibroinflammation, LIC and liver fat respectively. Up until the development of an algorithm developed by researchers at the Oxford Centre for Clinical Magnetic Resonance Research (OCMR) to correct for the confounding effects of high liver iron content (LIC) on measurement of liver T_1 values[186], studies had often tended to exclude those with high LIC. This algorithm measures iron using T_2* maps and transforms the acquired T_1 maps appropriately to counter the confounding effect of iron, yielding the iron corrected T_1 (cT_1).

The application of this algorithm to measurements of liver T_1 in a cohort of 79 prospectively recruited patients with CLD resulted in excellent diagnostic accuracy for diagnosis of any stage of Ishak fibrosis ($F \geq 1$) against no fibrosis and correlated with r = 0.68 against morphometric fibrosis. Since then, liver cT_1 has been shown in the form of a liver and inflammation (LIF) score from 0-4 to predict clinical outcomes in patients with CLD[187], with a negative predictive values of 100% for any patients with a LIF score < 2. Liver cT_1 has been shown to have excellent repeatability and reproducibility metrics[188], to have a tightly distributed range of values in a nominally healthy population[189], and has been used as a secondary endpoint in clinical trials investigating therapeutics for NASH[190]. It has also been shown to correlate with severity of steatohepatitis and differentiate between simple steatosis and NASH with an AUC of 0.80 and identify significant NAFLD with AUC 0.89[191]. Uncorrected liver T_1 was shown to increase significantly in rats administered with CCl_4 for 4 and 12 weeks and liver PDFF was shown to increase in rats fed a methionine and choline deficient (MCD) diet, with no decreases in liver T_2* in the CCl_4 treated rats, but a significant decrease observed in the MCD diet rats, indicating steatosis to have a confounding effect on T_2*.

Liver cT_1 is one of the imaging modalities being assessed in the large multicentre LITMUS trial (Liver Testing Marker Utility in Steatohepatitis) to identify the best non-invasive markers in the diagnosis of fibrosis in patients with NAFLD and in diagnosis of NAFLD vs NASH. Liver cT_1 was the primary outcome measure in a trial assessing liver health (minimum functional liver remnant (FLR)) pre and post liver resection surgery called the HepaT1ca study[192], and is being utilised in a multicentre NAFLD health economics trial (RADIcAL, ClinicalTrials.gov Identifier: NCT03165201) as well as others. Corrected T_1 mapping is also being assessed in multiple organs (liver, spleen, pancreas, kidney) for the holistic assessment of type 2 diabetes mellitus (NCT04114682).

The effects of iron and inflammation on liver T_1 assessment of fibrosis have been reported[156,164]. Out of these two confounding factors, liver iron has been accounted and corrected for by cT_1, but the effect of the inflammatory component on the T_1 signal cannot yet be separated from the fibrotic component.

Liver cT_1 is provided by Perspectum Diagnostics Ltd through their Liver*MultiScan*™ service and since these publications, the iron correction algorithm has been refined and another algorithm has been incorporated that accounts for variations in liver T_1 across magnetic field strengths and MRI scanner vendors by standardising to a reference scanner and field strength.

Liver and spleen cT_1 were the primary variables of interest in this thesis.

1.5 Aims

Previous work has shown liver cT_1 in the form of the LIF score to predict clinical outcomes in patients with CLDs over a relatively short follow-up period[187]. The LIF score has been abandoned and the algorithm to calculate liver cT_1 has been updated significantly. Further evidence of liver cT_1's prognostic value in larger cohorts over longer follow was therefore required. A landmark was achieved in the development of direct acting antiviral (DAA) treatments for chronic HCV, but their effect on disease regression is not well studied – therefore we aimed to assess the ability of liver cT_1 to quantitatively assess changes over the medium term. A pilot study from our centre had shown that spleen cT_1 correlated well with HVPG[193] and another study had shown spleen T_1 to be associated with HVPG[117]; another aim therefore was to investigate these findings further by collating MR data from multiple centres for the prediction of PH. The establishment of a normal range of spleen T_1/cT_1 was also necessary to assess the applicability to the wider population.

The aims of this thesis therefore were:

1. To assess the prognostic ability of the new, standardised liver cT_1 to predict clinical outcomes over longer time and in a larger cohort than previously reported
2. To assess the ability of liver cT_1 to track disease regression in patients undergoing DAA treatment for chronic HCV
3. To assess the ability of spleen T_1 and cT_1 to diagnose CSPH in multiple centres
4. To investigate spleen T_1 and cT_1 in healthy volunteers and participants with no known liver-related illness to establish a normal range of spleen T_1 and cT_1

Chapter 2: General Theory and Methods

2.1 Nuclear Magnetic Resonance theory

The phenomenon of nuclear magnetic resonance (NMR) and the subsequent techniques of spectroscopy (MRS) and imaging (MRI) are non-ionising, non-invasive, versatile and repeatable methods to probe for biological, chemical and physical information both *in vitro* and *in vivo*.

The quantum mechanical phenomenon of nuclear spin is at the heart of NMR. Atomic nuclei experience a nuclear spin if they have an odd mass number and/or an odd atomic number. The quantum mechanical nature of the nuclear spin can be explained by using classical mechanics. The nucleus is considered to be spinning on its own axis with constant velocity about a point and thus with constant angular momentum when not subjected to any external force. Since the nucleus carries an electric charge, and is 'spinning', a current loop and thus a magnetic field is formed. The defining attribute of this magnetic field is its magnetic dipole moment (μ), which is classically a product of the current and the area of the loop (Figure 1A). As a result, the nucleus can be thought of as a minute bar magnet.

The work presented in this thesis is limited to the investigations of hydrogen nuclei (protons) only.

Under normal conditions, the protons are randomly aligned (Figure 1B) , and the vector sum of their individual magnetic moments, i.e. the net magnetisation is zero. However, when an external, static magnetic field is applied (B_0), the protons become aligned in one of two ways in relation to the magnetic field: parallel or antiparallel (Figure 2A). There is a discrete energy difference between these two alignments.

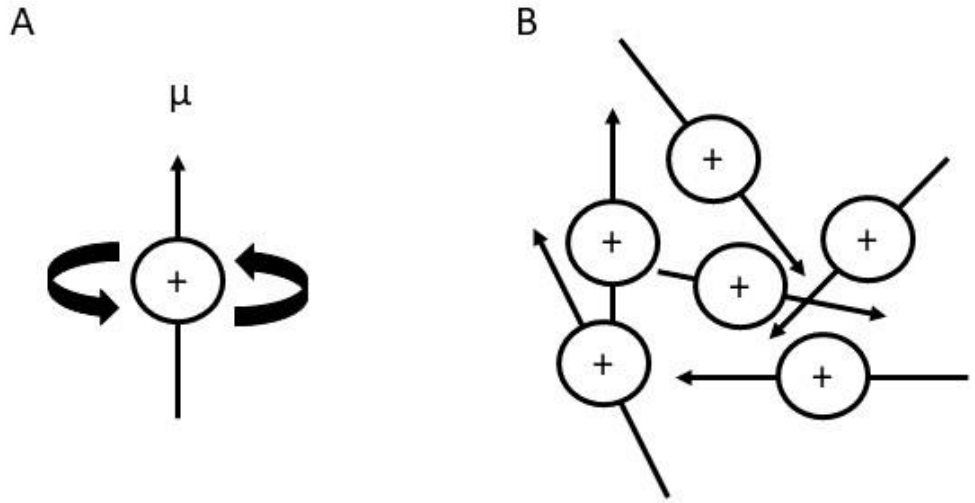


Figure 1: **Protons in the absence of an external magnetic field.** A) a proton revolving around its axis of rotation with magnetic dipole moment μ in the direction of the arrow. B) Random alignment of protons' magnetic moments in the absence of any external magnetic field.

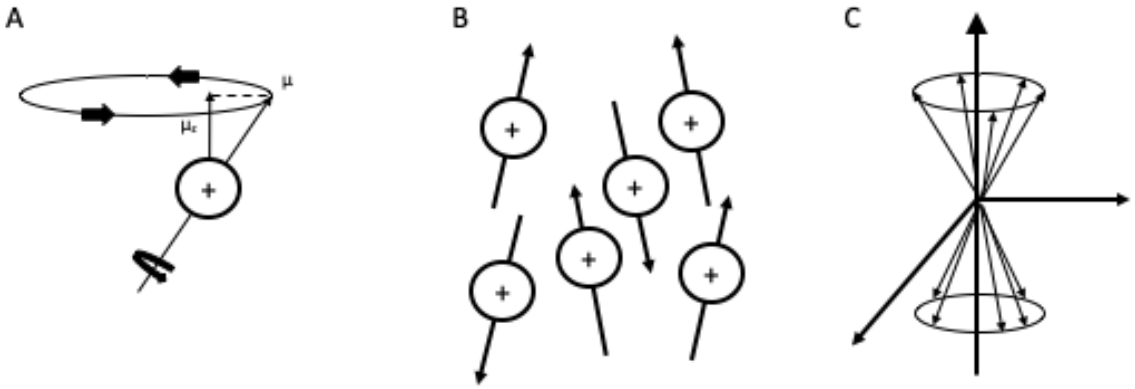


Figure 2: **Protons in the presence of an external magnetic field.** A) a proton's precession around its axis in the presence of an external magnetic field, with the contribution of the magnetic moment in the z direction depicted by μ_z, B) alignment of protons' magnetic moments in the presence of a static external magnetic field (B_0), C) contribution of the different protons alignments to the net magnetisation M_z.

The populations of the two different proton alignments in the presence of a static magnetic field are determined by the Boltzmann distribution[194]. The ratio of the protons at lower (parallel, $N_{\uparrow}$) and higher (antiparallel $N_{\downarrow}$) states are dependent on their energy difference ΔE and surrounding temperature T and are expressed in equation 1 below.

$$\frac{N_{\uparrow}}{N_{\downarrow}} = e^{-\frac{\Delta E}{k_B T}} = e^{-\frac{\hbar \gamma B_z}{k_B T}}, \quad (1)$$

where k_b is the Boltzmann constant. Equation 1 can be approximated by applying the Maclaurin series expansion and taking the first two terms yielding equation 2.

$$\frac{N_{\uparrow}}{N_{\downarrow}} \approx 1 - \frac{\hbar \gamma B_z}{k_B T} \quad (2)$$

The subsequent difference in protons in the parallel and anti-parallel states (n) is extremely small and therefore can be approximated to be equal. n is represented by equation 3 where N represents the total number of protons and Bz represents the contribution of the B field in the z direction

$$n \approx \frac{N \hbar \gamma B_z}{2 k_B T}. \quad (3)$$

In general, there will be more protons in the lower energy state, however at body temperatures, at which *in vivo* investigations are generally carried out, the difference in populations is minimal. Consequently, the net magnetisation (M_0) is very small.

Although the alignment of the two spin states described above may lead one to think that they are completely made parallel/antiparallel each individual proton is actually at a slight

angle as depicted in Figure 2B and the protons are distributed as in Figure 2C. When any of these protons is placed in the static B_0 field, it experiences a torque, therefore M_0 (as the simple vector sum of the moments) feels a torque, and precesses around the B_0 field (Figure 2A) with a specific resonant frequency (called the Larmor frequency). This is also the frequency required to excite the protons from the parallel state to the antiparallel state. M_0 can be visualised as a motionless vector in the same direction as considerably stronger B_0; this therefore makes its detection difficult without some sort of perturbation. There must be some component of the magnetisation in the transverse plane to be detected. To achieve this, an oscillating magnetic field B_1 is applied in the transverse plane with frequency of oscillation equal to that required to excite the protons from the lower to higher energy states. B_1 is applied as a pulse with frequency in the radiofrequency (RF) range (hereafter termed RF pulse).

The effect of applying B_1 can be thought of as 'flipping' M_0 into the transverse plane. The extent to which it is flipped can be described by the nutation angle (angle through which the magnetisation vector is rotated, also called the 'flip angle'), which depends on how long B_1 is applied for as well as its amplitude. The application of this B_1 field has two effects on the protons; firstly, the protons become more equally populated as the flip angle gets closer to 90°, secondly the protons attain coherence (are in phase), thereby generating transverse magnetisation.

The longitudinal and transverse magnetisation components can be written with respect to the flip angle θ as such:

$$M_z = M_0 \cos \theta, \qquad (4)$$

and:

$$M_{xy} = M_0 \sin \theta, \quad (5)$$

where M_z is the remaining longitudinal magnetisation, M_{xy} is the component now in the transverse plane.

The longitudinal magnetisation relaxes back towards the z-axis and the transverse component decays to zero after the B_1 pulse is removed. The recovery and relaxation time evolution of the longitudinal and transverse components of the net magnetisation vector are the solutions to the Bloch equations[59]. The Bloch equation describes the time evolution of the net magnetisation vector and includes terms for the precessional motion and the recovery of magnetisation:

$$\frac{d\vec{M}(t)}{dt} = \gamma\vec{M} \times \vec{B} - \frac{\vec{M}_x(t) + \vec{M}_y(t)}{T_2} - \frac{\vec{M}_z(t) - \vec{M}_0}{T_1} \quad (6)$$

The penultimate and ultimate components of equation 6 describe the dephasing of protons in the transverse plane (x-y) and the longitudinal (z direction) recovery of net magnetisation, respectively. These processes are called T_1 and T_2 relaxation.

Longitudinal, or 'T_1' relaxation back to the equilibrium state M_0 after a 90° excitation is described by solving equation 6 for Mz:

$$M_z(t) = M_0(1 - e^{-t/T1}) + M_z(0)\, e^{-t/T1} = M_0(1 - e^{-t/T1}), \quad (7)$$

where T_1 is the characteristic time constant (defined by the time at which point the magnetisation has recovered to 63% of the original M_z) unique to different types of compounds and their chemical environment and remembering that $M_z = 0$ at $t = 0$ (when the magnetisation is entirely in the transverse direction). T_1 indicates not only the time taken to return to M_0 but also the time taken for the protons to initially become magnetised by B_0. The mechanism governed by this time constant is defined by the thermal interactions between the nuclei and the surrounding molecular environment. The nuclei lose energy through dissipation, fall back to the lower energy state and align with the B_0 field. As a result, T_1 is highly B_0 dependent (it increases with increasing B_0). T_1 is also temperature dependent, but for *in vivo* studies as described exclusively in this thesis, body temperature is fairly homogenous and so this effect is negligible. The mechanism continues until the magnetisation has returned to equilibrium.

T_1 times of protons are highly dependent on their 'tumbling rate', which is in turn determined by their restrictedness/freedom of their movement. When the tumbling rate is close to the Larmor frequency, T_1 relaxation is most efficient and therefore shortest. Water molecules whose movement is restricted for example water bound to proteins or other large molecules have low tumbling rates. Water compartmentalised in tissue will have higher tumbling rates, and water that is unrestricted such as in blood will have the highest tumbling rates.

Magnetic field strength also affects T_1 relaxation times of macromolecules (e.g. proteins) and water in restricted pools (e.g. intracellular water). However, because unrestricted/free water molecules (e.g. blood) have tumbling rates that span a wide range of frequencies, they are not affected so much by an increase in field strength. On the other hand, restricted water molecules contain a much smaller percentage of protons that can tumble at the higher

frequency required by the higher field strength, resulting in longer T_1 times as depicted in Figure 3 that shows the spectral density functions of free, restricted and bound water molecules.

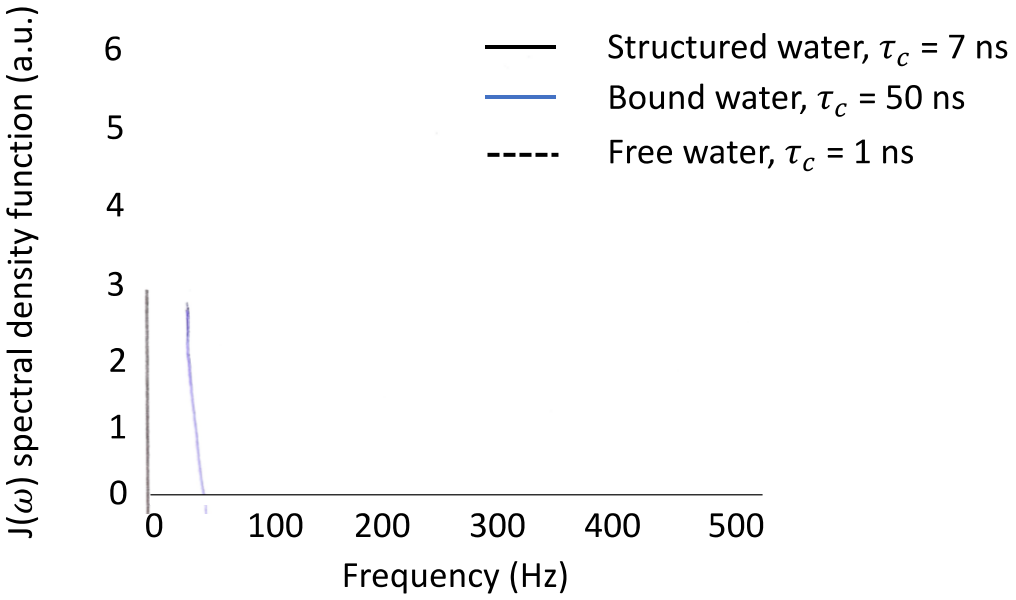


Figure 3: Spectral density functions of different water pools

The mechanism behind the decay of transverse magnetisation, or 'T_2' relaxation is dependent on the interactions between the excited protons themselves and governed by changes in entropy. As the protons interact, they fall out of phase with one another, destroying the transverse magnetisation as time passes. No energy is lost through this process. The mechanism is described by:

$$M_{xy}(t) = M_{xy}(0)\, e^{-t/T2}, \qquad (8)$$

where T_2 is the characteristic time constant. $T_2 \leq T_1$ since $|M(t)|$ cannot be larger than M_0. When the effects of magnetic field inhomogeneities are taken into account, the transverse

magnetisation decays more quickly, since localised regions of different field strength cause the protons to dephase faster and have a shorter characteristic time constant called T_2*.

T_1 times of protons (at fixed magnetic field strength and temperature) are dependent on their surroundings, in particular their chemical and electrical environment. In general, protons that are 'freer' i.e. are in continuous and more rapid motion, have longer T_1 (i.e. water in the cerebrospinal column) than those that are constrained by their physical environment (e.g. intracellular water or water bound to macromolecules).

2.2 Magnetic Resonance Imaging

In order to create an image from the NMR signal, one needs to know the precise locations of the different protons and therefore required making their resonant frequencies position dependent. This is done by employing magnetic field gradients that cause the static magnetic field B_0 to vary linearly with position. This allows a slab or 'slice' of protons to be excited, and those protons within the slice to be given their own position dependent phase. This allows an image or a 'map' of the signals from the different protons to be formed.

2.2.1 T_1 mapping

T_1 mapping can be performed using a number of techniques[195,196], and in this thesis the shortened modified Look Locker Inversion recovery (shMOLLI)[197] sequence was used. shMOLLI samples the T_1 relaxation curve after a 180° excitation, called an inversion, used to flip the longitudinal magnetisation. An adiabatic hyperbolic secant inversion pulse is used in order to minimise effects of B_0 field inhomogeneity.

The longitudinal magnetisation is inverted a certain number of times during the acquisition process. Each inversion is followed by a certain number of balanced steady-state free precession (bSSFP)[198] readouts each. The shMOLLI sequence can adopt many 'schemes' which differ by the number of acquisitions and/or time between inversion pulses. An inversion pulse followed by 5 bSSFP readouts followed by two sets of 1 RR period followed by 1 bSSFP readout would be termed a 5(1)1(1)1 scheme (Figure 4), with the numbers outside brackets referring to the number of bSSFP acquisitions and the bracketed numbers referring to the recovery period between TIs. An exponential curve is fitted to those points, estimating the T_1.

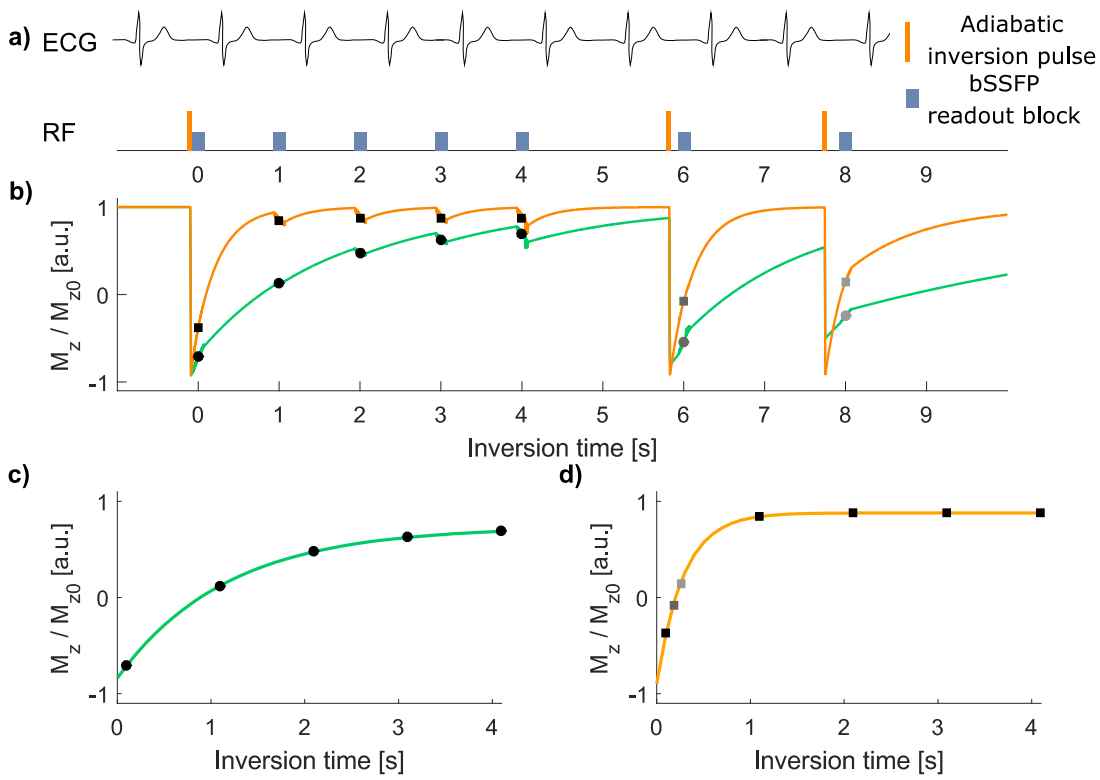


Figure 4: Simplified shMOLLI pulse sequence diagram. Adiabatic inversion pulses precede bSSFP readout blocks (comprising series of 35° pulses), which are synced to end-diastole a). The bSSFP readouts sample the time evolution of M_Z b). Green line represents a tissue with long T_1, and orange represents tissue with a short T_1. Individual readouts are re-ordered according to inversion time and are fitted to a monoexponential recovery curve. The conditional fitting algorithm fits five points for long T_1s c) and all seven points for short T_1 fittings (d). Fig reproduced from Mozes 2018[199] with permission.

shMOLLI is electrocardiogram (ECG) gated and is performed in a single breath-hold to reduce cardiorespiratory effects – acquisitions occur in a window between RR intervals (time between continuous R-waves). This fact that shMOLLI is cardiac-gated is a result of the sequence originally being meant for use in cardiac imaging. It has been repurposed for liver imaging, so retains the ECG gating, but in general this is not strictly necessary to image the liver. The original MOLLI sequence from which the shMOLLI was derived was shown to be affected by heart rate[200]. This heart rate dependence had an average error of 9%, which was reduced to a constant 4% error in underestimation of T_1 times using the shMOLLI sequence[197].

An SSFP pulse is formed by multiple RF pulses in quick succession over a constant TR that is less than T_2 of the tissue being imaged. During the short TR interval, free induction decays (FIDs) and echoes overlap creating a magnetisation that never totally disappears. When the scanner gradients' moment total zero over a whole TR, a balanced SSFP (bSSFP) sequence is achieved. No dephasing over a TR interval results in the frequency dependence of a bSSFP pulse as shown in Figure 5. Lastly, when the TE is exactly half of the TR, the bSSFP sequence takes on a spin-echo style sequence (Figure 6).

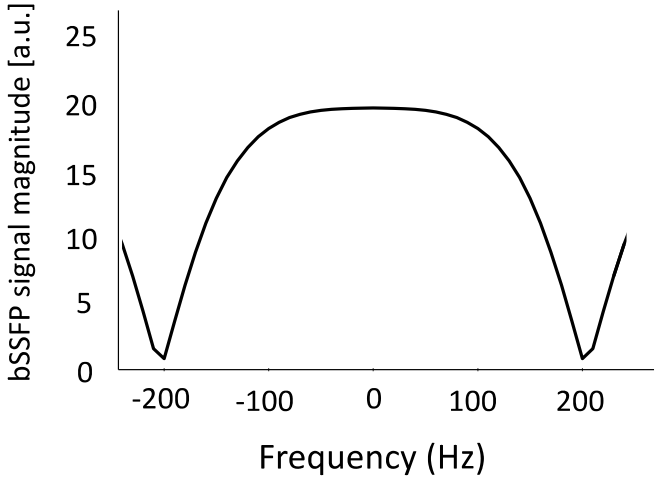


Figure 5: Frequency dependence of the bSSFP signal with TR = 2.3 ms and TE = 1.15 ms.

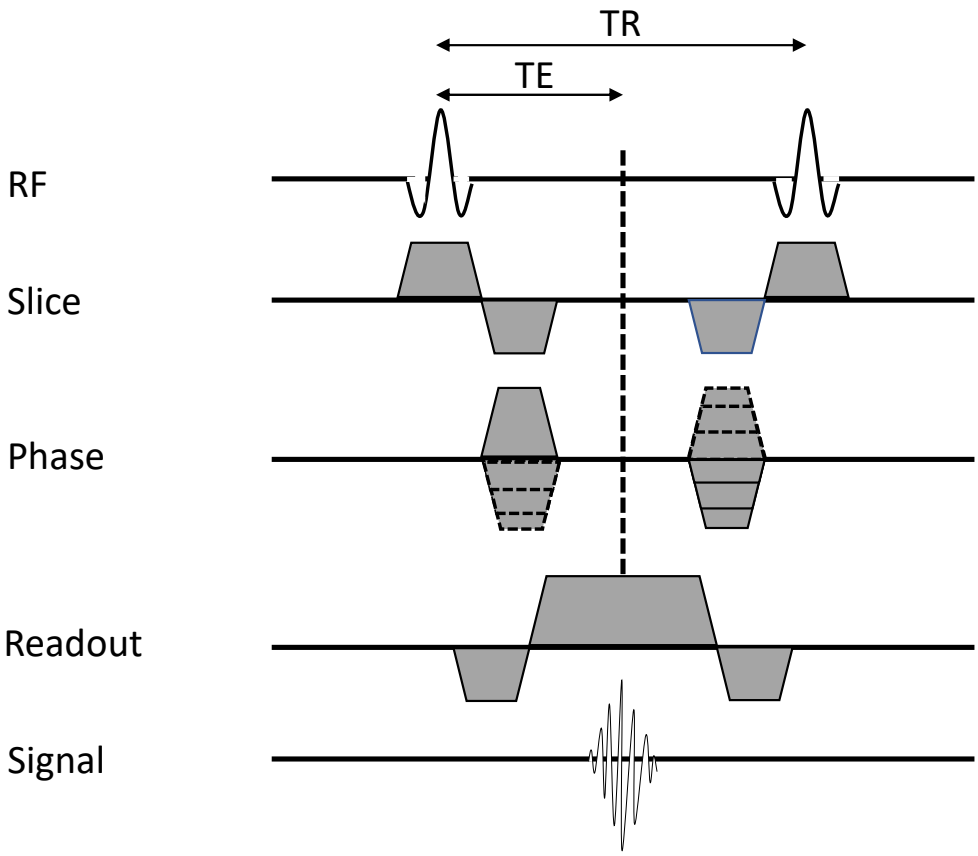


Figure 6: One TR interval of balanced steady-state free precession sequence, with a symmetric readout.

BSSFP sequences establish a coherent steady state whose maximum signal intensity depends on the T_2/T_1 ratio of the imaged tissue. For example, fat and blood have high T_2/T_1 ratios and therefore appear bright on bSSFP sequences, whereas liver tissue has a lower ratio so will appear darker in comparison. Greatest contrast would therefore be achieved at a 90° readout, but for shorter scan times often flip angles of 45° or less are used; in this thesis 35° were used.

Sequences that employ a bSSFP readout are frequency-dependent and therefore are strongly influenced by local magnetic field inhomogeneities.

Liver fat is a major confounder of liver T_1 values, when measured using the shMOLLI sequence and can lead to artificially elevated T_1 values in patients with fatty liver. Fat has a shorter T_1 than water and a resonant frequency of Hz. The chemical shift difference between water and

the main fat peak is 3.4 ppm, which at 3T represents a frequency of 447 Hz. If the TR of the shMOLLI scheme used is close to 1/447 or approx. 2.23 ms, and TE half of that, fat and water will be completely out of phase. At a different TRs the phase overlap between the bSSFP pass bands of fat and water is different and the signal is a complex sum of the water and fat signals (Figure 7). This leads to a measured T_1 that can be longer than the individual water and fat components.

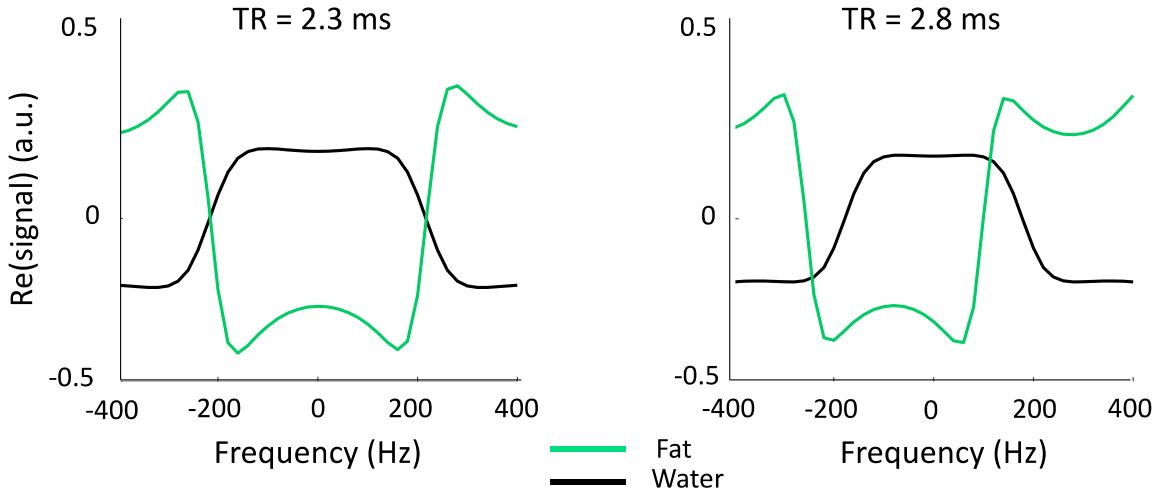


Figure 7: Simulated off resonance dependence of the bSSFP signal. Water has T_1 900 ms, T_2 30ms, main fat peak has T_1 320 ms, T_2 60 ms and chemical shift difference 3.4ppm. TE = TR/2. Fig adapted with permission from Mozes 2018[199].

In the original MOLLI scheme, all acquisitions were pooled by their effective TI and fitted to a monoexponential recovery curve. In a 5(1)1(1)1 shMOLLI scheme, bSSFP readouts are fitted using a conditional fitting algorithm. If the T_1 of the tissue is short ($<0.4T_{RR}$) all 7 points are used, if $0.4T_{RR} < T_1 < T_{RR}$ the first 6 points are used, and if the T_1 is longer, then only the first 5 points acquired are necessary. In practice since liver T_1 are on the longer side, only fitting the first 5 points is usually necessary.

2.2.2 T_2* relaxation

T_2* maps are acquired using a multi-echo spoiled gradient echo (GRE) sequence. These maps are used to quantify iron concentration. Proton density fat fraction (PDFF) maps can be extracted from multiple-echo GRE images using the DIXON[71] or IDEAL[201] methods.

Multi-echo GRE acquisition was used with parameters: TR: 500 ms, TEs: 1.25, 2.46, 3.69, 4.92, 6.15, 7.38, 8.61, 9.84 ms, 6 mm slice thickness, GRAPPA acceleration factor 2, flip angle: 20°. DIXON method used for fat fraction analysis. FOV: patient dependent, matrix size 104 x 128.

The echo times (TEs) were chosen to correspond to times at which the combined signal from fat and water were in and out of phase so that proton density fat fraction maps could be generated. Maps with fat and water "in-phase" and "out-of-phase" can be added and subtracted to derive the fat map of the liver only. As described earlier in 2.2.1, the TEs are chosen based on the chemical shift δ of 3.4 ppm at 3 T, which corresponds to 447 Hz = 2.23 ms TR and that TE = TR/2. We then have initial TE_{out} and TE_{in} of 1.15 ms and 2.3 ms.

2.2.3 Iron-corrected T_1 (cT_1)

T_1 is a measure of water content, of which the extracellular component rises with increasing tissue fibrosis and inflammation. T_1 is extensively used in cardiac imaging as a surrogate marker of fibrosis and inflammation. The liver however is a major store of iron, and liver iron content (LIC) can reach especially high levels in liver disease. Iron is paramagnetic and iron deposition causes rapid dephasing of the transverse magnetisation. This leads to a significantly shorter tissue T_2 and T_2*. It has been shown that the shMOLLI T_1 mapping sequence has a T_2 dependence[202]; therefore, the *measurement* of tissue T_1 by shMOLLI is affected by T_2 and T_2* values. High iron with short T_2* results in an artificially lower measured

T_1. A model for correcting this confounding effect of iron on measured T_1 values was developed by Tunnicliffe *et al.*[186] and is applied via the Liver*MultiScan*™ software (Perspectum Diagnostics Ltd (PD), Oxford, UK) to obtain the T_1 value that would be measured if the liver iron concentration was that of a normal value.

This correction models the competing effects of increasing fibrosis (represented by increasing extracellular fraction) with increasing iron (decreasing T_2*). Look-up tables are generated that allow the calculation of 'iron-corrected' T_1 values given an input of T_2* and measured T_1.

Liver*MultiScan*™ now also applies an algorithm to the T_1 maps that allows standardisation across field strength and vendor. Differences in the MRI sequences used on different platforms produce systematic differences in fitted T_1 values. These quantitative differences were resolved by using phantom measurements to generate linear mapping functions to standardise the values obtained on one system to those from another at the same magnetic field strength (1.5T or 3T). All 3T systems were linearly mapped to the Siemens Prisma 3T, and all 1.5T systems to the Siemens AvantoFit 1.5T. The field strength correction is an almost linear correction based on mapping values between the two reference scanners at 1.5T and 3T. It is worth noting that as field strength increases, the precision of the T_1 fitting changes. This may therefore also affect the precision of the final corrected T_1 and highlights the complexities involved in scaling T_1 values with field strength.

The combination of the iron correction and standardisation algorithm yields the corrected T_1 (cT_1). Refinement of the iron correction algorithm and development of the standardisation algorithm both took place after the initial publications by Banerjee[203] and Pavlides[187].

2.2.4 Image analysis

T_1 and T_2* maps were analysed using Liver*MultiScan*™ for the purpose of generating and analysing cT_1 maps by the following steps:

1) Three regions of interest (ROIs) were manually placed on T_2* colour maps to obtain a representative iron concentration within the organ of interest (either liver or spleen).
2) This iron concentration was fed into an algorithm that transforms the measured T_1 map into a cT_1 map in such a way that corrects[186] for any underestimation caused by increased iron concentration.
3) EITHER
    a. 3 ROIs were placed on the cT_1 map over parts of the liver/spleen that were considered representative of the organ. This meant avoiding blood vessels, organ boundaries and potential cysts/ascites that were present. A single ROI measurement calculated the mean of the cT_1 pixel values within it; the mean of three ROIs was taken as the final cT_1 value. (Technique used in Chapter 3 as this was current PD reporting method and analysis was performed according to PD's current protocol at the time)

    OR

    b. On the cT_1 map, the organ of interest was segmented semi-automatically, removing the blood vessels and other abnormalities (e.g. cysts/ascites). The mean value of cT_1 within the organ was calculated and taken as the final cT_1 value. (Techniques used in Chapters 4, 5 and 6).
4) An output report for visualisation and MRI metrics was produced (Figure 8), and numerical metrics could also be parsed straight into a Microsoft Excel spreadsheet.

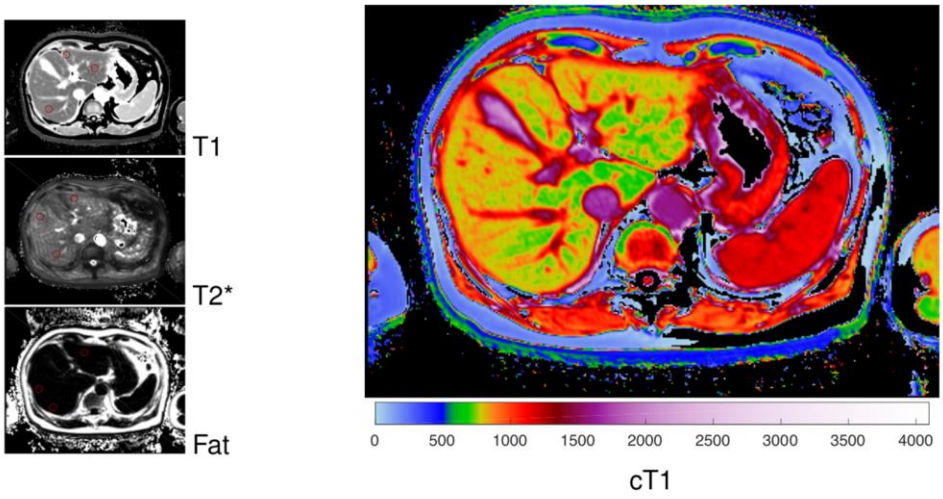
As part of fitting data to T_1 and T_2* maps, R^2 maps are also generated which are a measure of the goodness of fit of the signal intensity to the relaxation curve. When placing ROIs in the T_1 and T_1* maps, mean R^2 values were also calculated. T_2* map R^2 values had to be > 0.95, and T_1 map R^2 values had to be > 0.99 for all liver measurements to be considered valid.

A)

Statistics Summary

Region 2

Fat: **0.9 %**	Normal range: <5.6% [1]
Iron: **1.0 mg/g dry weight liver**	Normal range: <1.8mg/g [2]
cT1: **787 ms**	Reference interval: 632ms - 795ms [3]



T1 image quality: **Satisfactory**
T2* image quality: **Satisfactory**
Fat image quality: **Satisfactory**

B)

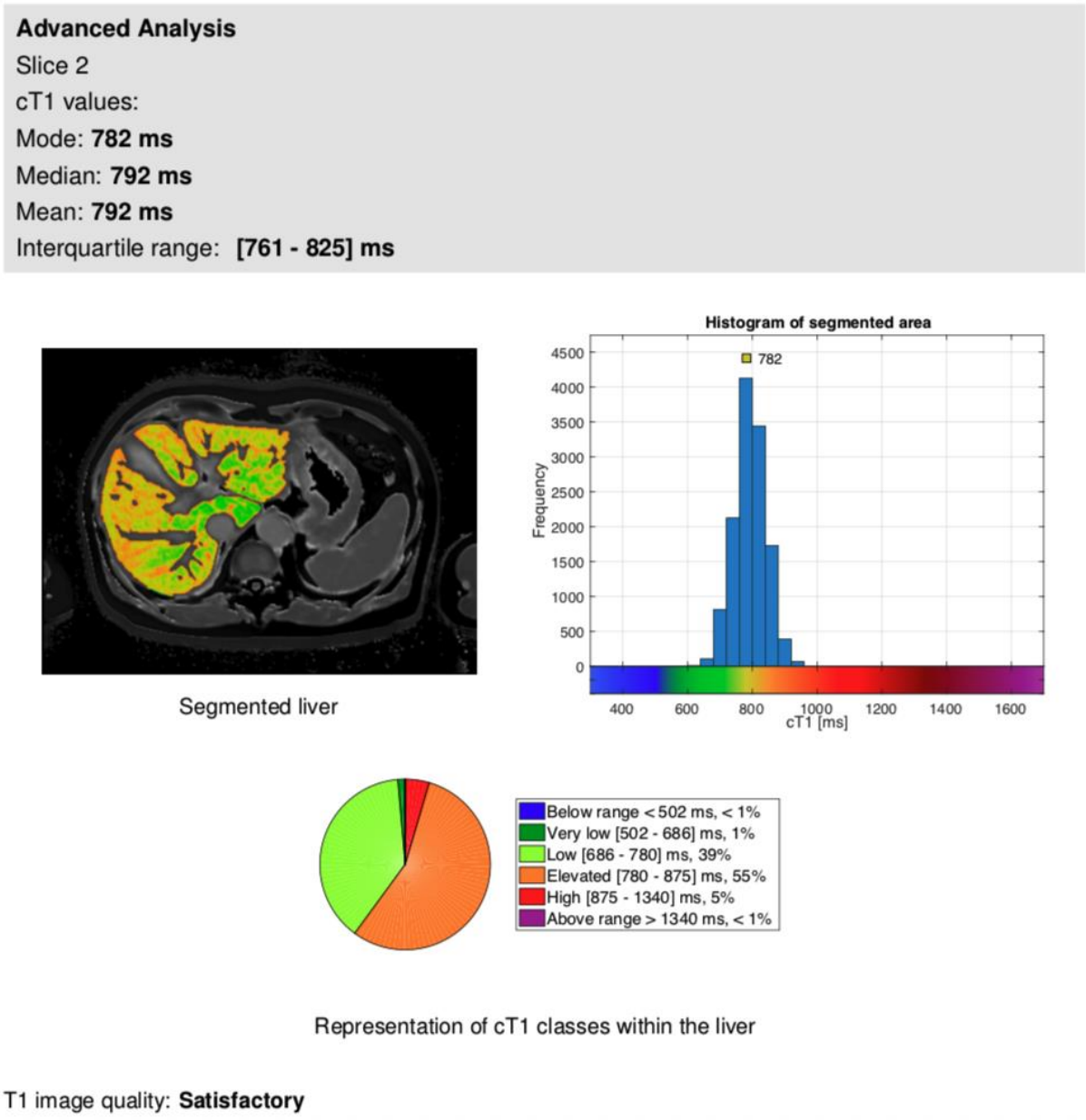


Figure 8: Example of a Liver*MultiScan*™ report and associated metrics. A) Report output using region of interest (ROI) analysis method. B) Report output using liver segmentation analysis method. Patient aged 69 with histologically confirmed Ishak fibrosis stage 1 (mild fibrosis).

2.2.5 United Kingdom Biobank (UKBB) imaging protocol

The UKBB population health study collects data from 500,000 participants and has a magnetic resonance imaging sub study aiming to collect data from a subset of 100,000 participants,

making it the world's largest imaging study. These participants are chosen by proximity to the designated UKBB imaging centres (Cheadle, Stockport, UK). The imaging sub study includes brain, heart and abdominal MRI[189]. MRI data from the UKBB are used in Chapter 6 and described here are the MRI parameters used for T_1 and T_2* mapping.

T_1 mapping: shMOLLI 5(1)1(1)1 sequence with scan parameters: Trigger delay set to 50 ms, first TI: 170 ms, with five increments of 50 ms each, TR = 4.8, TE: 1.93 ms, slice thickness 8.0 mm, flip angle: 35° and FOV: 440 x 330 mm^2.

T_2* mapping: GRE scan with parameters: TR: 27 ms, TEs: 2.38, 4.76, 7.14, 9.52, 11.90, 14.28, 16.66, 19.04, 21.42 and 23.80 ms, 6 mm slice thickness, GRAPPA acceleration factor 2, flip angle 20° and 6°, FOV: 400 x 400 mm^2.

All scans are carried out at 1.5T, and all maps were taken from the same transverse slice through the abdomen, obtaining good visual representations of both the liver and spleen.

2.3 Magnetic Resonance Spectroscopy

2.3.1 Theory

The Larmor frequency of protons is affected by their chemical environment. This induced change in the Larmor frequency is called the chemical shift δ (δ has units of ppm (parts per million)) and occurs when the electronic properties of the chemical compound screen the nuclei from the external B_0 field, causing the nuclei to experience an altered field strength, and thus a different resonance frequency. The consequence of the above is that hydrogen nuclei in different chemical environments each have different resonant frequencies. Relative quantification of the contributions from these various proton-containing compounds can be achieved by measurement of the signal from each. The compounds of interest were those

lipids that contribute to overall liver fat. At 3 Tesla magnetic field strength the resonant frequencies of water and methylene (main lipid) are 4.7 and 1.3 ppm respectively. The proton density fat fraction (PDFF) can be calculated by comparing the relative intensities of the fat and water peaks of a spectroscopic acquisition.

2.3.2 Acquisition and protocol

Localiser images in the sagittal, coronal and transverse planes were performed to guide the placement of the voxel used to acquire spectroscopy data. The voxel was placed avoiding biliary structures and blood vessels (Figure 9). In patients who had repeated scans, a screenshot of the location of the voxel was taken to enable the voxel to be placed at the same location for all follow-up scans.

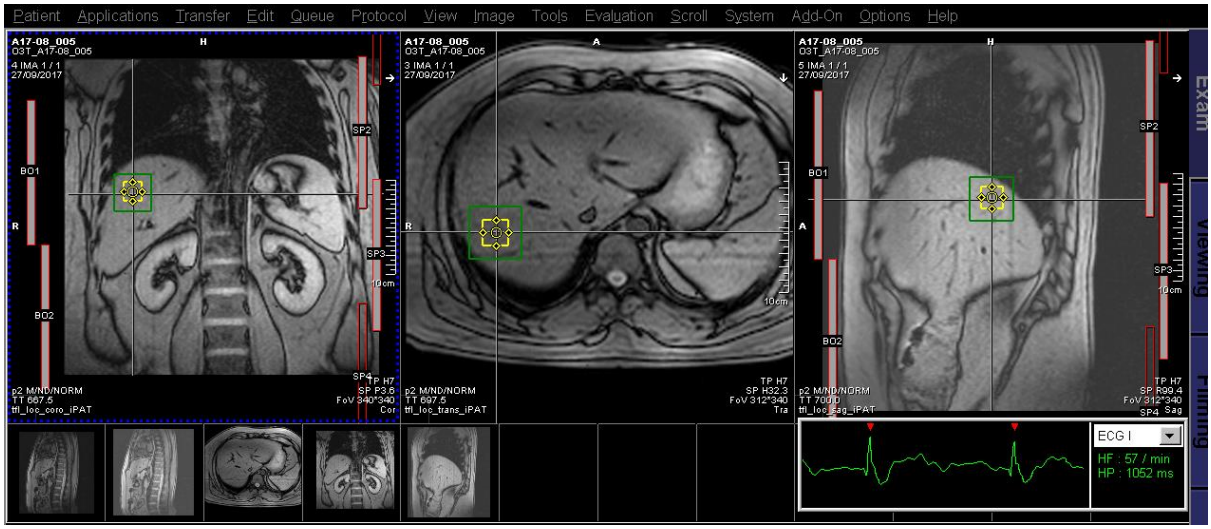


Figure 9: Placement of voxels for proton spectroscopy (MRS) to quantify liver fat fraction.

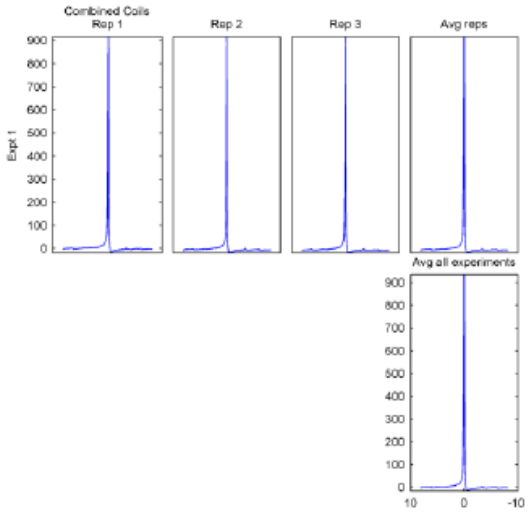
A STimulated Echo Acquisition Method (STEAM)[204] sequence was used with the following parameters: voxel size 20 x 20 x 20 mm^3 (shim volume 40 x 40 x 40 mm^3), TE = 10 ms, mixing time = 7 ms, TR: 2 s for water-suppressed and TR: 4s for unsuppressed data with 1024 points acquired at bandwidth 50 Hz. 3 scans of 5 repetitions with water suppression were performed to obtain good SNR, and 1 scan of 3 repetitions with water suppression turned off was performed to capture the amount of water in the voxel of interest. The sequence is a breath-hold sequence, therefore entails 4 breath-holds in total.

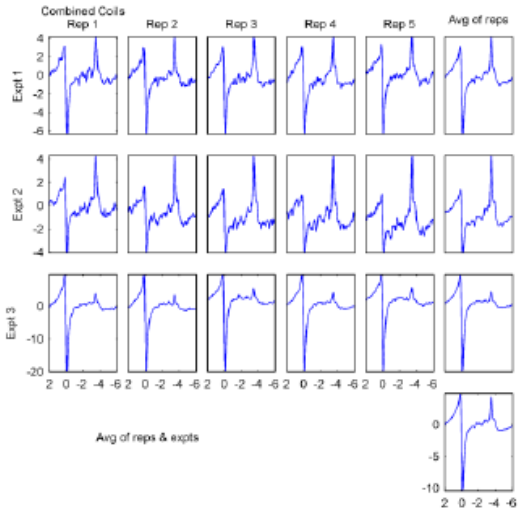
2.3.4 Analysis

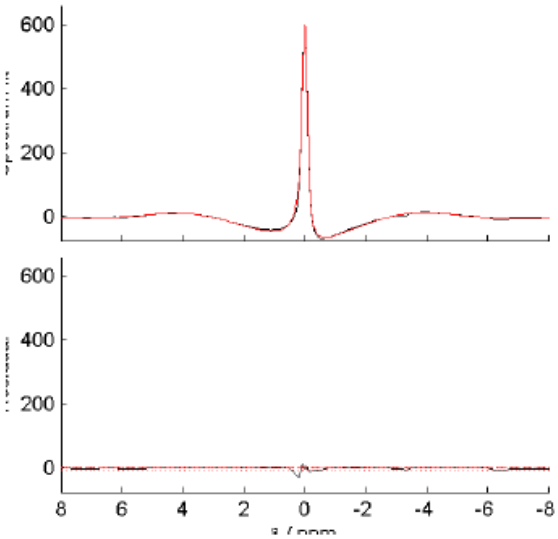
Spectroscopic acquisitions were combined and fitted using the OXSA toolbox[205] implementation of the AMARES algorithm[206] with an in-house MATLAB (The Mathworks, Natick, MA, USA) script[207]. Examples of output reports are shown in Figure 10.

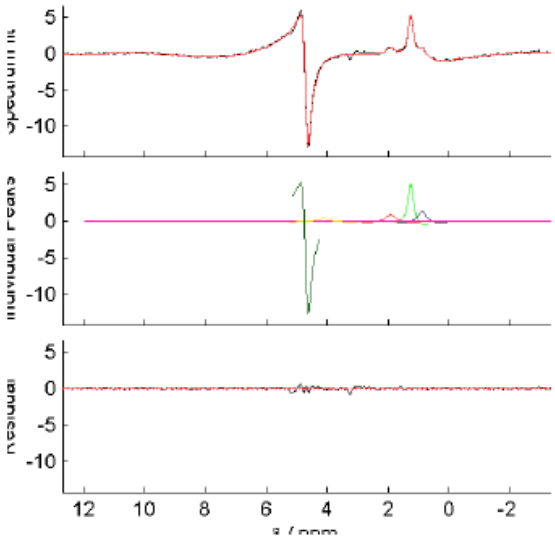
A)

PDFF	L/W CRLB	Residual/Water	R/W CRLB	Fit Residual	Fit Residual
%	%	%	%	(Unsup)	(Sup)
1.75	± 4.89	2.56	± 0.86	0.01	0.02









B)

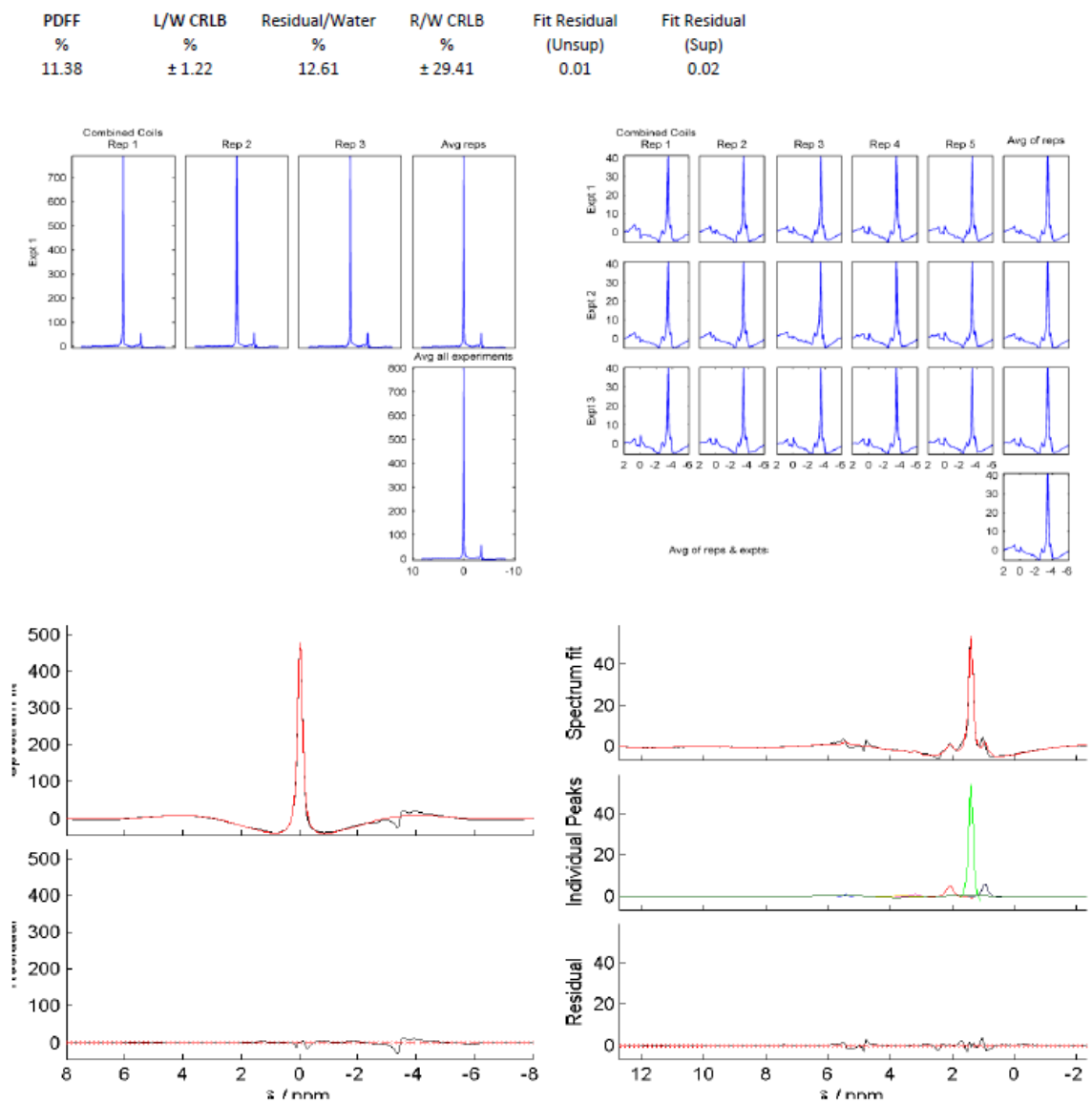


Figure 10: Spectroscopy tool output detailing proton density fat fraction (PDFF). Cramer-Rao lower bounds (CRLB) and representations of acquired spectra are shown in A) a patient with normal liver fat fraction and B) a patient with elevated liver fat fraction.

Liver fat was expressed as a ratio of the sum of the 6 lipid peaks in the spectrum to the sum of the 6 lipid peaks plus the water peak as shown in Figure 10.

Presence of iron affects both liver T_1 and liver T_2. The T_2 of water is shorter than the T_2 of lipid peaks, giving rise to faster decaying of the water signal than the lipid signal. In addition to this, the presence of iron will also shorten the measured water T_2 more so than the lipid T_2, since iron does not mix with fat[208]. These effects can lead to overestimation of the fat fraction of the liver[209]. Therefore, a correction for the presence of iron (determined by the T_2* value) calculated from the literature[82,210,211] was used to obtain more accurate fat fraction values.

2.4 Scanning practicalities

All patients and healthy volunteers that were scanned at OCMR during the course of this thesis underwent scanning at 3T using a Siemens Tim Trio (Erlangen, Germany). Patients from collaborations where the scans took place at other institutions have their scanning conditions described in the relevant chapters.

Different studies took place at different field strengths. There has been patient data to validate combining liver T_1/cT_1 measures across field strength[188], but no data to validate combining spleen T_1/cT_1 between field strength.

All patients were scanned having fasted for at least 4 hours to avoid the postprandial 'meal effect' which increases the measured T_1 due to increase in blood flow water to the gut. For the same reason, patients were also asked to not drink water up until 1 hour before their study assessment. In addition, depending on how long one is scanned after a meal, there can be increased concentration of glycogen which will reduce the T_1 measurement[212].

Scanning was performed either by a trained radiographer or by the author whilst supervised by a trained scanner operator.

2.5 Transient elastography (TE)

Liver stiffness was measured by TE (Fibroscan™, Paris, France). Transient elastography provides a one-dimensional image of liver stiffness[213]. Shear waves are created by a hand-held probe that is placed on the patients' right-hand side of the abdomen. This shear wave's speed of propagation is then imaged or 'tracked' as it passes through the body via an ultrasound beam. From the assessment of the shear wave speed, and the Young's Modulus (E) is calculated using the relationship (Hooke's law) between velocity and density;

$$E = 3\rho V^2 \tag{9}$$

where ρ is density of tissue and V is velocity of the shear wave. Transient elastography assumes that the tissue is homogenous, isotropic and purely elastic. TE has been shown to correlate with increasing fibrosis stage(ref) and with risk of clinical outcomes[214,215].

Since TE is a measurement of the viscoelastic properties of the liver, any variable that affects those properties can subsequently affect the stiffness measurements. These confounding variables include inflammation[216], cholestasis[217] and the effect of ingesting a meal[218,219], amongst others. TE has also been shown to be dependent on operator experience and studies have shown that between 100 and 500 examinations may be required to be considered 'experienced'[41,220].

All TE examinations were performed by a suitably trained operator. A minimum of 10 valid readings were required and the median value was taken from these. Unreliable readings were

defined as having an IQR/median > 0.3, as per the manufacturer's guidelines. The success rate was defined as the number of successful readings divided by the total number of attempts.

2.6 Blood sampling

Blood samples obtained at OCMR were drawn on the same day as the MRI visit unless otherwise stated. Samples were tested for Aspartate aminotransferase (AST), alanine aminotransferase (ALT), albumin, alkaline phosphatase (ALP), bilirubin, gamma-glutamyl transferase (GGT), lipid profile (triglycerides, cholesterol, high density lipoprotein (HDL), low density lipoprotein (LDL)), iron profile (ferritin, transferrin, transferrin saturation, serum iron) prothrombin time (PT), platelet count, and analysed by the Oxford University Hospitals Foundation Trust Haematology and Biochemistry laboratory.

2.7 Statistical analysis

All statistical analysis was performed and all plots generated by the author of this thesis using R Statistical software[221] and RStudio. Specific statistical analysis performed is detailed in each chapter.

Chapter 3. Prognostic ability of liver cT_1

The work in this Chapter has been submitted as a paper to:

Jayaswal A, Levick C, Selvaraj E, Dennis A, Collier J, Cobbold J, Booth J, Tunnicliffe E, Kelly M, Barnes E, Neubauer S, Banerjee R, Pavlides M. Liver cT_1 predicts clinical outcomes in patients with chronic liver disease: Further evidence from 754 years of follow-up.

And has been presented as abstracts at the following:

Jayaswal A, Pavlides M, Banerjee R, Kelly M, Barnes, Neubauer S, (2019). Liver cT_1 predicts clinical outcomes in patients with chronic liver disease. Journal of Hepatology, 156, S1325 - S1325.

Jayaswal A, Pavlides M, Banerjee R, Kelly M, Barnes, Neubauer S, (2018). Liver cT_1 can predict clinical outcomes in patients with chronic liver disease, equivalent to biopsy. Hepatology 68(S1), 607A–608A.

3.1 Abstract:

Background & Aims: Liver cT_1 has been previously reported as a predictor of clinical outcomes over 27 months. The aim of this study was to examine liver cT_1's prognostic value over extended follow-up of a larger cohort.

Methods: Patients with unselected liver disease aetiologies had baseline MR assessments and were followed up for development of clinical outcomes by review of their medical records, blinded to the imaging results. The pre-specified clinical endpoints were all-cause mortality, liver-related mortality, and liver events (ascites, variceal bleeding, hepatic encephalopathy, hepatocellular carcinoma, liver transplantation and liver mortality). Liver cT_1's prognostic value was assessed for mortality and liver event-free survival.

Results: Two hundred and eighteen patients (61% male) with median age of 54 years were followed up for 754 patient-years (median (IQR) 42 (23-57) months). The main diagnoses were NAFLD (41%), viral hepatitis (VH; 25%) and alcohol related liver disease (ArLD; 14%). During follow-up 23 new liver events, 13 liver-related deaths and 7 non-liver deaths occurred.

In 201 patients with valid MR scans, liver cT_1 > 825 ms predicted all-cause mortality (p = 0.003), liver-related mortality (p = 0.008) and liver event-free survival (p = 0.005). For liver event-free survival, cT_1 > 825ms had a hazard ratio of 4.9 (95% CI: 1.4-16.5) and correctly classified 19/22 (86%) liver events. In NAFLD/ArLD/VH patients (n = 164), cT_1 > 825ms correctly classified 18/18 liver events (p = 0.0009).

Conclusions: We provide further evidence that liver cT_1 can predict clinical outcomes in a larger patient cohort over longer follow-up and performs well in patients with NAFLD, ArLD and VH.

3.2 Introduction

Chronic liver diseases (CLDs) affect an estimated 1.5 billion people worldwide[1], with the predominant causes being non-alcoholic fatty liver disease (NAFLD)[3], chronic viral hepatitis (VH) B and C[4] and alcohol related liver disease (ArLD)[5]. Liver fibrosis is the final common pathway of injury in CLDs and over time accumulation of scarring can lead to liver cirrhosis. CLDs are reported to induce cirrhosis in approximately 633,000 patients per year in the United States[6]. Once fibrosis progresses to cirrhosis, each year approximately 5-7% will become decompensated[7] (develop variceal bleeding, ascites, hepatic encephalopathy). Cirrhosis is the 11th most common cause of mortality worldwide[8] and liver cancer is reported as the 4th most common cause of cancer-related mortality[9].

In the face of this public health epidemic there is an unmet clinical need to stratify disease severity in patients with CLD and to flag patients at risk of decompensation and development of hepatocellular carcinoma (HCC) as early as possible.

Liver biopsy is used routinely in clinical practice to assess fibrosis stage and inform prognosis[222], but is unsuitable for routine longitudinal follow-up of patients due to its invasive nature, associated risk and cost.

The magnetic resonance imaging (MRI) parameter T_1, when corrected for iron content (cT_1)[186] can quantify extracellular water content, which rises with fibrosis and inflammation. Liver cT_1 correlates with liver fibrosis[203] and severity of steatohepatitis[191], has excellent repeatability and reproducibility[188], and is used in the UK Biobank population health study as the reference for liver fibroinflammatory disease[189]. Liver cT_1, when stratified into groups according to a liver inflammation and fibrosis (LIF) score, was found to predict clinical outcomes in a general hepatology outpatient setting[187]. Liver T_1 has also been shown to be associated with

development of clinical outcomes in other cohorts[184,223]. The LIF score is no longer utilised by Perspectum since the FDA has preferred the use of liver cT_1 and refinement of the algorithm used to generate liver cT_1 now enables it to be standardised across field strengths and vendors. Pavlides et al.'s study[187] had a relatively short follow-up time and small cohort. For these reasons, longer term follow-up of larger patient cohorts was needed to corroborate these findings and validate cT_1 as a surrogate endpoint for clinical outcomes, to allow confident stratification of patients by prognosis.

3.3 Aims

The primary aim of this study was to assess the prognostic value of liver cT_1 in a larger cohort of patients with CLD, over longer follow-up. Secondary aims were to assess the relative prognostic value of other liver-related biomarkers, including liver fat measured by MR proton spectroscopy (^{1}H-MRS), liver iron measured by T_2*, transient elastography (TE) and histological fibrosis.

3.4 Methods

3.4.1 Patient selection

This study follows up patients from two cohort studies assessing the role of MRI in liver disease assessment (details in Table 1). Patients referred for a clinically indicated liver biopsy, or with a known diagnosis of liver cirrhosis were eligible to take part. Patients were recruited from hepatology services at Oxford and Reading between May 2011 and July 2017. Patients with contraindications to MRI scanning were excluded. The studies were approved by the

United Kingdom National Research Ethics Service (UK NRES references 13/SC/0243 and 11/H0504/2) and were conducted according to the declaration of Helsinki. All patients gave written and informed consent. Prognostic results from some of the patients from these cohort studies have been reported previously[187].

At baseline, patients underwent multiparametric MRI scans and TE following a minimum of 4 hours fasting.

Table 1: Description of patient cohort

Study 1 Rapid Imaging Assessment of the Liver (RIAL) (Ref: 11/H0504/2)	Study 2: Non-Invasive and COmprehensive Liver Imaging Assessment (NICOLA) (Ref: 13/SC/0243)
Patients with suspected liver disease of any aetiology attending liver clinic and scheduled for a clinically indicated liver biopsy having paired MR assessment (n = 129), 9 of whom did not undergo biopsy	Patients undergoing weight loss surgery who consented to liver biopsy during the weight loss procedure (n = 19; 1 did not undergo biopsy). Patients with suspected NAFLD/NASH scheduled to undergo clinically indicated biopsy (n = 16; 1 did not undergo biopsy). Patients with a diagnosis of cirrhosis. Diagnosis of cirrhosis based on histology or clinical or radiological findings (n = 27; 14 not biopsied). Patients scheduled for either clinically indicated transjugular biopsy and hepatic venous pressure gradient measurements, or clinically indicated percutaneous biopsy for suspected cirrhosis, who agreed to undergo transjugular biopsy instead to obtain HVPG measurements (n = 27; 2 did not undergo biopsy).

3.4.2 MRI protocol and analysis

Transverse abdominal shMOLLI images were acquired with parameters: 5(1)1(1)1 scheme, TI: 110 ms, 110 ms + RR, 110 ms + 2RR, 110 ms + 3RR, 110 ms + 4RR, 190 ms and 270 ms, TR/TE: 234.97/1.05 ms, TR of individual bSSFP readouts 2.5 ms, slice thickness 8.0 mm, GRAPPA acceleration factor 2, flip angle: 35°. FOV: patient dependent, matrix size: 144 x 192.

T_2* maps were acquired with a multi-echo GRE acquisition was used with parameters: TR: 500 ms, TEs: 1.25, 2.46, 3.69, 4.92, 6.15, 7.38, 8.61, 9.84 ms, 6 mm slice thickness, GRAPPA acceleration factor 2, flip angle: 20°. DIXON method used for fat fraction analysis. FOV: patient dependent, matrix size 104 x 128.

Magnetic resonance spectroscopy (MRS) was performed using a STimulated Echo Acquisition Method (STEAM)[204] sequence with the following parameters: voxel size 20 x 20 x 20 mm^3 (shim volume 40 x 40 x 40 mm^3), TE = 10 ms, mixing time = 7 ms, TR: 2 s for water-suppressed and TR: 4s for unsuppressed data with 1024 points acquired at bandwidth 50 Hz. 3 scans of 5 repetitions with water suppression were performed and 1 scan of 3 repetitions with water suppression turned off was performed.

All MRI analysis was carried out as described in Chapter 2.

In previous publications[187], cT_1 was tri-linearly mapped to a discrete scoring system graded from 0-4 named the LIF score in an attempt to align with histological assessment gradations. The LIF score is no longer supported by Liver*MultiScan*™, and here we report cT_1 values.

3.4.3 Histological assessment

All liver biopsies were performed as part of the patients' clinical care using 18G needles for percutaneous biopsies and 19G needles for transjugular biopsies. All biopsies were included

in the analysis as they were used to inform clinical care irrespective of the core length or number of portal tracts. All biopsies were assessed by a specialist liver histopathologist blinded to the imaging data, for the Ishak stage.

3.4.4 Transient elastography

TE measurements of liver stiffness were performed using Fibroscan® (Echosens, Paris) by trained operators. Both the M and XL probes were used. For a successful measurement, 10 valid readings were required and as per recommended guidelines, unreliable readings were defined as having an IQR/median > 0.3.

3.4.5 Clinical follow-up

Outcome data were extracted from the patients' electronic medical records. The clinical endpoints were all-cause mortality, liver-related mortality, and liver events (a composite endpoint comprising new ascites, variceal bleeding, HCC, hepatic encephalopathy, liver transplantation and liver-related mortality). Patient records were reviewed by two researchers blinded to the patients' MRI and liver-related biomarker results until the patients' latest medical evaluation or until they died, with any disagreements adjudicated by a senior clinician. Patients were considered as "lost to follow-up" if they did not return for any clinical follow-up after their baseline assessments. Where multiple liver events occurred, only the first event was counted in analysis. In patients with decompensation at baseline, new liver events had to be distinct from any previous events prior to baseline scan.

3.4.6 Statistical analysis

Baseline statistics are described as mean ± standard deviation (SD) for normally distributed variables and median (interquartile range (IQR)) for non-normally distributed variables. The primary outcome was liver event-free survival. Liver-related and all-cause mortality were

secondary outcomes. Survival analysis was carried out on all the available data for each outcome with each biomarker. Kaplan-Meier curves were compared using the log-rank test to evaluate significance of survival differences of binary and grouped cut-offs. All p values quoted for binary cut-offs were calculated using the log-rank test. Cox proportional hazards analysis was used to calculate hazard ratios (HRs) for grouped and continuous variables. Where impossible to obtain a hazard ratio in cut-off analysis (due to zero events occurring in a group), sensitivity, specificity, positive predictive value (PPV) and negative predictive values (NPV) were reported. Receiver Operator Curve (ROC) analysis was performed for all continuous variables.

The primary variable of interest was liver cT_1. Cut-off values were pre-defined with the cT_1 cut-off (825 ms) corresponding to the previously reported LIF score of 2[187]. Liver cT_1 was also grouped by minimum-median, median-3rd quartile, and 3rd quartile-maximum of cohort cT_1 values. The median, 95th percentile and additional cut-offs (770 and 800 ms) derived from internal Perspectum Diagnostics (PD) data were assessed (770 ms – cT_1 in the range 770-825 identified as borderline fibroinflammation and 800 ms – indicative of 'higher risk' in patients with suspected NAFLD). To determine the importance of the iron correction, we also assessed uncorrected liver T_1 at 825 ms and the best cut-off (Youden index) for prediction of events as there were no previously reported cut-offs at 3T.

Secondary variables were liver iron as measured by T_2*, liver fat as measured by ^{1}H-MRS, liver stiffness as measured by TE and Ishak fibrosis stage. The cut off for T_2* from our previous study[203] was used to define two severity categories for iron (high: T_2* ≤ 12.5 ms, low: T_2* > 12.5 ms) which represents the best threshold for identifying those with any grade of histological iron deposition and corresponds to a dry weight of iron of 1.7 mg/g[224]. The cut-

off for fat assessed was 5.6%[225]. A cut-off of 8 kPa for TE was used, which is reported to correspond to presence of marked fibrosis[226]. TE was also split into 4 groups (< 8kPa, 8-20 kPa, 20-39 kPa and ≥ 40kPa) for a more precise stratification of risk, similarly to previous studies[31,215]. Fibrosis severity was defined as mild (Ishak F0-2), moderate (Ishak F3-4) and severe (Ishak F5-6).

Given the spatial heterogeneity of disease often seen in autoimmune hepatitis (AIH) and biliary liver disease, subgroup analysis was conducted for liver event-free survival in patients with only the 3 main liver disease aetiologies (NAFLD, ArLD and VH) where disease distribution is more homogenous. Subgroup analysis was also conducted in patients with only compensated disease at baseline. Effect of treatment for underlying liver disease, presence of type 2 diabetes, body mass index (BMI) > 30 kg/m^2 and age were also analysed for prediction of outcomes.

Separate analysis was conducted only including data from patients that had *both* MRI *and* TE attempted. Unreliable results were included, and failed attempts were assigned either as false positive or false negative depending on the patient's outcome, generating 3 x 2 tables[227]. This analysis aimed to address the influence of technical rate of failure/unreliability.

Statistical significance was set at $p < 0.05$. Analysis was performed, and plots generated using R Statistical software[221].

3.5 Results

3.5.1 Cohort characteristics

235 patients were included, of which 17 were lost to follow up (Figure 11). In total, 218 patients were followed up for a total of 754 patient-years with a median (IQR) follow-up period of 42 (23-57) months.

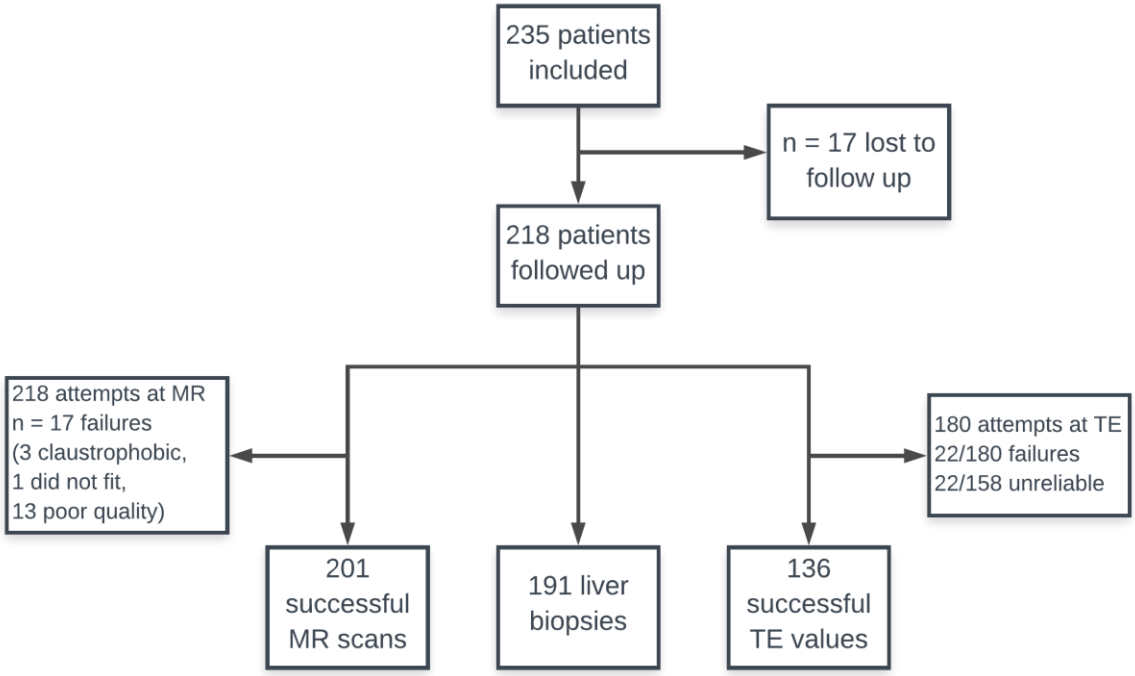


Figure 11: Flow diagram of study protocol. Poor quality data comprised images that were degraded significantly due to excess deposition of liver iron or images that did not pass quality control checks due to mistriggering of the shMOLLI sequence or the presence of image artefacts.

Baseline characteristics for the whole cohort and for those who had/did not have clinical events are shown in Table 2. The three most common aetiologies of chronic liver disease were NAFLD (n = 89, 41%), ArLD (n = 30, 14%) and VH (n = 55, 25%). Overall, 191 (88%) patients had liver biopsy, in whom fibrosis was mild, moderate and severe in 96 (50%), 30 (16%) and 65 (34%) respectively. Fifty-nine patients had their underlying liver disease treated during the

follow up period – these included patients with chronic hepatitis C achieving a sustained virological response (SVR) (n = 29), patients with chronic hepatitis B achieving viral suppression (n = 8), patients regularly consuming harmful amounts of alcohol achieving complete abstinence (n = 3), and patients undergoing bariatric surgery who lost >10% body weight (n = 19).

There were 23 new liver events. Liver-related mortality occurred in 13 patients and 7 patients died from non liver-related causes (all-cause mortality n = 20, Table 3). Of 21 patients with decompensation at baseline, 12 went on to develop a further liver-related event, including death. Concordance of liver-related event reporting between adjudicators was 93%.

When stratified into those who did and did not have clinical events, patients who had events were older and had higher prevalence of baseline decompensation, higher prevalence of ArLD, more severe histological fibrosis, significant differences in blood tests, and higher serum biomarkers, liver stiffness and liver cT_1 (Table 2).

Table 2: Baseline patient characteristics

	All (n=218)	Event-free (n=190)	With events (n=28)	*p* value
Age; years (median; IQR)	54 (45-60)	52 (44-59)	60 (56-68)	< 0.001
Male; n (%)	137 (63)	116 (61)	21 (75)	0.155
BMI; kg/m^2 (median; IQR)	28.1 (24.8-33.3)	27.9 (24.7-32.9)	28.3 (24.8-32.2)	0.897
T2DM; n (%)	46 (21)	38 (20)	8 (29)	0.245
Excess alcohol use; n (%)	39 (18)	28 (15)	11 (39)	0.0009
Decompensation at baseline; n (%)	21 (10)	9 (5)	12 (43)	< 0.0001
Treatment during follow-up period; n (%)	59 (27)	55 (29)	4 (14)	0.104
Primary disease aetiology; n (%)				
Viral hepatitis (VH)	55 (25)	51 (27)	4 (14)	0.155
NAFLD	89 (41)	81 (43)	8 (29)	0.159
ArLD	30 (14)	21 (11)	9 (32)	0.011
PSC/PBC/AIH	18 (8)	15 (8)	3 (11)	0.616
Other*	26 (12)	22 (11)	4 (14)	0.090
Histological fibrosis; Ishak stage; n (%)				
Mild; F0-2	96 (44)	93 (49)	3 (11)	< 0.0001
Moderate; F3-4	30 (14)	29 (15)	1 (4)	
Severe; F5-6	65 (30)	45 (24)	20 (71)	
No biopsy	27 (12)	23 (12)	4 (14)	
Blood results; (median; IQR)				

ALT (IU/L)	45 (29-87)	47 (30-91)	38 (25-67)	0.247
AST (IU/L)	41 (29-70)	38 (28-65)	58 (43-101)	< 0.001
Albumin (g/L)	44 (40-46)	44 (41-46)	38 (32-42)	< 0.0001
ALP (IU/L)	177 (130-250)	170 (125-232)	308 (181-389)	< 0.001
Bilirubin (μmol/L)	11 (8-17)	10 (7-15)	18 (13-35)	< 0.0001
GGT (IU/L)	78 (36-171)	67 (33-126)	287 (158-407)	< 0.0001
Platelets (x 10^{-9} L)	190 (144-260)	198 (155-267)	124 (88-157)	< 0.0001
Serum biomarkers; (median; IQR)				
FIB-4	1.68 (0.96-3.15)	1.52 (0.90-2.60)	4.98 (4.16-10.33)	< 0.0001
AST/ALT	0.89 (0.61-1.27)	0.87 (0.60-1.16)	1.67 (1.21-2.00)	< 0.0001
APRI	0.57 (0.32-1.00)	0.50 (0.30-0.89)	1.06 (0.76-2.15)	< 0.0001
ELF	9.6 (8.7-10.8)	9.3 (8.6-10.5)	11.9 (10.9-12.8)	< 0.0001
TE (kPa) (median, IQR)	9.0 (6.0-18.7)	8.1 (5.7-14.3)	32.0 (13.4-42.2)	< 0.0001
Liver cT_1 (ms) (median, IQR)	840 (788-930)	831 (785-924)	923 (845-1114)	< 0.0001
Liver fat (%) (median, IQR)	4.45 (1.38-10.64)	4.90(1.40-10.80)	4.19 (1.22-8.65)	0.583
T_2* (ms) median (IQR)	19.3 (14.0-22.3)	19.2 (13.9-22.4)	20.6 (14.6-21.2)	0.748

P values quoted for the differences between patients with events and no events using the Wilcoxon non-parametric test. *Other included cryptogenic cirrhosis, no specific features, sarcoid, cholestasis and haemochromatosis. Abbreviations: SD; standard deviation, IQR; interquartile range, BMI; body mass index, T2DM; type 2 diabetes, NAFLD; non-alcoholic fatty liver disease, ArLD; alcohol-related liver disease, PSC; primary sclerosing cholangitis, PBC; primary biliary cholangitis, AIH; autoimmune hepatitis; ALT; alanine aminotransferase, AST; aspartate aminotransferase, ALP; alkaline phosphatase, GGT; gamma-glutamyltransferase, FIB-4; fibrosis-4 score APRI; AST- platelet ratio index, cT_1; iron-corrected T_1.

Table 3: Table of clinical events

liver cT_1 (ms)	Diagnosis	Index event (months)	Patient mortality (months)	Decomp-ensated	ISHAK	Fibro-scan (kPa)
840	NAFLD	HE (15)	Yes (23)	No	6	44
978	ArLD	LRD (2)	Yes (2)	Yes	5	nm
1148	ArLD	ascites (7)	Yes (9)	Yes	nm	75
806	ArLD	None	Yes – NLRD (62)	Yes	3	nm
1125	NAFLD	Ascites (21)	Yes (64)	No	6	39.1
896	Viral hep	HCC (10)	Yes (18)	No	6	35.3
766	PSC	HE (47)	Yes (64)	No	2	29.9
994	NAFLD	None	Yes – NLRD (43)	No	6	12.8
1081	NAFLD	LRD (6)	Yes (6)	Yes	5	nm
945	ArLD	Ascites (14)	No	No	6	5.2
871	NAFLD	Ascites (22)	Yes (34)	No	6	15.1
900	Viral hep	HCC/ascites (17)	Yes – NLRD (39)	No	6	27.7
840	Viral hep	ascites (77)	No	No	5	nm
806	PSC	LT (6)	No	Yes	5	20.6
861	ArLD	HCC (7)	No	No	nm	69.1
1015	AIH	LT (6)	No	Yes	6	75
1284	ArLD	Ascites (3)	Yes (3)	Yes	6	34.0
1186	ArLD	Variceal bleed (19)	Yes (20)	No	6	na

1261	ArLD	SBP (41)	Yes (41)	Yes	nm	65.2
1163	Viral hep	Ascites (6)	No	No	6	na
827	NAFLD	Ascites (5)	Yes – NLRD (12)	Yes	2	9.0
869	NAFLD	HCC (1)	Yes (45)	Yes	5	36.9
1068	Cryptogenic	None	Yes – NLRD (32)	No	6	34.3
1139	ArLD	HE (4)	Yes (59)	Yes	6	53.2
765	Cryptogenic	variceal bleed (5)	No	Yes	5	11.7
na	HH	None	Yes – NLRD (56)	No	1	na
na	HH	HCC (85)	No	No	nm	6.1
900	NAFLD	None	Yes – NLRD (65)	No	6	nm

Liver cT_1 value classed as na if failed, Ishak score classed as nm if not clinically indicated, transient elastography (TE) classed as nm if not measured or na if failure/unreliable reading. **Abbreviations:** HE – hepatic encephalopathy, NLRD – non liver-related death, LT – liver transplant, ArLD, alcohol-related liver disease, NAFLD, non-alcoholic fatty liver disease, HCC – hepatocellular carcinoma, SBP, spontaneous bacterial peritonitis, PSC – primary sclerosing cholangitis, PBC, primary biliary cholangitis, AIH, autoimmune hepatitis, HH – haemochromatosis.

3.5.2 Survival analysis by biomarker

Table 4: Survival analysis results using binary cut-offs

Test	HR	95% CI	P value	Sens (%)	Spec (%)	PPV (%)	NPV (%)
cT_1 > 825 ms	4.869	1.441-16.46	0.0048	86.4	46.9	16.7	96.6
cT_1 > 1127	11.24	4.348-29.06	<0.0001	27.3	98.3	66.7	91.7
T_1 > 825 ms	1.799	0.733-4.413	0.19	68.2	48.3	13.9	92.6
T_1 > 1082 ms	9.429	3.903-22.78	<0.0001	36.4	96.1	53.3	92.5
Liver fat >5.6%	0.695	0.291-1.657	0.41	36.4	53.7	8.9	87.2
T_2* < 12.5 ms	0.981	0.332-2.900	0.97	18.2	80.6	10.3	89.0
TE > 8 kPa	6.116	1.387-26.97	0.0064	87.5	47.5	18.2	96.6
Ishak ≥ F5	17.65	4.075-76.46	<0.0001	89.5	72.4	27.4	98.3
Subset of patients with NAFLD, ArLD and VH							
cT_1 > 825 ms	-	-	0.0007	100.0	43.2	17.8	100.0
cT_1 > 1127	12.38	4.558-33.62	<0.0001	33.3	97.9	66.7	92.3
T_1 > 825 ms	2.477	0.815-7.528	0.098	77.8	44.5	14.7	94.2
T_1 > 1082 ms	10.80	4.157-28.04	<0.0001	44.4	95.2	53.3	93.3
Liver fat >5.6%	0.7851	0.3114-1.98	0.50	44.4	47.2	9.5	87.2
T_2* < 12.5 ms	1.062	0.349-3.227	0.920	22.2	78.1	11.1	89.1
TE > 8 kPa	8.417	1.077-65.8	0.015	90.9	48.5	16.7	97.9
Ishak ≥ F5	29.56	3.881-225.2	<0.0001	93.3	73.7	28.0	99.0

Abbreviations: HR: hazard ratio, sens: sensitivity, spec: specificity, PPV: positive predictive value, NPV: negative predictive value, NAFLD: non-alcoholic fatty liver disease, ArLD: alcohol-related liver disease, VH: viral hepatitis

Liver cT_1

MR scanning was successful in 201 of 218 (92%) patients, but failed in 17 (3 had claustrophobia, 1 did not fit in the scanner and 13 scans could not generate reliable cT_1).

On univariate Cox regression, liver cT_1 had HR 1.005 ($p < 0.0001$) for liver event-free survival (0.5% increase in risk for every 1 ms increase in cT_1; equivalent to HR 1.65; 65% increased risk per 100 ms) and predicted liver event-free survival with an area under the curve (AUC) of 0.73 (95% CI: 0.61-0.84).

Liver cT_1 > 825 ms could predict liver event-free survival ($p = 0.005$, Table 4), all-cause mortality ($p = 0.003$) and liver-related mortality ($p = 0.008$) (Figure 12A-C). Liver cT_1 > 825 ms correctly predicted 19/22 (86%) liver events, 17/19 (89%) deaths and 12/13 (92%) liver-related deaths. Liver cT_1 > 825 ms had a sensitivity of 86%, specificity of 47%, PPV 17%, and NPV 97% for liver event-free survival.

Liver cT_1 when stratified into 3 groups by quartile-derived cut-offs, showed significant increases in risk of liver events with increasing cT_1 ($p = 0.003$) (Figure 12D, Table 5). Using the 95^{th} percentile (1127 ms) cT_1 could predict liver event-free survival ($p < 0.001$) with sensitivity 27%, specificity 98%, PPV 67% and NPV 92%.

In patients with compensated liver disease only, liver cT_1 > 825 ms could predict liver event-free survival ($p = 0.02$, HR 7.9, 95% CI: 1-61, n = 11 events) and all-cause mortality ($p = 0.02$, n = 10 deaths), but not liver related mortality ($p = 0.1$, n = 6 deaths).

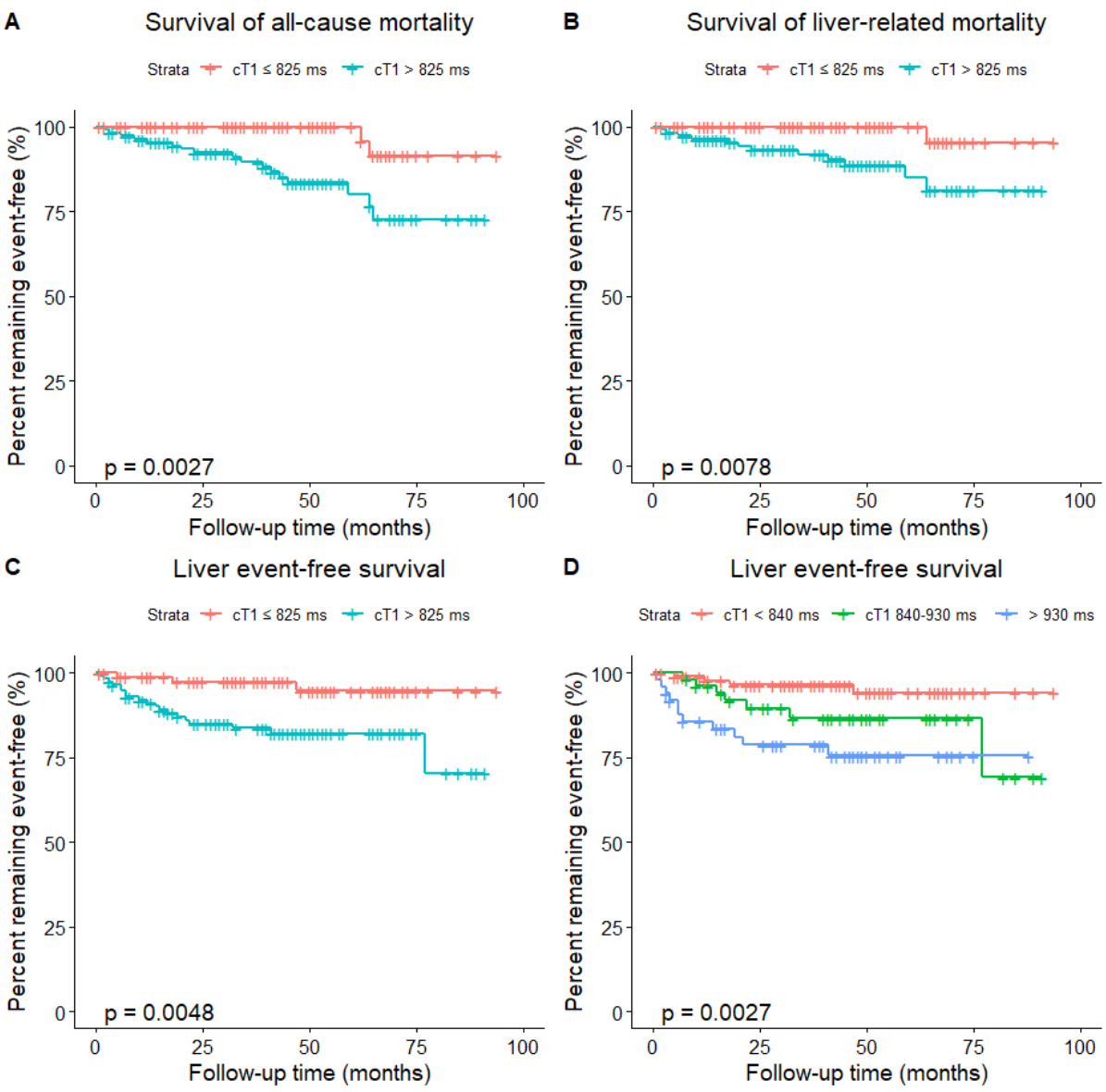


Figure 12: **Kaplan-Meier plots for clinical outcomes stratified by liver cT_1.** Plots show A) survival from all-cause mortality, B) survival from liver-related mortality, C) liver event-free survival stratified by liver cT_1 > 825 ms and D) liver event-free survival stratified by cT_1 quartiles (< 840 ms, 840-930 ms, > 930 ms) in n=201 patients with chronic liver disease. P values for differences between survival curves were generated by the log-rank test.

Table 5: Liver event-free survival assessed by grouped liver cT_1 cut-offs.

	Hazard ratio (HR)	95% CI	P value
cT_1 < 840 ms (n=94)	Reference	-	-
840 – 930 ms (n=50)	3.12	0.91 – 10.67	0.070
cT_1 > 930 ms (n=48)	6.08	1.93 – 19.13	0.002

Grouped cut-offs derived from: median (840 ms), 3rd quartile (930 ms)

NAFLD/ArLD/VH subgroup

In the NAFLD/ArLD/VH subgroup (n = 164) liver cT_1 had HR 1.007 (equivalent to HR 2.01; 101% increase per 100 ms, $p < 0.0001$) and AUC 0.77 (95% CI: 0.67-0.88) for liver event-free survival.

Liver cT_1 > 825 ms predicted liver event-free survival in the subset of NAFLD, ArLD and VH patients ($p = 0.0007$), correctly predicting all events (Figure 13A, Table 4).

In patients with compensated liver disease only, liver cT_1 > 825 ms could predict liver event-free survival ($p = 0.01$, n = 10 events), correctly predicting all events.

When stratified into 3 groups by quartile-derived cut-offs, liver cT_1 showed significant increases in risk of liver events with increasing thresholds ($p = 0.0018$) (Figure 13B, Table 6).

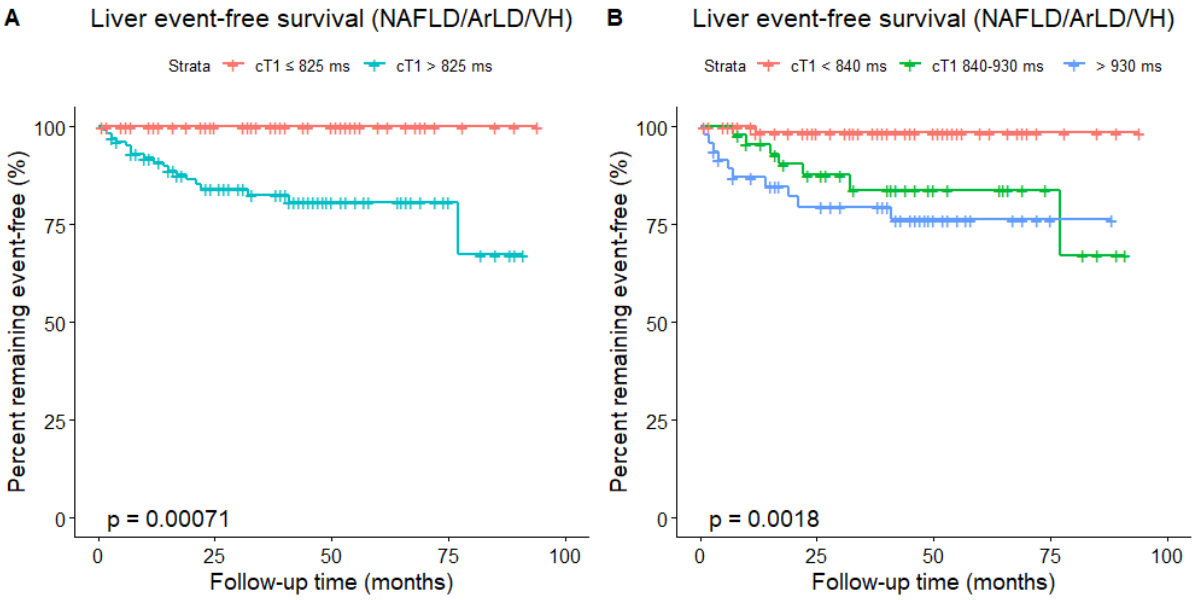


Figure 13*:* Kaplan-Meier plots to show liver event-free survival stratified by liver cT_1 in patients with NAFLD, ArLD and viral hepatitis (VH). Patients (n=164) were stratified by A) liver cT_1 > 825 ms and B) by cT_1 groups of < 840 ms, 840-930 ms and > 930 ms. P values for differences between survival curves were generated by the log-rank test.

Table 6: Liver event-free survival assessed by grouped liver cT_1 cut-offs in patients with NAFLD, ArLD and VH

	Hazard ratio (HR)	95% CI	P value
cT_1 < 840 ms (n=73)	Reference	-	-
840 – 930 ms (n=44)	11	1.3 – 88	0.026
cT_1 > 930 ms (n=47)	17	2.2 – 133	0.007

Grouped cut-offs derived from: median (840 ms), 3rd quartile (930 ms). **Abbreviations:** NAFLD – non-alcoholic fatty liver disease, ArLD - alcohol-related liver disease, VH - viral hepatitis

Liver T_1

Uncorrected liver T_1 had HR 1.004 (p = 0.001, 95% CI: 1.002-1.007) and AUC 0.66 (95% CI: 0.51-0.80) for liver event-free survival. Uncorrected liver T_1 > 825 ms was not predictive of liver event-free survival (p = 0.19, Table 4), liver-related mortality (p = 0.07) or all-cause mortality (p = 0.05) and classified 15/22 (77%) liver events correctly. There were 4 patients that developed liver events correctly identified by liver cT_1 > 825 ms who were not identified by uncorrected T_1 > 825 ms. The best T_1 cut-off for event-free survival was 1082 ms (p < 0.0001, Table 4) and predicted 8/22 (36%) of liver events.

T_1 > 825 ms was unable to predict liver event-free survival in the subgroup of NAFLD/ArLD/VH (p = 0.1, Table 4).

Liver fat and liver iron

Neither liver iron nor liver fat were predictive of clinical outcomes in either the whole cohort or the subset of NAFLD, ArLD and VH (See Table 4 for binary cut-off, data not shown for iron or fat as continuous variables).

Histology

Biopsies were performed in 191/218 (88%) patients. In those with biopsy there were 19 liver events, 11 liver-related deaths and 7 non-liver related deaths. As a semi-continuous variable Ishak stage could predict liver event-free survival with HR of 1.9 between stages (95% CI: 1.38-2.72, p = 0.0001) and an AUC 0.81 (95% CI: 0.73-0.89).

Ishak stage ≥ F5 predicted liver event-free survival (p < 0.0001, Figure 14A, Table 4), all-cause mortality (p < 0.0001), and liver-related mortality (p < 0.0001). Patients with F3-4 on biopsy showed no greater risk of liver-related events than with F0-2.

When grouped by mild, moderate and severe fibrosis, Ishak fibrosis stage showed significant increases in risk of liver events with increasing thresholds (p = 0.002, Figure 14B, Table 7).

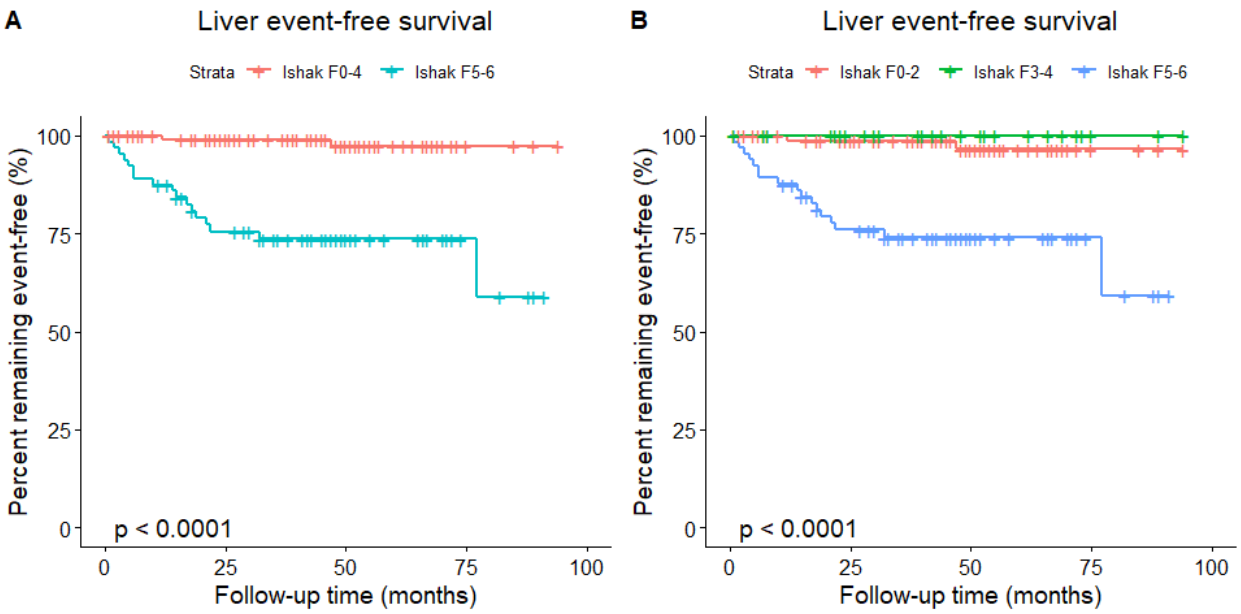


Figure 14: Kaplan-Meier curves for liver event-free survival stratified by histological fibrosis stage. Patients were stratified by A) Ishak stage ≥ F5 and B) mild, moderate and severe fibrosis.

Table 7: Liver event-free survival by Ishak fibrosis stage.

	Hazard ratio (HR)	95% CI	P value
Mild vs Moderate vs Severe	4.64 (per increasing group) *	2.003 – 10.75	< 0.0001

Patients were stratified by: mild (F0-2), moderate (F3-4), and severe (F5-6) fibrosis stage. *one group (F3-4) had zero events in it, so only global HR and p value could be calculated as direct comparisons cannot generate HRs.

NAFLD/ArLD/VH subgroup

In the NAFLD/ArLD/VH subgroup, Ishak stage could predict liver event-free survival with a HR of 2.2 (p = 0.0005, 95% CI: 1.41-3.41) and an AUC 0.84 (95% CI: 0.76-0.92). Ishak fibrosis stage ≥ F5 was predictive of liver event-free survival (p < 0.0001, Figure 15A, Table 4), correctly predicting 14/15 (93%) events. When grouped by mild, moderate and severe fibrosis, Ishak fibrosis stage showed significant increases in risk of liver events with increasing thresholds (p = 0.002, Figure 15B, Table 8).

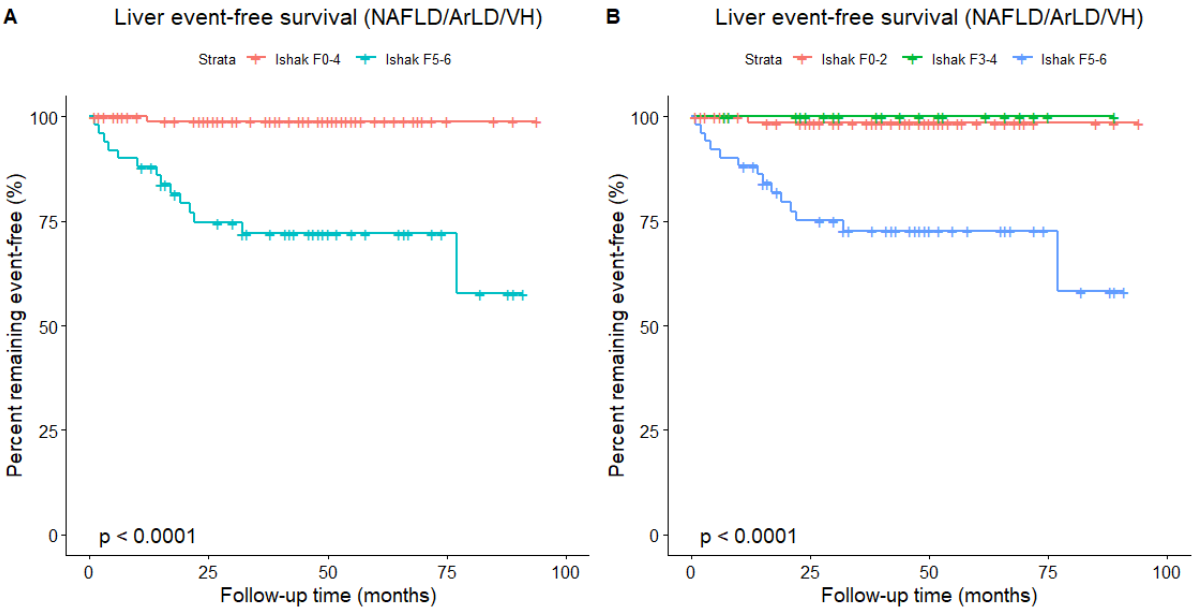


Figure 15: Kaplan-Meier curves for liver event-free survival stratified by histological fibrosis stage in patients with NAFLD/ArLD/VH. Patients were stratified by A) Ishak stage ≥ F5 and B) mild, moderate and severe fibrosis.

Table 8: Liver event-free survival by Ishak fibrosis stage in patients with NAFLD/ArLD/VH.

	Hazard ratio (HR)	95% CI	P value
Mild vs Moderate vs Severe	6.48 (per increasing group)*	1.978 – 21.23	0.002

Patients were stratified by: mild (F0-2), moderate (F3-4), and severe (F5-6) fibrosis stage. *one group (F3-4) had zero events in it, so only global HR and p value could be calculated as direct comparisons cannot generate HRs.

Transient elastography

TE was attempted in 180/218 patients with technical failure observed in 22 (12%). From the remaining 158 measurements, unreliable values were observed in 22 (13%) with 136 (76%) reliable TE measurements remaining. As a continuous variable TE could predict liver event-free survival with a HR of 1.049 (95% CI: 1.029-1.069) and AUC 0.80 (95% CI: 0.67-0.93).

TE > 8 kPa could predict liver event-free survival (p = 0.006, correctly predicting 14/16 (88%) liver events, Figure 16A, Table 4), all-cause mortality (p = 0.001) and liver-related mortality (p = 0.008).

When stratified into groups, TE showed significant increases in risk of liver events with increasing TE thresholds (p = 0.0002, Figure 16B, Table 9).

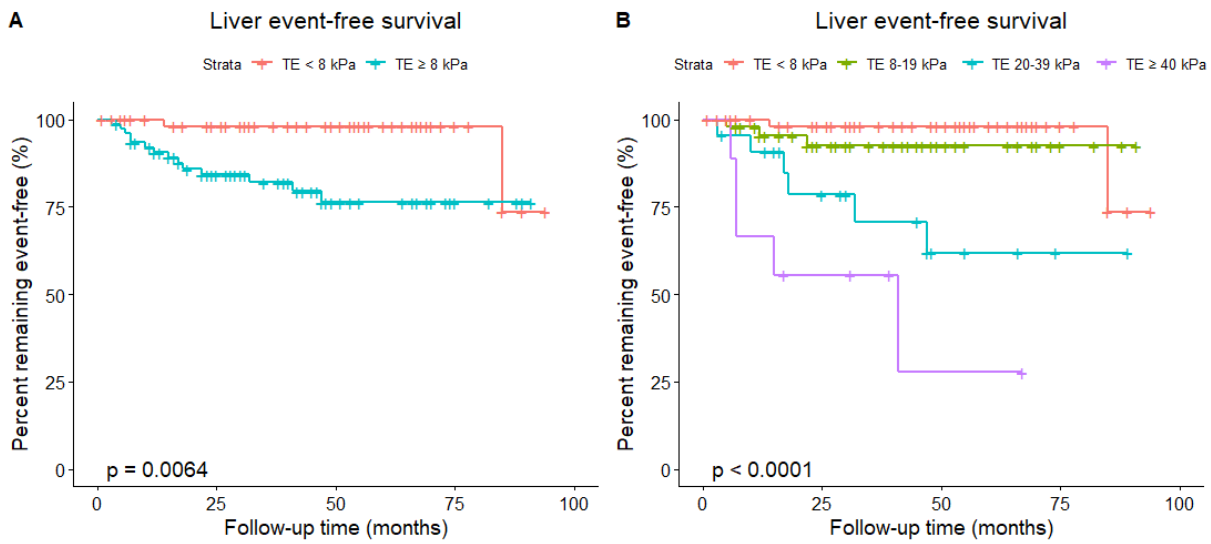


Figure 16: Kaplan-Meier curves for liver event-free survival stratified by TE score. Patients stratified by A) TE 8kPa and B) TE < 8kPa, 8-19 kPa, 20-39 kPa, ≥ 40 kPa.

Table 9: Liver event-free survival assessed by grouped TE cut-offs

	Hazard ratio (HR)	95% CI	P value
TE < 8 kPa (n=55)	Reference	-	-
8 – 20 kPa (n=46)	2.1	0.35 - 12.6	0.415
20 – 39 kPa (n=22)	9.4	1.89 - 46.5	0.006
TE ≥ 40 kPa (n=9)	26.7	5.04 - 141.2	< 0.001

NAFLD/ArLD/VH subgroup

In the NAFLD/ArLD/VH combined subgroup, TE was attempted in 142/174, failed in 17/142 (12%) and was unreliable in 17/125 (14%) (total failed and/or unreliable in 33/142 (24%)). In the 108 patients that it performed accurately, TE could predict liver event-free survival with a HR of 1.056 (95% CI: 1.031-1.083) and AUC 0.83 (95% CI: 0.67-0.99). TE > 8 kPa was able to predict liver event-free survival (p = 0.015,

Figure **17**A, Table 4), correctly predicting 10/11 (91%) events. When stratified into groups, TE showed significant increase in risk of liver events with increasing TE thresholds (p < 0.001,

Figure **17**B, Table 10).

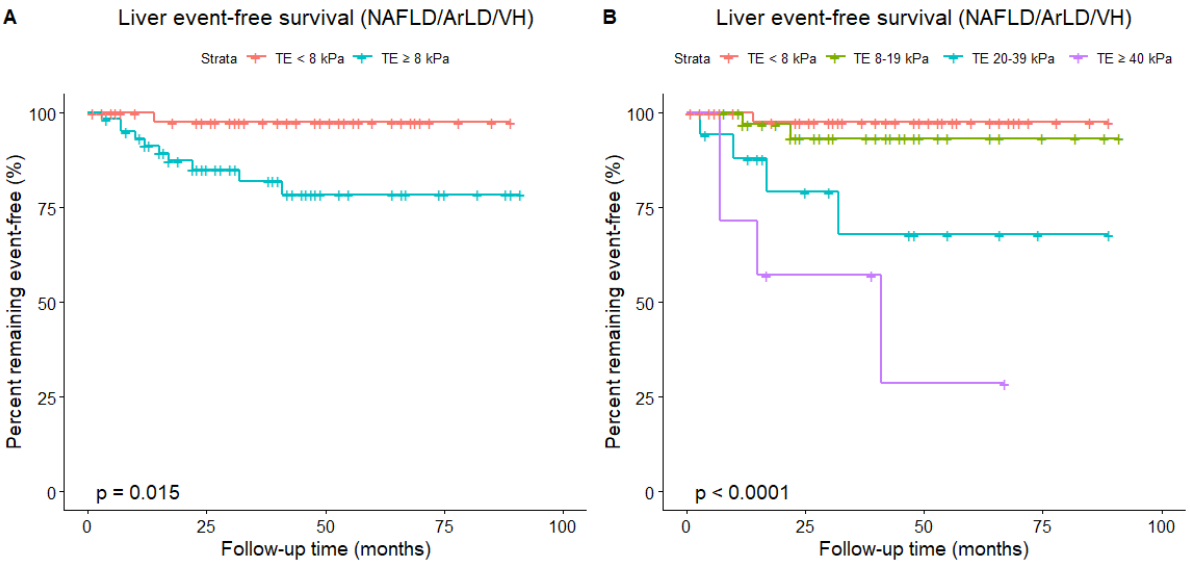


Figure 17: Kaplan-Meier curves for liver event-free survival stratified by TE score in patients with NAFLD/ArLD/VH. Patients stratified by a) TE 8kPa and b) TE < 8kPa, 8-19 kPa, 20-39 kPa, ≥ 40 kPa.

Table 10: Liver event-free survival assessed by grouped TE cut-offs in patients with NAFLD/ArLD/VH.

	Hazard ratio (HR)	95% CI	P value
TE < 8 kPa (n=48)	Reference	-	-
8 – 20 kPa (n=36)	2.7	0.24-29.0	0.426
20 – 39 kPa (n=17)	12.9	1.42-114	0.022
TE ≥ 40 kPa (n=7)	33.6	3.74-302	0.002

Analysis including failed and unreliable results for TE and MR (3 x 2 table analysis)

In 180 patients with both TE and MRI attempted (Figure 18), there were 21 liver events, 12 liver-related deaths and 6 non liver-related deaths. Liver cT_1 > 825 ms could predict liver event-free survival (p = 0.03), but TE > 8 kPa could not (p = 0.14, Table 11).

Table 11: Additional analysis assessing influence of failure/unreliability rates on MRI and TE results

Test	HR	95% CI	P value	Sens (%)	Spec (%)	PPV (%)	NPV (%)
cT_1 > 825 ms	3.1	1.0-9.1	0.03	81.0	43.4	15.9	94.5
TE > 8 kPa	2.10	0.76-5.76	0.14	76.2	39.6	14.3	92.6

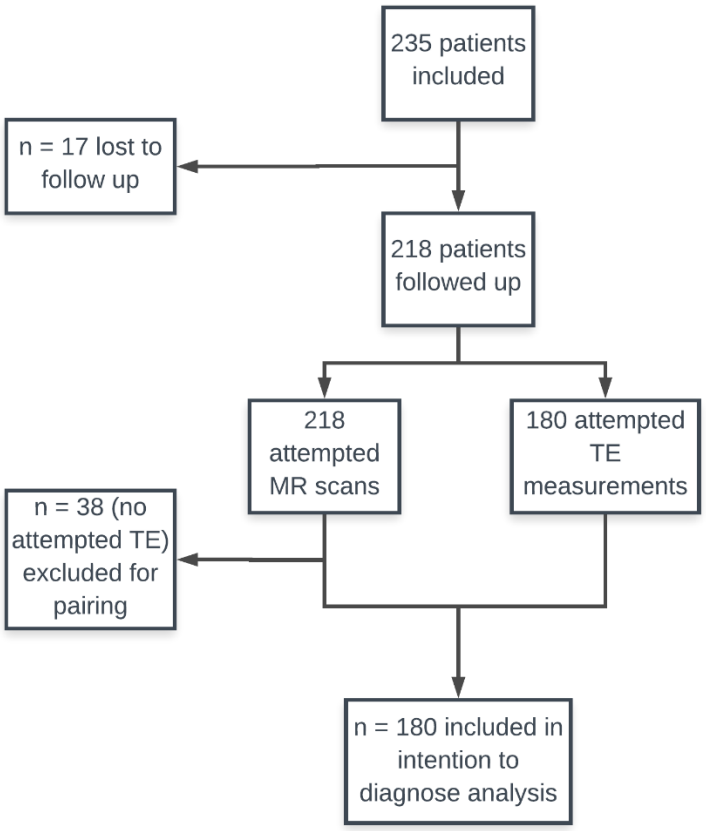


Figure 18: Flow diagram for direct comparison between liver cT_1 and transient elastography (TE).

Traditional risk factors

Neither presence of Type 2 diabetes nor BMI > 30 kg/m^2 could predict liver event-free survival, liver-related or all-cause mortality (data not shown). Age as a continuous variable was predictive of liver event-free survival (p = 0.003, HR: 1.061, 95% CI: 1.02-1.1), liver-related mortality and all-cause mortality (data not shown). Treatment for underlying liver disease was predictive of survival of liver mortality (p = 0.02), but not liver event-free survival (p = 0.07) or all-cause mortality (p = 0.06).

Example images

See Figure 19 for example images of patients who had and did not have clinical outcomes.

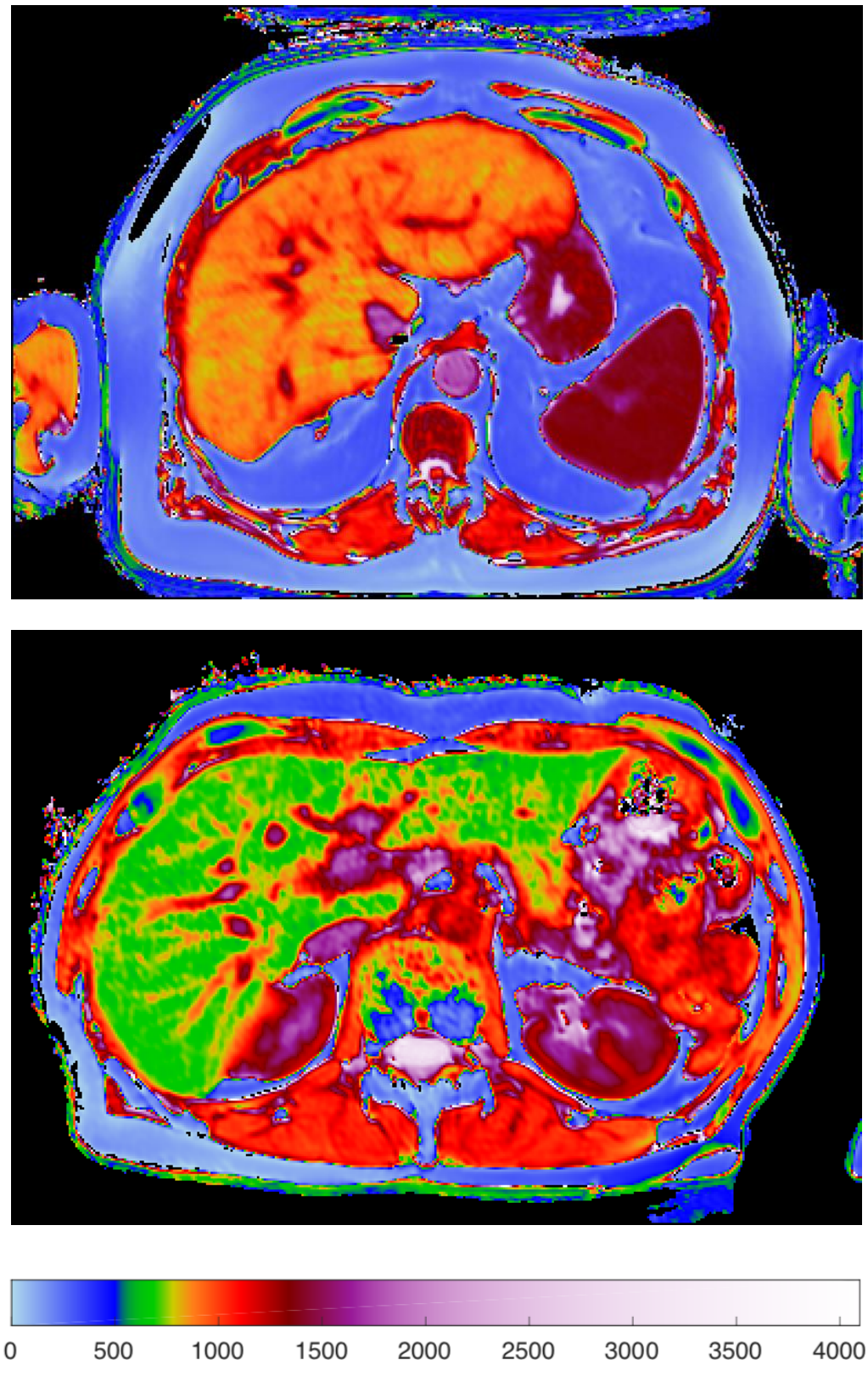


Figure 19: Example of abdominal iron-corrected T_1 (cT_1) colour maps. (Top) a patient with compensated cirrhosis (Ishak F6) at time of scan who went on to develop hepatocellular carcinoma, ascites and death; (bottom) a patient with no histological or clinical evidence of cirrhosis at time of scan who did not develop any clinical event. Scale in milliseconds (ms).

3.6 Discussion:

We have shown further evidence with 754 patient-years of follow-up that liver cT_1 can predict all-cause mortality, liver related mortality and liver event-free survival. Liver cT_1 > 825 ms correctly predicted 19/22 (86%) of liver events and performed even better when the patient cohort was restricted to NAFLD, ArLD and VH, predicting all liver events.

Higher liver T_1 has been shown in a previous study to be associated with development of liver-related clinical outcomes, albeit in a cohort of patients with CLD devoid of high liver iron content[184]. In our study there were four patients correctly classified by liver cT_1 > 825 ms who went on to develop liver-related events but were wrongly classified as to be at low risk of a liver event by their uncorrected T_1. This reduced the classification percentage from 86% to 68%. The best T_1 cut-off was high (1082 ms) and therefore, although it was highly specific, it performed poorly as a screening tool, missing 14/22 events. These results underscore the utility of the iron-correction method, especially considering the prevalence of elevated liver iron content in the general population[86] and particularly in patients with ArLD and NAFLD. T_1 was unable to predict clinical outcomes at 825 ms in our study but could do so as a continuous variable, albeit with lower AUC and HR than cT_1.

Neither ^{1}H-MRS quantitation of liver fat, nor liver iron content as measured by MRI T_2* could predict clinical outcomes. Although excess liver fat can accelerate fibrosis progression in certain patients[228], it has been shown that histological liver fibrosis is the only parameter independently associated with poorer outcomes[222]. The distribution of iron deposition (hepatocellular or parenchymal) may be important, but this cannot be assessed by MRI, which may explain T_2*'s lack of prognostic ability.

Here, liver cT_1 had almost identical sensitivity, specificity, PPV and NPV to TE. However, in the additional analysis that assessed the influence of failed/unreliable results on prognostic accuracy, TE ≥ 8 kPa was unable to predict liver event-free survival, whereas liver cT_1 > 825 ms could. This was not a true intention-to-diagnose analysis as there were no data collected on patients who were excluded at the outset due to MRI contraindications, therefore the success rates of MRI and TE in the general CLD population may be different. A distinct advantage of cT_1 therefore is its technical success rate, reported as 98% vs 85% for TE – including patients with obesity and ascites[229] – and excellent test-retest repeatability and reproducibility metrics[188] compared to a high coefficient of variation for TE[230]. In clinical practice, if a TE score is unobtainable, patients are often biopsied. Liver cT_1 would be a useful diagnostic tool at this point instead of TE – or as 2nd line test after TE in failed cases – to avoid unnecessary biopsies, perhaps most usefully in NAFLD/ArLD/VH, where we have shown it has the most prognostic value and as liver cT_1 has been shown to be cost effective in the non-invasive risk stratification[231] and diagnostic pathway of NAFLD[232].

Our reported TE failure rates are higher than in other prognostic studies conducted in cohorts with no or smaller (< 13.5%) proportions of NAFLD patients[31,214,215] and comparable to studies in NAFLD[233]. Our cohort contained a significant number (89/218) of NAFLD patients, (including patients scheduled for bariatric surgery), in which it is reported that TE performs less well and can reach failure/unreliability rates of 20%[41] – an important consideration with NAFLD now the most prevalent CLD. Liver cT_1 failed mainly in those who either were unable to enter the scanner (claustrophobia) or with extremely high iron (haemochromatosis), degrading the T_2* signal. The TE failure rate from our study was much higher than in other studies. This could represent an anomaly but could also be due to the fact that a majority of these examinations took place between 2014-2016; with evidence showing that TE failure/unreliability rates have

decreased considerably since then with advances in probe technology. It is likely therefore that these rates are not representative of current failure/unreliability rates. Additionally, the experience of the researchers performing the tests may have played a part. In current clinical practice usually TE is performed by trained nurses who perform a large number of TE examinations a week. In this study TE was performed by suitably trained operators as part of the research team, albeit with less day-to-day experience than a dedicated Fibroscan nurse. This factor is likely to have affected the unreliability results more than the failure rates as it may take more attempts to obtain reliable readings, but absolute failure may be dependent on the technical aspects of the probe and/or the skin to capsule ratio.

In comparison to Ishak fibrosis stage F5-6, which was able to predict 17/19 (89%) of liver events, cT_1 predicted 19/22 (86%) events. This suggests that liver cT_1 performs similarly in excluding patients at low risk of developing clinical outcomes as biopsy. Liver $cT_1 > 825$ was shown to have a similarly high NPV to biopsy for the exclusion of patients with liver events (96% vs 98%), indicating its suitability as a screening test. Liver $cT_1 > 1127$ ms had the highest specificity (98%) and PPV (67%) of all binary cut-offs.

When we compare our results to those of particular blood serum-based biomarkers, we find it compares favourably to most, largely due to the variability of their prognostic capability. FIB-4 and ELF have performed well in tertiary care settings[30,33,234] with FIB-4 often performing better in patients with hepatitis C[30,235] than NAFLD[233]/ArLD[35]. Indeed, a systematic review and meta-analysis of ELF studies (personal communication from the LITMUS (Liver Investigation: Testing Marker Utility in Steatohepatitis) group) has indicated that current cut-offs for staging of liver fibrosis in NAFLD/NASH may be too high for accurate diagnosis of fibrosis in a lower prevalence setting, indicating lower cut-offs may also be required for prediction of clinical

outcomes. Liver cT_1 is likely to have a greater impact earlier rather than later in the screening process for the detection of early liver fibroinflammatory disease and help in monitoring these patients noninvasively. Studies of even larger cohorts may bring out these differences in the future.

Additionally, when we only included patients with compensated disease at baseline, we observed a higher HR of 7.8 for liver event-free survival compared to 4.9 when including those with evidence of decompensation when stratifying patients by liver cT_1 > 825 ms. This is important, as the utility of a non-invasive tool to predict those at risk would be of most use in compensated patients who do not outwardly exhibit signs of CLD. There were only 10 events in compensated patients however, so larger numbers are needed to verify this.

An unexpected observation was that liver cT_1 > 825 ms cut-off provided a greater separation when assessing all-cause mortality than liver mortality. There were by definition more all-cause mortality events which increased the statistical significance. The greater separation is most likely explained by the sheer chance that a few more non liver-related deaths were correctly classified as at-risk. In larger cohorts with more mortality cases, we would expect this separation to be higher for liver-related mortality. Another reason may be that often even if there is a non-liver related health problem, a side effect may be liver and/or other internal organ inflammation. Additionally, 2 patients died from non-liver related cancer – some cancer therapies can be hepatotoxic and may have contributed to the higher liver cT_1 in those.

We have shown liver cT_1 performs especially well in the most prevalent CLDs, namely NAFLD, ArLD and VH, with cT_1 > 825 ms correctly identifying all liver events. These particular diseases represent a significant financial cost to the UK's social and healthcare systems[15]. Accurate, cost-effective stratification of these patients with non-invasive biomarkers would allow

significant savings in time and money and an increase in patient comfort. There is also an unmet need for reliable non-invasive endpoints in NASH drug development. Liver cT_1 can differentiate simple steatosis from NASH[191], and the prognostic value of liver cT_1 shown in this study provides further evidence to support the use of liver cT_1 as a surrogate endpoint in NASH drug trials. Indeed, liver cT_1 is shown later in this thesis and in a recent study by Harrison et al.[190] to be able to monitor treatment response.

Limitations of this study included no fixed follow-up timepoint. This means that the AUC analyses, sensitivities, specificities, PPVs and NPVs generated in cut-off analysis should be taken with caution because unlike in Cox proportional hazards and Kaplan-Meier analysis, the effect of patients being censored at different timepoints is not taken into account. Our patient cohort included a wide range of liver disease severity from patients with mild fibrosis to those with cirrhosis and decompensation at baseline. A prognostic test might not necessarily be indicated in patients with decompensation, so we performed subgroup analysis excluding these patients to address this. Future studies should examine the prognostic value of liver cT_1 in prespecified liver disease severity groups and particularly those with compensated cirrhosis.

In conclusion, liver iron-corrected T_1 (cT_1) can identify patients at risk of all-cause mortality, liver-related mortality and liver event-free survival in a cohort with mixed CLD aetiologies and is particularly accurate in predicting liver event-free survival in patients with NAFLD, ArLD and VH. Further multicentre studies should be carried out to validate our results.

Chapter 4: Changes in liver cT_1 after direct acting antiviral treatment for chronic hepatitis C virus

The work in this Chapter has been submitted as a paper to:

Jayaswal A, Levick C, Collier J, Tunnicliffe E, Kelly M, Neubauer S, Barnes E, Pavlides M. Liver cT_1 decreases following treatment by direct acting antiviral therapy in patients with chronic hepatitis C virus.

and has been presented as an abstract at:

Jayaswal A, Levick C, Pavlides M, Collier J, Kelly M, Neubauer S, Barnes E, (2017). Multiparametric magnetic resonance imaging can quantitatively assess fibroinflammatory response to direct acting antiviral treatment of hepatitis C virus. Hepatology 66, 336A-337A.

4.1 Abstract:

Background & Aims: Direct acting antiviral therapies (DAAs) for treatment of chronic hepatitis C virus (HCV) have excellent rates of viral eradication (>95%), but their effect on regression of liver fibrosis is unclear. The primary aims were to use magnetic resonance imaging (MRI) and spectroscopy (MRS) to evaluate changes in liver cT_1, liver fat and liver iron content (LIC) in patients with chronic HCV following treatment with DAAs.

Patients & Methods: In this prospective study, 15 patients with chronic HCV approved to start treatment with DAAs were consecutively recruited from hepatology outpatient clinics at the John Radcliffe Hospital, Oxford, UK, between December 2014 and September 2017. All patients with a most recent clinical transient elastography (TE) measurement of > 8 kPa were invited to take part in the study. Patients underwent MRI and MRS at baseline (before treatment), and at 24 weeks and 48 weeks after the end of treatment (EoT).

Results: Liver cT_1 showed significant decreases from baseline to 24 weeks EoT (876 vs 806 ms, p = 0.002, n = 15), from baseline to 48 weeks EoT (876 vs 788 ms, p = 0.0002, n = 13) and from 24 weeks EoT to 48 weeks EoT (806 vs 788ms, p = 0.016, n = 13). Between baseline and 48 weeks EoT significant reductions in liver fat and increase in LIC were observed.

Conclusions: Liver cT_1 can quantitatively assess changes in fibroinflammation in patients with chronic HCV undergoing treatment by DAAs. Due to the relatively short follow-up period, the reduction in cT_1 is likely due to a reduction in inflammation rather than regression of fibrosis.

4.2 Introduction

Globally, chronic hepatitis C virus (HCV) affects approximately 71 million people, a prevalence of 1% of the world's population[4]. Chronic HCV is associated with liver cirrhosis, development of hepatocellular carcinoma (HCC), liver failure and death. Recently, direct acting antiviral therapies (DAAs) have transformed the ability to treat and cure HCV, particularly in those with advanced fibrosis and cirrhosis since these patients were largely intolerant of previous interferon (IFN) based regimens. These treatments have high rates of success at eradicating the virus[2-14] (> 95%), even in patients with cirrhosis and in patients who have relapsed following earlier attempts at cure.

However, current evidence suggests that after achieving a sustained virological response (SVR) through DAA treatment, risk of developing HCC is reduced, but not removed completely[249–251].

There is no clear-cut set of risk factors for development of HCC post viral eradication. A study by Iannou et al.[252] identified severity of Fibrosis-4 (FIB-4) score > 3.25 as a risk factor for development of HCC post treatment for HCV in patients with cirrhosis and fibrosis. Another study developed models for HCC risk with age, platelet count, serum aspartate aminotransferase/alanine aminotransferase (AST/ALT) ratio and albumin being the main contributors to risk, based on 1,412 cases of HCC diagnosed during a mean follow-up of 2.5 years[253].

There is a growing body of evidence to support regression of non-invasive biomarkers of liver fibrosis in significant proportions of patients following DAA treatment[185,254–260]. However, it is

unclear from these results whether these changes are due to regression of inflammation or fibrosis, or both. Furthermore, it is also unclear whether changes in fibroinflammation reduce risk of HCC development. One study has assessed paired liver biopsy in 15 patients pre and post DAA therapy[261]. This study showed maintenance of histological fibrosis and inflammation in some patients and regression in other patients. Furthermore, in this same study changes in non-invasive markers overestimated histological regression of fibrosis.

Another factor affecting risk of HCC development post viral eradication is epigenetic change. Epigenetic aberrations are linked to tumour development. Chronic HCV has been shown to induce epigenetic changes, which persist following viral eradication by DAAs[262,263]. It is postulated therefore that some patients, regardless of whether a reversal in either fibrosis or inflammation takes place, may still be at risk of HCC development due to persistent epigenetic alterations.

As a result, patients with liver cirrhosis prior to treatment are recommended for HCC surveillance which includes ultrasound combined with alpha-fetoprotein (AFP) performed every 6 months indefinitely. AFP is suboptimal for follow-up on its own due to low sensitivity[264,265], and abdominal ultrasound can miss small HCCs. If our magnetic resonance imaging (MRI) techniques can monitor disease regression/progression in these patients and assess their future risk of HCC development, the need for close ultrasound follow-up could be removed. MRI may also be able to detect lesions that might be otherwise missed.

4.3 Aims

The main aim of this chapter was to evaluate changes in liver MR parameters (liver cT_1 and liver iron content (LIC) using MRI and liver fat using MR spectroscopy (MRS)) in patients with chronic HCV following treatment with DAAs.

Secondary aims were to assess changes in other non-invasive biomarkers of liver disease including transient elastography (TE), simple blood tests and serum-based fibrosis scores. Patients were also followed up by their patient records for the duration of the study for development of HCC and other liver decompensation events.

4.4 Methods

4.4.1 Patient selection

In this prospective study, patients with chronic HCV (as defined by positive HCV RNA assays on more than two occasions 6 months apart) were consecutively recruited from hepatology outpatient clinics at the John Radcliffe Hospital, Oxford, UK, between December 2014 and September 2017. All patients who were approved to start treatment with DAAs by a multi-disciplinary team, with the most recent clinical TE measurement of > 8 kPa, were invited to take part in the study. (Patient demographics and treatment regimens are detailed in Table 12). Patients underwent study assessments comprising MRI, MR spectroscopy (MRS), TE and blood sampling at baseline (before treatment), and at 24 weeks and 48 weeks after the end of treatment (EoT).

The study was approved by the local Research Ethics Committee (ref no: 13/SC/0234) and conforms to the declaration of Helsinki. All patients gave written, informed consent to take part.

4.4.2 Patient follow-up

Patients were also followed up for the development/recurrence of HCC and development/recurrence of liver decompensation events (ascites, variceal bleeding, hepatic encephalopathy, liver transplantation and liver-related mortality) for the duration of the study by their electronic patient records.

4.4.3 MR protocol

Transverse abdominal shMOLLI images were acquired with parameters: 5(1)1(1)1 scheme, TI: 110 ms, 110 ms + RR, 110 ms + 2RR, 110 ms + 3RR, 110 ms + 4RR, 190 ms and 270 ms, TR/TE: 234.97/1.05 ms, TR of individual bSSFP readouts 2.5 ms, slice thickness 8.0 mm, GRAPPA acceleration factor 2, flip angle: 35°. FOV: patient dependent, matrix size: 144 x 192.

T_2* maps were acquired with a multi-echo GRE acquisition was used with parameters: TR: 500 ms, TEs: 1.25, 2.46, 3.69, 4.92, 6.15, 7.38, 8.61, 9.84 ms, 6 mm slice thickness, GRAPPA acceleration factor 2, flip angle: 20°. DIXON method used for fat fraction analysis. FOV: patient dependent, matrix size 104 x 128.

Magnetic resonance spectroscopy (MRS) was performed using a STimulated Echo Acquisition Method (STEAM)[204] sequence with the following parameters: voxel size 20 x 20 x 20 mm^3 (shim volume 40 x 40 x 40 mm^3), TE = 10 ms, mixing time = 7 ms, TR: 2 s for water-suppressed and TR: 4s for unsuppressed data with 1024 points acquired at bandwidth 50 Hz. 3 scans of 5 repetitions with water suppression were performed and 1 scan of 3 repetitions with water suppression turned off was performed.

4.4.4 MR analysis

Liver*MultiScan*™ was used to calculate cT_1, T_2* and liver iron content (LIC). MRS data were analysed to quantify liver fat as described in Chapter 2.

4.4.5 Blood sampling

Blood samples were drawn on the same day as the MRI scan. Samples were analysed for ALT, AST, alkaline phosphatase (ALP), gamma-glutamyl transferase (GGT), C-reactive protein (CRP), bilirubin, albumin, cholesterol, high density lipoprotein (HDL) and low-density lipoprotein (LDL) and platelet count. Simple scores for serum-based fibrosis markers FIB-4[266], aspartate aminotransferase to platelet ratio index (APRI)[267] and AST/ALT ratio were calculated.

Diabetes status, alcohol consumption, evidence of decompensation (presence of ascites, hepatic encephalopathy and varices), height, weight, and body mass index (BMI) measurements were recorded.

4.4.6 Transient Elastography

TE (Fibroscan®, Echosens, Paris, France) was performed by a trained professional, and a valid reading was defined according to the manufacturer's criteria (10 valid shots and IQR/median ratio > 0.3). TE was performed on the same day as the MRI scan and the blood tests where possible, alternatively the most recent clinically performed value was used (max. 1 weeks difference in time).

4.4.7 Statistical analysis

The primary outcome was the change in liver cT_1 following DAA treatment. Secondary outcomes included changes in liver stiffness measured by TE, liver fat, LIC, FIB-4, APRI and

AST/ALT scores and other blood-based biomarkers of liver fibrosis and fat. Comparison between study timepoints was performed using the non-parametric paired Wilcoxon test. Since the paired Wilcoxon test was used to evaluate significance of changes in variables from one study timepoint to the next, only participants who had scans at both timepoints were included in the analysis (i.e. 15 patients at timepoint 1, but only 13 at timepoint 3; only 13 participants' data would be included in the Wilcoxon test for significance).

The cohort was split into patients with and without cirrhosis (defined by Ishak fibrosis stage ≥ F5 or clinical decompensation at baseline assessment) and temporal changes in liver cT_1 were compared between groups. The cohort was split by genotype (GT) and liver fat compared between GT1, GT2 and GT3 patients.

Correlations between liver cT_1 and TE and between changes in liver cT_1 and liver fat were investigated using the Spearman's rank correlation coefficient. Normality of variables was assessed by the Shapiro-Wilk test. Significance level was set at 0.05. We did not correct for multiple comparisons using the Bonferroni correction as there were small numbers of patients and many of the tests were dependent tests. Median and interquartile ranges (IQR) were reported for overall patient characteristics due to the small numbers involved.

4.5 Results

Seventeen patients were recruited. One patient was excluded due to inability to hold their breath during the MRI scan, another patient did not return for any follow-up visits. All patients achieved SVR as confirmed by undetectable HCV RNA assay. Fifteen returned for follow-up visits at 24 weeks EoT and 13 (2 lost to follow-up) returned at 48 weeks EoT (Figure 20).

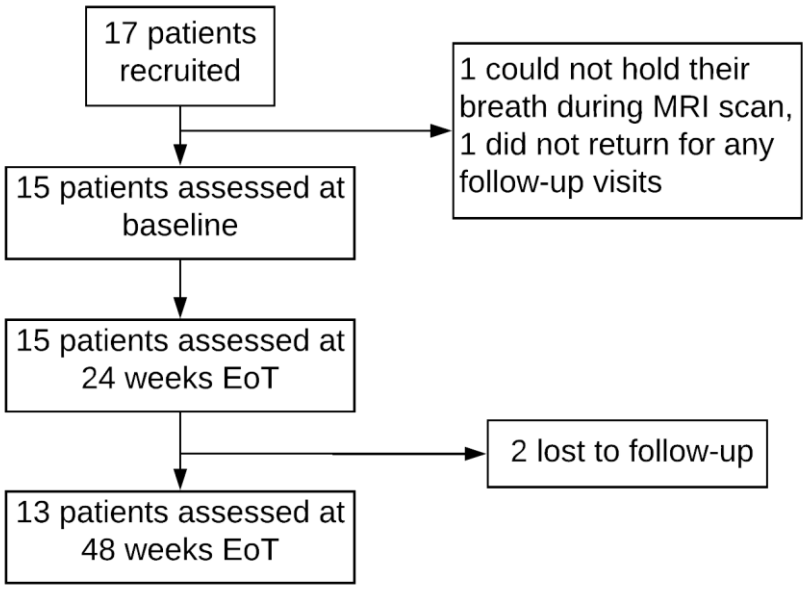


Figure 20: Study flow diagram. Abbreviations: EoT – end of treatment.

4.5.1 Patient characteristics

Patients had median (IQR) age 58 (52-60), were 75% (10/15) male, with median (IQR) BMI 26 (24-28) kg/m^2. 7 (47%) patients had genotype 1 (GT1) infection, 3 (20%) had GT2 and 5 (33%) had GT3. 4/15 (27%) of patients self-reported as drinking alcohol in excess of the recommended limits and 7/15 (47%) had cirrhosis at baseline. Individual patients' details are reported in Table 12 and changes in blood tests outlined in Table 13, changes in non-invasive scores outlined in Table 14, and changes in MR parameters outlined in Table 15.

Table 12: Individual baseline patient characteristics and self-reported alcohol consumption at each study visit

DAA treatment used	HCV genotype	Treatment duration (weeks)	Age	Sex	BMI	Alcohol consumption			Baseline cirrhosis	Baseline decompensation
						Baseline	24 weeks	48 weeks		
Sofosbuvir, Ledipasvir + Ribavarin	3	24	60	M	27.1	0	0	0	Yes	No
Ombitasvir, Paritaprevir, Ritonavir, Dasabuvir + Ribavarin	1	12	57	M	18.9	110	140	140	Yes	No
Sofosbuvir, Ledipasvir + Ribavarin	1	12	54	F	26.8	14	5	5	Yes	No
Sofosbuvir, Ledipasvir + Ribavarin	1	12	60	F	23.2	0	0	0	Yes	Yes (HCC)
Sofosbuvir + Ribavarin	2	12	67	M	24.8	0	0	0	No	No
Daclatasvir, Sofosbuvir + Ribavarin	3	12	58	M	27.0	0	0	0	Yes	Yes (ascites)
Ombitasvir, Paritaprevir, Ritonavir, Dasabuvir + Ribavarin	1	12	48	M	22.6	0	0	14	No	No
Elbasivir, Grazoprevir	1	12	68	F	24.9	0	0	0	No	No
Glecaprevir, Pibrentasvir	3	8	58	M	23.9	14	14	14	No	No
Glecaprevir, Pibrentasvir	2	8	65	F	32.1	0	0	0	Yes	No
Glecaprevir, Pibrentasvir	3	12	45	M	30.3	0	0	0	No	No
Glecaprevir, Pibrentasvir	2	12	59	M	28.4	130	65	130	No	No
Elbasivir, Grazoprevir	1	12	28	F	22.9	14	14	14	No	No
Elbasivir, Grazoprevir + Ribavarin	1	16	53	M	25.7	140	50	30	Yes	Yes (ascites)
Glecaprevir, Pibrentasvir	3	12	52	M	41.5	32	30	28	No	No

Table 13: Simple blood test results at baseline, 24 weeks EoT and 48 weeks EoT.

	Baseline (*n=15*) median (IQR)	24 weeks EoT (*n=15*) median (IQR)	48 weeks EoT (*n=13*) median (IQR)	p value (baseline vs 48 weeks)
Bilirubin (mmol/L)	12 (8-16)	12 (10-18)	9 (9-13)	0.0658
ALT (IU/L)	120 (59-134)	28 (21-34)	23 (17-29)	**< 0.001**
ALP (IU/L)	96 (70-152)	71 (60-105)	77 (64-93)	**0.0137**
Albumin (g/L)	36 (34-37)	39 (37-40)	38 (36-39)	**0.0172**
GGT (IU/L)	107 (88-135)	36 (24-71)	33 (24-73)	**< 0.001**
AST (IU/L)	94 (67-140)	32 (29-42)	29 (25-35)	**0.0078**
Trigs (mmol/L)	0.99 (0.70-1.64)	0.78 (0.62-1.00)	0.70 (0.61-0.97)	0.0840
HDL (mmol/L)	1.1 (0.9-1.4)	1.4 (1.2-1.6)	1.5 (1.1-1.6)	**0.0453**
LDL (mmol/L)	2.5 (2.2-2.8)	2.5 (2.3-2.9)	2.5 (2.3-2.7)	0.8885
Ratio	4.22 (3.07-5.00)	2.90 (2.60-4.03)	2.94 (2.71-3.63)	**0.0098**
Platelets ($x10^9$/L)	192 (126-239)	223 (127-233)	233 (113-286)	0.0661

Abbreviations: ALT (Alanine aminotransferase), ALP (alkaline phosphatase), AST (aspartate aminotransferase), GGT (gamma-glutamyl transferase), Trigs (Triglycerides), HDL (high density lipoprotein), LDL (low density lipoproteins), FIB-4 (fibrosis-4), APRI (AST to platelet ratio index).

Since the paired Wilcoxon test was used to evaluate significance of changes in variables from one study timepoint to the next, only participants who had scans at both timepoints were included in the analysis (i.e. 15 patients at timepoint 1, but only 13 at timepoint 3; only 13 participants' data would be included in the Wilcoxon test for significance).

Table 14: Non-invasive liver tests at baseline, 24 weeks EoT and 48 weeks EoT.

Test (normal range)	Baseline (*n=15*) median (IQR)	24 weeks EoT (*n=15*) median (IQR)	48 weeks EoT (*n=13*) median (IQR)	p value (baseline vs 48 weeks)
Child-Pugh score	6 (5-6)	5 (5-5)	5 (5-5)	0.3458
FIB-4	2.51 (1.96-4.48)	2.27 (1.57-3.73)	1.39 (1.16-4.06)	**0.0156**
AST/ALT (0-1)	1.158 (0.703-1.196)	1.400 (1.160-1.467)	1.29 (1.052-1.771)	0.0781
APRI	1.185 (0.776-2.299)	0.406 (0.259-0.541)	0.219 (0.181-0.658)	**0.0156**
TE (kPa) (0-7)	19.5 (8.2-24.3)	11.9 (7.4-19.2)	10.1 (7.4-18.7)	**0.002**

Abbreviations: FIB-4 (fibrosis – 4), AST/ALT score (aspartate aminotransferase/alanine aminotransferase ratio), APRI (aspartate aminotransferase to platelet ratio index), TE (transient elastography). N.B. on normal ranges: FIB-4 is not validated in patients with normal blood scores, therefore no 'normal' range can be described; the same is true for APRI. Child-Pugh score is developed to categorise stages of cirrhosis rather than distinguish between normal and fibrotic patients. TE range is based on evidence that 90-95% of healthy people will have a TE score <7 kPa.

Table 15*:* Magnetic resonance parameters at baseline, 24 weeks EoT and 48 weeks EoT

	Baseline (*n=15*) median (IQR)	24 weeks EoT (*n=15*) median (IQR)	48 weeks EoT (*n=13*) median (IQR)	p value (baseline vs 48 weeks)
Liver cT_1 (ms)	876 (832-901)	806 (772-836)	788 (767-807)	**0.0002**
Liver fat (%)	5.17 (2.64-6.03)	2.54 (1.63-8.48)	2.65 (1.71-5.03)	**0.027**
Liver iron (mg/g)	0.913 (0.884-0.994)	1.019 (0.922-1.161)	0.950 (0.895-1.129)	**0.021**

4.5.2 Liver cT_1

Liver cT_1 showed a significant decrease from baseline to 24 weeks EoT (876 vs 806 ms, p = 0.002, n = 15, Figure 21), from baseline to 48 weeks EoT (876 vs 788 ms, p = 0.0002, n = 13) and from 24 weeks EoT to 48 weeks EoT (806 vs 788ms, p = 0.016, n = 13,

Table **15**, Figure 21A). Example images at baseline and 24 weeks EoT are shown in Figure 22.

4.5.3 Liver fat

Liver fat showed a trend towards decrease from baseline to 24 weeks EoT (5.17% vs 2.54%, p = 0.064, n = 12) reaching a significant decrease between baseline and 48 weeks EoT (5.17% vs 2.65%, p = 0.027, n = 11) and a significant decrease between 24 and 48 weeks EoT (median 2.54% vs 2.65%, mean 5.72% vs 3.62%, p = 0.039,

Table **15**, Figure 21C). There was no correlation between change in liver fat and change in cT_1. Liver fat did not vary between genotypes with mean values for genotypes 1 (n = 7), 2 (n = 3) and 3 (n = 4) at baseline being 6.59%, 7.05% and 6.26% respectively. GT1 vs GT2 (p = 0.833), GT2 vs GT3 (p = 0.629), GT1 vs GT3 (p = 0.649). Liver fat also did not vary between genotypes at 48 weeks EoT, with mean values for genotypes 1 (n = 7), 2 (n = 2) and 3 (n = 2) being 2.81%, 5.78%, and 4.31% respectively. GT1 vs GT2 (p = 0.333), GT2 vs GT3 (p = 1.00), GT1 vs GT3 (p = 0.5).

4.5.4 Liver Iron

LIC increased between baseline and 24 weeks EoT (0.913 vs 1.019 mg/g dry weight, p = 0.022, n = 15) and between baseline and 48 weeks EoT (0.900 vs 0.950 mg/g, p = 0.021, n = 13), but there was no significant change between 24 and 48 weeks EoT (p = 0.470, Table 15, Figure 21D).

4.5.5 Transient Elastography

TE showed significant decreases between baseline and 24 weeks EoT (19.1 vs 12.5 kPa, p = 0.002, n = 12), between baseline and 48 weeks EoT (21.4 vs 12.3 kPa, p = 0.002, n = 10) and between 24 and 48 weeks EoT (14.5 vs 12.9 kPa, p = 0.042, n = 12, Table 14, Figure 21B).

TE measurements were significantly correlated with liver cT_1 (r = 0.60, p < 0.0001, n = 39, Figure 23). In measurements taken before treatment, there was a moderate but significant correlation (r = 0.547, p = 0.05, n = 12) and in post-treatment measurements there was a significant, but weaker correlation between (r = 0.134, p = 0.003, n = 27). Reductions in cT_1 correlated moderately (r = 0.500, p = 0.015, n = 23) with reductions in TE.

4.5.6 Blood tests and biochemistry

From both baseline to 24 weeks EoT and baseline to 48 weeks EoT, significant reductions were observed in ALT ($p < 0.001$ and $p < 0.001$), AST ($p = 0.004$ and $p = 0.008$), albumin ($p = 0.005$ and $p = 0.017$), ALP ($p = 0.014$ and $p = 0.014$), GGT ($p < 0.001$ and $p < 0.001$), FIB-4 ($p = 0.023$ and $p = 0.016$), APRI ($p = 0.008$ and $p = 0.016$), and a significant increase in HDL were observed ($p = 0.028$ and $p = 0.045$).

From baseline to 24 weeks EoT significant reductions were observed in AST/ALT ($p = 0.027$) and ferritin ($p = 0.023$). From baseline to 48 weeks EOT significant reductions were observed in cholesterol:HDL ratio ($p = 0.010$).

Between 24 weeks and 48 weeks EoT, AST, ALT, GGT and APRI showed a trend towards decrease ($p = 0.059$, $n = 11$), ($p = 0.077$, $n = 12$), ($p = 0.076$, $n = 13$) and ($p = 0.064$, $n = 10$) respectively but no significant reductions were observed. A significant increase in platelet count was observed ($p = 0.041$, $n = 10$).

4.5.7 Patient follow-up

No *de novo* occurrences of HCC were observed following DAA treatment. One patient had been diagnosed with HCC and had undergone ablation therapy in November 2011, 4 years before DAA was started. The HCC recurred prior to DAA therapy (July 2015) and was again ablated (June 2016) after completion of DAA therapy (November 2015). This HCC did not recur during the follow-up of the study. This patient's liver cT_1 was 882 ms at baseline, reduced to 864 ms at 24 weeks EoT and then 829 ms by 48 weeks EoT.

4.5.8 Patients with cirrhosis vs patients without cirrhosis

Seven patients had clinical or histologically diagnosed cirrhosis. There was no significant difference in mean baseline cT_1 between patients with and without cirrhosis (863 vs 866 ms, p = 0.694). Liver cT_1 was significantly higher in the group with cirrhosis at 24 weeks EoT (840 ms vs 789 ms, p = 0.014, Figure 24), with only a significant drop in cT_1 in the non-cirrhotic group (859 vs 767 ms, p = 0.008) not in the group with cirrhosis (863 vs 827 ms, p = 0.375). There was no significant difference in mean cT_1 at 48 weeks between patients with and without cirrhosis (p = 0.234) nor any difference in magnitude of reduction in cT_1 between 24 and 48 weeks EoT. Both groups were overall significantly reduced at 48 weeks EoT from baseline (p = 0.031 for both).

4.5.9 Effects of excess alcohol consumption

Patients who had self-reported as drinking alcohol in excess of the recommended limits throughout the study (n = 4) had significantly higher levels of LIC (1.55 mg/g vs 0.96 mg/g dry weight, p < 0.001), significantly higher levels of liver fat (9.19% vs 3.91%, p = 0.004) and significantly higher GGT levels (133 vs 59 IU/L, p = 0.011).

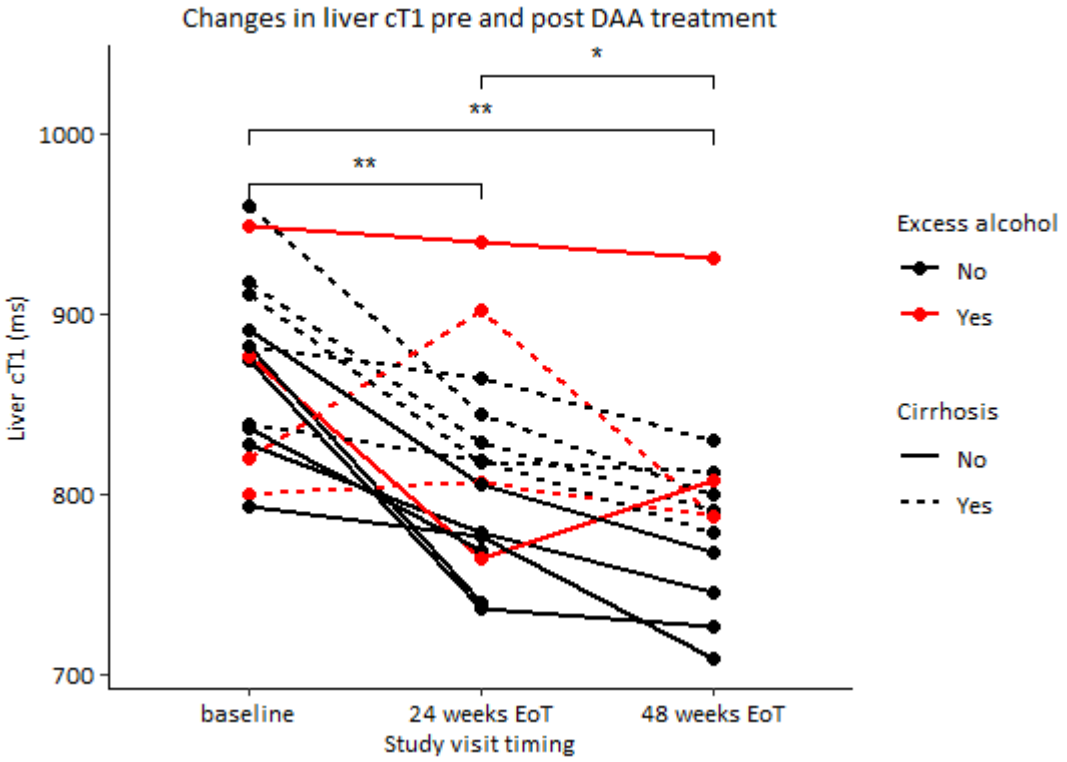
When these patients were excluded from the analysis, liver cT_1 showed a significant decrease from baseline to 24 weeks EoT (874 vs 798 ms, p < 0.001, n = 11), from baseline to 48 weeks EoT (877 vs 773 ms, p = 0.004, n = 9) and from 24 to 48 weeks EoT (808 vs 773, p = 0.004, n = 9).

Liver fat showed a significant decrease from baseline to 24 weeks EoT (6.38 vs 4.47%, p = 0.004, n = 9), from baseline to 48 weeks EoT (5.17 vs 2.53%, p = 0.031, n = 6) but not from 24 to 48 weeks EoT (p = 0.313).

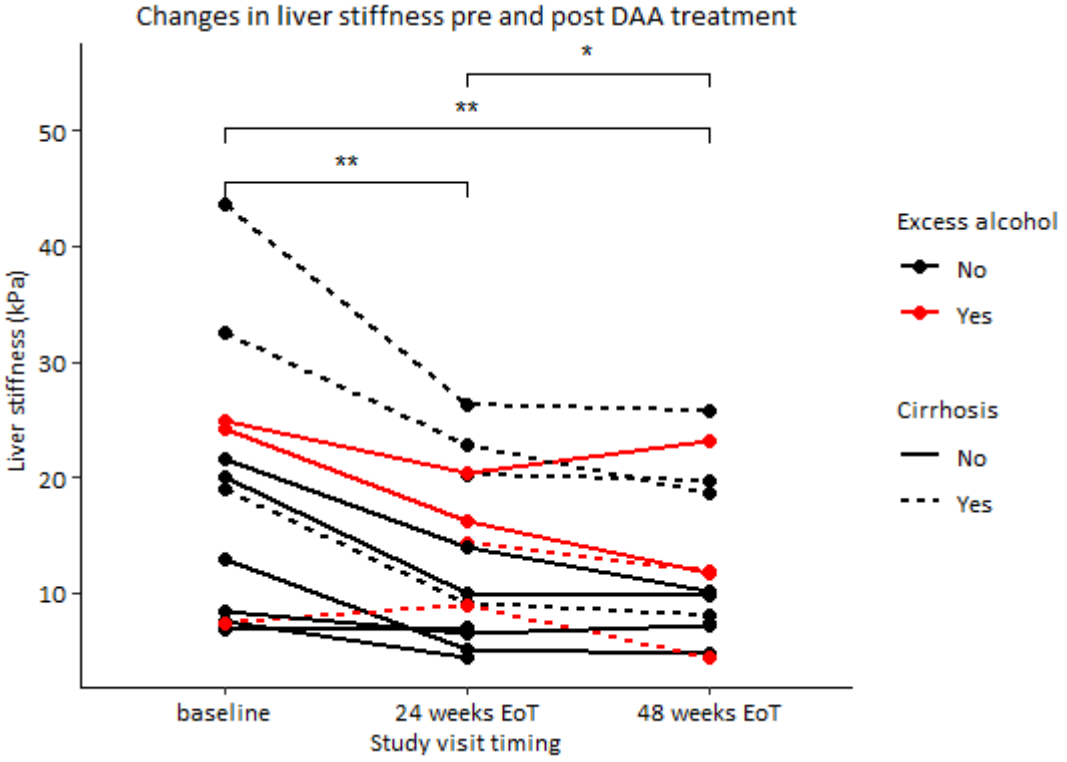
TE showed a significant decrease from baseline to 24 weeks EoT (19.1 vs 11.7 kPa, p = 0.008, n = 9), baseline to 48 weeks EoT (22.5 vs 12.0 kPa, p = 0.016, n = 7) and showed a trend towards decrease from 24 to 48 weeks EoT (p = 0.08, n = 8).

LIC did not change between any timepoints.

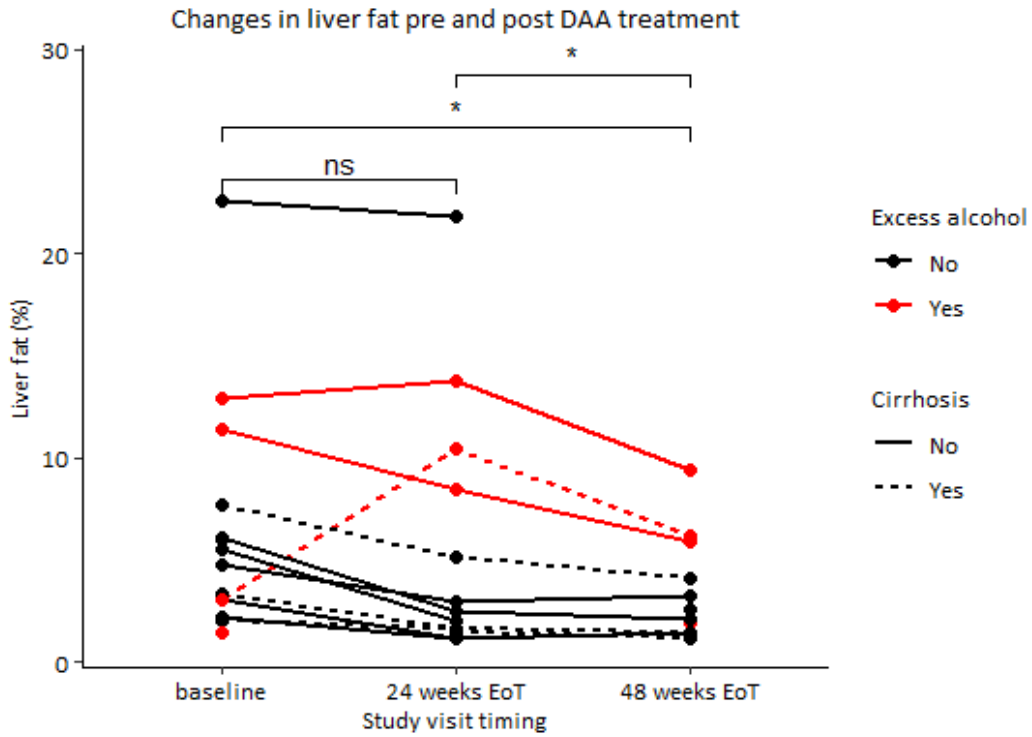
A


Changes in liver cT1 pre and post DAA treatment
Liver cT1 (ms)
1000
900
800
700
*
**
**
baseline
24 weeks EoT
48 weeks EoT
Study visit timing
Excess alcohol
No
Yes
Cirrhosis
No
Yes

B


Changes in liver stiffness pre and post DAA treatment
Liver stiffness (kPa)
50
40
30
20
10
*
**
**
baseline
24 weeks EoT
48 weeks EoT
Study visit timing
Excess alcohol
No
Yes
Cirrhosis
No
Yes

C



D

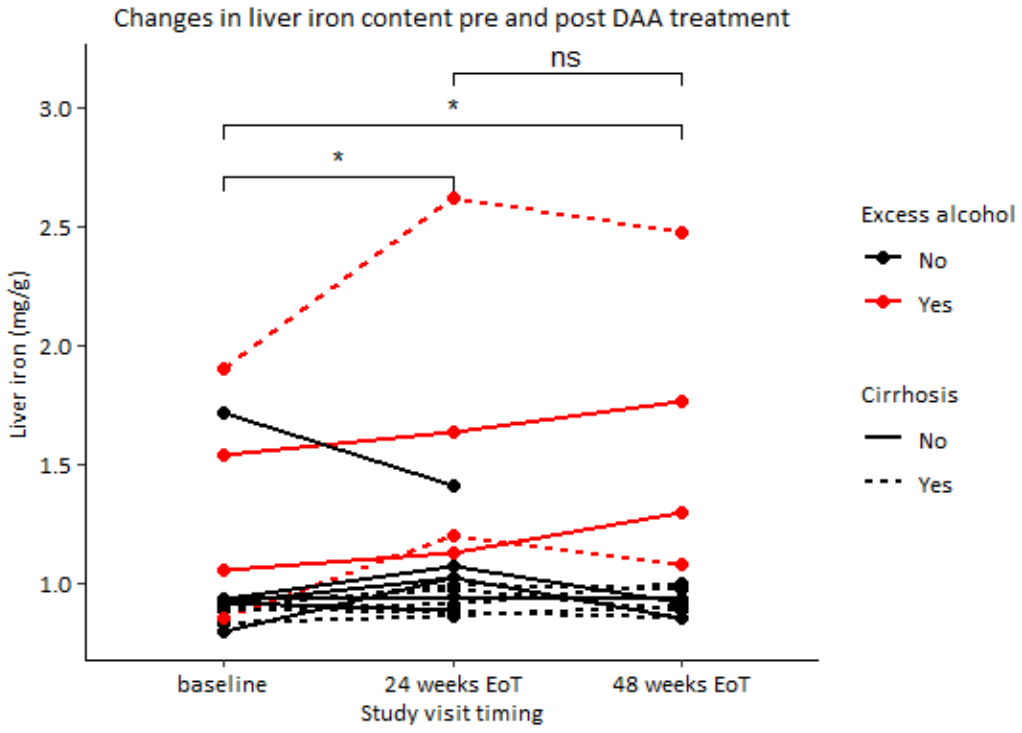


Figure 21: Response to DAA treatment from baseline to 24 weeks EoT, from baseline to 48 weeks EoT and from 24 weeks EoT to 48 weeks EoT. Plots show response to treatment by A) liver cT_1, B) transient elastography (TE), C) liver fat, D) liver iron content. * denotes $p < 0.05$, ** denotes $p < 0.01$, ns denotes p is not significant. EoT – end of treatment, DAA – direct acting antiviral therapy.

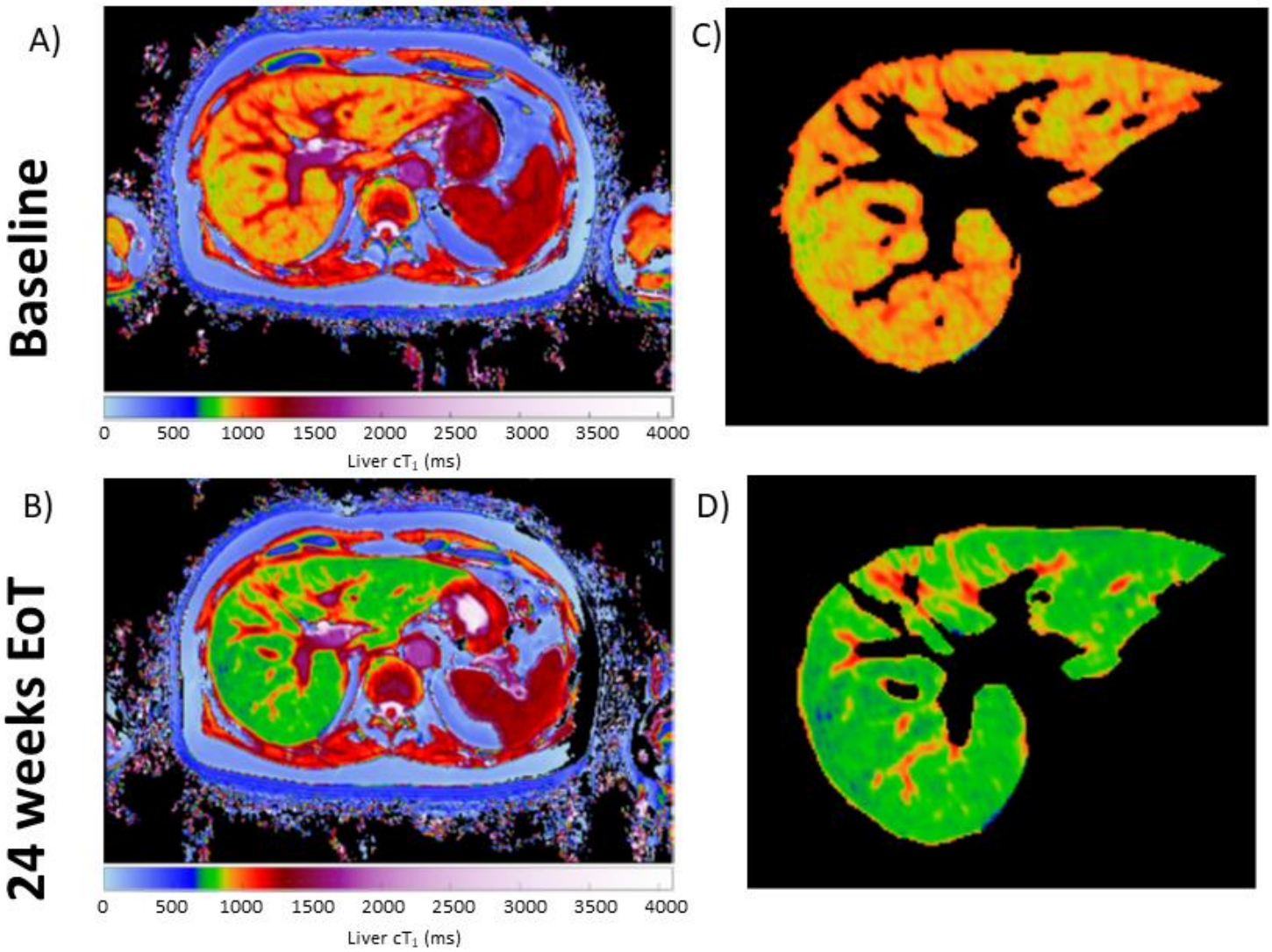


Figure 22: Abdominal cT_1 images and segmentation masks pre and post treatment. A-B) Abdominal cT_1 images from a patient with chronic hepatitis C virus at baseline and at 24 weeks after end of treatment (EoT) respectively. C-D) Semi-automatically segmented liver cT_1 masks from the same patient at baseline and 24 weeks EoT. See Chapter 2 for details of liver segmentation method.

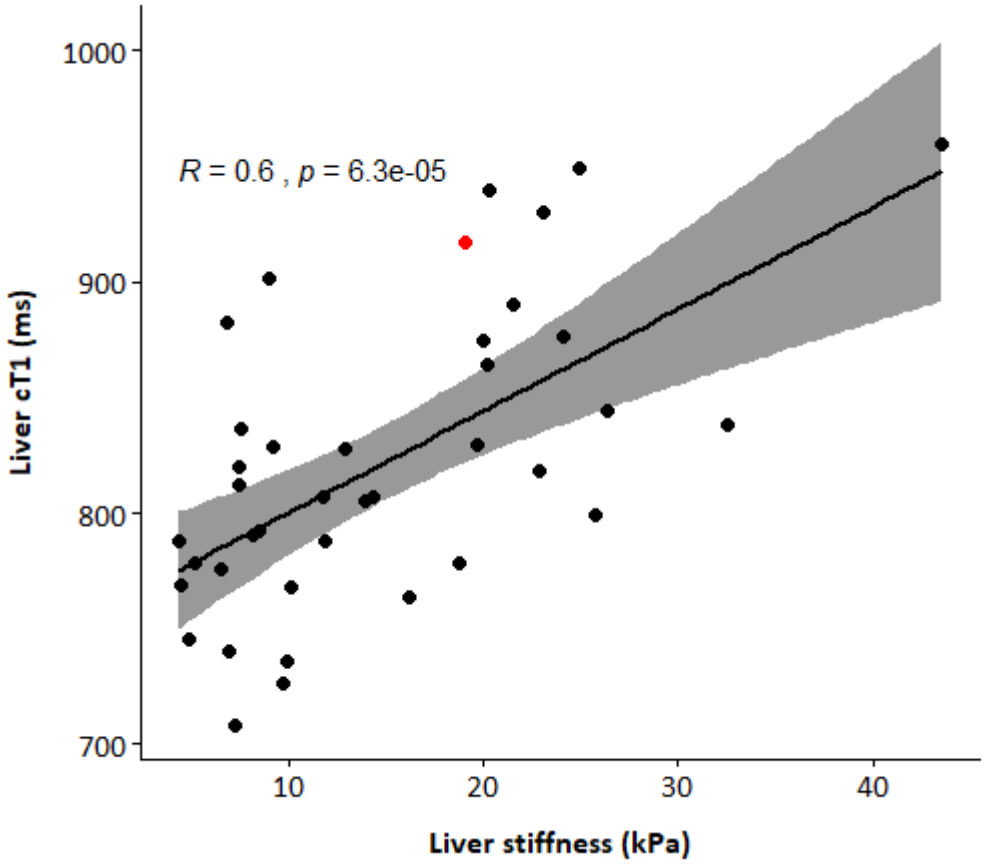


Figure 23: Liver cT_1 plotted against paired TE measurements. Both liver cT_1 and liver stiffness measured by transient elastography (TE) were taken on the same day, within an hour of each other, except one measurement which was taken within a week of MRI scan (coloured in red). Shaded area indicates 95% confidence interval, R denotes Spearman's rank correlation coefficient.

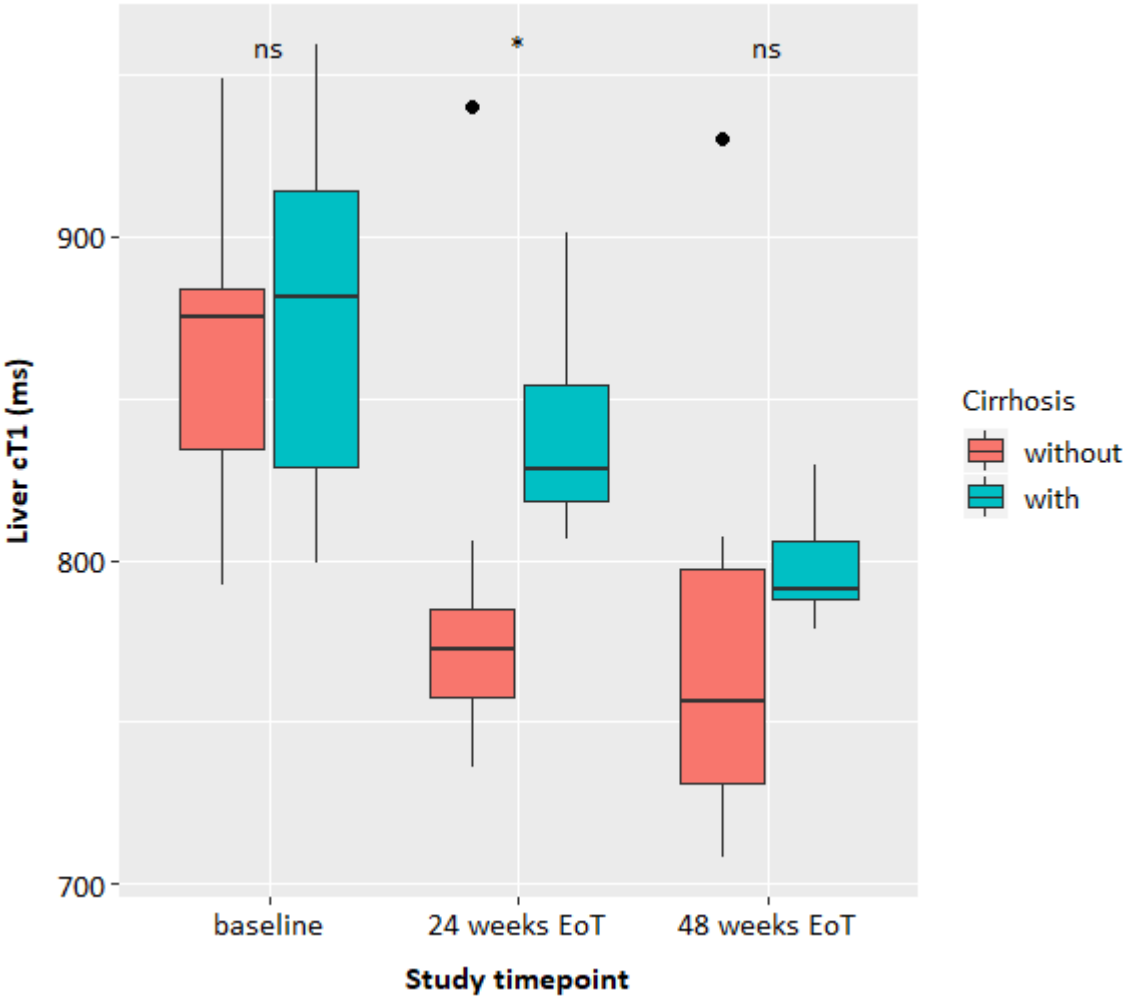


Figure 24: Comparison of liver cT_1 between patients with and without cirrhosis at baseline, 24 weeks EoT and 48 weeks EoT. * denotes $p < 0.05$. Cirrhosis was defined histologically (Ishak stage ≥ F5) or evidence of decompensation. EoT – end of treatment.

4.6 Discussion

In this study liver cT_1 has been used to track liver fibroinflammatory disease in the setting of DAA treatment of chronic HCV. We have shown a significant decrease in liver cT_1 in patients with cirrhosis and mild-to-moderate fibrosis following DAA treatments for chronic HCV from baseline to 24 and 48 weeks EoT. This suggests that there is both short and medium-term improvement in liver fibroinflammatory disease after successful DAA treatment for chronic HCV. We have observed that patients without cirrhosis show earlier improvements than those with cirrhosis, but the effect persists in patients with cirrhosis for longer, resulting in similar fibroinflammatory improvement in both groups of patients by 48 weeks EoT.

Another study with a similar number of patients (n = 17) had shown decreases in liver T_1 after DAA-achieved SVR[185]. This study followed up patients once after treatment, for varying, shorter periods than in this study (median 22 (IQR:3-79) days). We followed up at fixed timepoints of 24 and 48 weeks after EoT to assess longer term changes and included all patients with HCV undergoing DAA treatment. Our study investigated only those with baseline TE > 8kPa, thereby investigating higher-risk patients that are more likely to require close follow-up in a clinical setting. The study from the Nottingham group[185] did not report any exclusion criteria, representing a more real-world context; but it is unclear whether they excluded any participants with excess liver iron, as previously[184]. They did not report patients' alcohol consumption, as we did. Treatment in that study was by sofosbuvir plus ledipasvir or daclatasvir, with or without ribavarin; ours were a broader range of DAAs as prescribed by the clinical care team (see Table 12). In comparison to that study by Bradley et al[185], in which early reductions in liver T_1 of 35 ms (at 1.5 T) were observed, we found reductions at 24 weeks

EoT of liver cT_1 of 70 ms from our study (at 3 T) – this is to be expected as the magnitude of liver T_1 increases (almost linearly) with field strength. We postulate that the later, smaller reductions in cT_1 in our study may be due to a combination of residual/remaining inflammation and fibrosis reversal. However, it is more likely that most of the reduction in cT_1 was due to inflammation as liver fibrosis reversal has shown to be a slow process[268].

In this study we not only observed reductions in liver cT_1 but also in liver fat following DAA therapy. Changes in both these parameters have been observed in response to treatment for other chronic liver conditions. In particular, data from a recent open-label multi-centre study investigating the ability of a fat-reducing therapy (NGM282) for resolution of non-alcoholic steatohepatitis (NASH) showed that 12 weeks of treatment by NGM282 in different doses (1 mg and 3 mg daily, n = 24 and 19 patients in each group respectively) resulted in reductions of 78 and 82 ms in patients with histologically proven NASH[190]. We have observed similar decreases in liver cT_1 (88 ms), albeit over a longer time period. The main difference that could affect the significance of these changes are that in the NGM282 trial, patients had a high starting level of liver fat (17-19%), and experienced large decreases in liver fat (11 %), most of which occurred over a short period of time (6 weeks). It is known that high levels of liver fat can artificially affect shMOLLI-acquired T_1 (and therefore cT_1) measurements[269]. This may indicate an artificially large decrease in cT_1 in the NGM282 study, whereas in our study, although we observed small, significant changes in liver fat over 48 weeks (but not 24 weeks), our cohort had much lower baseline liver fat, with many starting below 5%, indicating the cT_1 reduction observed in this Chapter is independent of any fat effect. The NGM282 study also showed that histological responders had a more marked reduction in cT_1 than in non-

responders (92 vs 64 ms); this is similar to our data that show a more marked decrease in patients who did not drink alcohol in excess than in those who did, potentially inhibiting any histological improvement in their liver condition.

Patients without cirrhosis showed a steeper initial decrease in liver cT_1 than those with cirrhosis. We have also shown that liver cT_1 continues to decrease between 24 and 48 weeks EoT, albeit at a lower rate than between baseline and 24 weeks EoT. Single liver cT_1 measurements cannot separate the individual contributions of liver fibrosis and inflammation, and presently the only method that can do this is liver biopsy. However, these results indicate that relative temporal cT_1 response could differentiate between patients with cirrhosis vs non-cirrhosis and inflammation vs fibrosis reduction. Paired biopsy would be needed to confirm this but was outside the scope of this study. It also indicates that most of the improvement in fibroinflammation in patients without cirrhosis is observed by 24 weeks EoT.

If the cT_1 values of patients with and without cirrhosis decline to more similar levels over a longer follow-up it may indicate risk of HCC development is equal and may allow reduction in the frequency of ultrasound assessment for HCC in patients with cirrhosis – much larger cohorts and longer follow-up would be needed to fully assess this. For example, a retrospective study of 204 patients with PSC[270] showed that an increase in liver stiffness assessed by MRE was independently associated with risk of hepatic decompensation – a similar, but prospective study design would be appropriate for this.

We observed a weaker correlation between liver cT_1 and TE after treatment than prior to treatment ($r = 0.134$ vs 0.547). Whilst TE reportedly correlates well with histological liver

fibrosis[213,271,272], a study has shown it to have poorer diagnostic ability to discriminate F4 cirrhosis in patients post treatment for hepatitis C than in patients pre-treatment[273]. Using a cut-off of 12 kPa, in that study 21% of patients with TE < 12 kPa had F4 cirrhosis and in total 5/13 (38%) patients with F4 cirrhosis post-treatment were misclassified by that same TE cut-off. In our study, it is likely that a significant reduction in inflammation has affected the post-treatment correlation between liver cT_1 and TE.

Multiple studies have reported significant decreases in TE measurements and blood-based biomarkers in patients undergoing DAA treatment in as little as 12 weeks and up to as long as 12 months EoT using TE[254–260,274], in accordance with our findings. Liver stiffness assessed by magnetic resonance elastography (MRE) has shown to decrease after 6 months in a pilot study by Guo et al[275] of 8 kidney transplant recipients with chronic HCV being treated with DAAs and has also shown decreases in trials involving pharmacological therapy for NASH[276]. A study by Harrison et al[230] assessing liver cT_1, MRE and TE in response to GR-MD-02 (galactoarabino-rhamnogalaturonate) for therapeutic effect on NASH, showed no clinical effectiveness of the treatment. However, repeat measurements of cT_1, MRE and TE yielded coefficients of variation (CoV) for these 3 biomarkers of 3.1%, 11% and 40% respectively. This indicates that both MRE and TE may be less well suited to longitudinal studies as their variance may falsely indicate change in patient's liver health. This further bolsters the case for liver cT_1 to be used as a primary endpoint in clinical trials for NASH and/or antifibrotic therapies.

Liver fat reduced significantly between baseline and 48 weeks EoT. Lipid accumulation is intrinsically associated with HCV and is frequently noticed histologically in chronic HCV,

especially in genotype 3 infection. We did not observe significant differences in liver fat between genotypes; this is likely due to our small cohort. Historically, steatosis has been associated with lower SVR rates to interferon therapy[277] and shown to be an independent risk-factor for HCC development with chronic HCV-related cirrhosis[278]. Reduction in liver fat may therefore also improve patients' outcomes, but this is at odds with the results from Chapter 3. Our results suggest that DAA treatment may have liver fat-reducing properties regardless of whether patients have clinical steatosis or not.

A small but significant increase in LIC was observed from baseline to 48 weeks EoT. LIC is a surrogate marker, calculated from, and inversely proportional to the T_2* relaxation time. T_2* itself is a composite of the T_2 relaxation time and magnetic field inhomogeneity. It is likely that liver tissue T_2 has fallen (as observed by Bradley et al[185]) with improvement in liver fibroinflammation, causing a decrease in T_2* and increase in iron quantification in turn, rather than an increase in actual tissue iron. It is worth noting than when excluding patients who drank excess alcohol, no significant change in LIC was observed, but an increase was observed in each patient who consumed excess alcohol Figure 21D. This indicates that the increase in LIC is driven solely by the patients drinking excess alcohol.

Patients drinking excess alcohol had higher liver fat and LIC across all timepoints in the study; a study observed similar findings with respect to higher histological liver iron in patients with both ArLD and HCV than HCV alone[279]. The reference value for higher than normal iron is 1.7 mg/g[224] and our patients were at the higher end of the normal spectrum. It is known that patients with ArLD frequently present with evidence of iron overload, and even moderate

alcohol consumption can increase markers of iron storage[280]. The only participants whose liver cT_1 did not decrease between each time point were those that had been drinking heavily – it is unclear exactly how the potentially opposing effects of viral eradication and excess alcohol is affecting liver tissue and thus our liver cT_1 measurements.

Liver cT_1 > 825 ms has been shown in Chapter 3 of this thesis to predict higher risk of clinical liver-related outcomes including HCC. Only 2/15 patients had mean liver cT_1 > 825 ms at 48 weeks EoT (1 was drinking excess amounts of alcohol, the other had had HCC diagnosed and ablated twice). We observed no re-occurrence of HCC in any DAA-treated patient; however, this was in a small cohort and over a relatively short follow-up. Longer follow-up in larger cohorts is needed to assess the prognostic significance of this cut-off in these cohorts. Furthermore, the above cut-off may not apply to patients who have undergone fibro-inflammatory reversal as structural changes in the liver and residual fibroinflammation may require a different cut-off.

The greatest limitation of this study is the lack of paired histology, however MRS is well-validated against histology and liver cT_1 has been previously shown to correlate well with histological fibrosis[203] and due to sampling variation and interobserver variability[281,282], liver biopsy is not a true gold standard against which to measure. Furthermore, research has shown that the current staging system for fibrosis may not be optimal in the context of fibrosis regression in viral hepatitis[283].

In conclusion, this chapter has shown that liver T_1, when corrected for presence of iron (cT_1) can be used to assess fibroinflammatory response to DAA treatment for chronic HCV. Due to the relatively short follow-up period, the reduction in cT_1 is likely due to a reduction in inflammation, rather than regression of fibrosis, which has been shown to be a slow process[268]. Further studies should be carried out to investigate whether the same effect is seen when other insults to the liver e.g. excess alcohol consumption are stopped. Longer term follow-up of HCC (re)development is also needed to evaluate the prognostic capability of cT_1 in patients after DAA treatment.

Chapter 5. T_1 and cT_1 of abdominal organs for the assessment of portal hypertension

The work in section 2 of this chapter was submitted as a paper:

Chang J, *Jayaswal A*, Cheng L, Tay M, Gogna A, Tan H, Too C, Low A, Tan C. Multiparametric MRI of liver and spleen predicts clinically significant portal hypertension and variceal bleeding.

5.1 Abstract

Background & Aims: Chronic liver disease (CLD) can lead to portal hypertension (PH), which is associated with development of varices and hepatic decompensation. The current gold standard measure of PH is the hepatic venous pressure gradient (HVPG); patients also undergo endoscopy for detection and grading of varices. Both measures are invasive to different extents. Non-invasive markers of the spleen and liver have been shown to be associated with PH. We therefore aimed to assess spleen T_1/cT_1 and liver cT_1 for the assessment of PH in multiple centres.

Patients & Methods: Patients were recruited from 3 centres in Oxford (UK), Singapore and Cincinnati (US). Adult patients undergoing paired magnetic resonance imaging (MRI) and HVPG measurements were recruited from Oxford and Singapore. The primary aim was diagnosis of clinically significant PH (CSPH) defined by HVPG ≥ 10mmHg.

Results: *Oxford*: Spleen cT_1 correlated well with HVPG (r=0.46, p=0.03), and could predict CSPH with AUC 0.86, which rose to AUC 0.90 in logistic regression with platelet count. Liver cT_1 did not correlate with HVPG and could not predict CSPH. *Singapore*: Liver cT_1 could not predict CSPH but could predict HVPG > 12 mmHg (p=0.03). Spleen cT_1 did not correlate with HVPG and could not predict CSPH but could predict future variceal bleeding (AUC: 0.74), especially in patients on nonselective betablockers (AUC: 0.97). Similar spleen results were obtained when excluding patients with non-cirrhotic PH (NCPH), but liver cT_1 performed better for diagnosis of CSPH (AUC 0.81), with liver cT_1 + bilirubin on logistic regression predicting CSPH with AUC 0.93.

Conclusions: T_1/cT_1 mapping of abdominal organs either alone, or in combination with certain simple blood tests are promising biomarkers for the non-invasive assessment of PH and should be evaluated further in larger cohorts.

5.2 Introduction

Chronic liver disease (CLD) is associated with the development of liver fibrosis and eventually cirrhosis. Portal hypertension (PH), often a result of CLD, is defined by an abnormally elevated portal pressure gradient (PPG). The PPG is the difference in pressure between the portal vein and the inferior vena cava (IVC) and represents the portal perfusion pressure of the liver. Normal PPG levels are 1-5 mmHg; levels $\geq$ 10 mmHg indicate clinically significant portal hypertension (CSPH), the recognised threshold for a multitude of complications to begin arising[24]. In the Western world, the main cause of PH is liver cirrhosis. Other causes of PH are collectively termed non-cirrhotic portal hypertension (NCPH).

The direct effects of PH are the opening up of collaterals between the portal and systemic circulations, causing blood flow to bypass the liver.

Any occurrence or situation that causes either an increase in portal blood flow or increase in resistance to portal blood flow can potentially lead to PH. This is shown by applying Ohm's law $\Delta P = F \times R$, where P is pressure, F is blood flow and R is resistance to flow.

Portal pressure is normally maintained within a certain range (1-5 mmHg) by the liver sinusoids that adjust their resistance to any changes in portal flow. If, as in liver cirrhosis, these sinusoids and their connecting networks become damaged, intrahepatic resistance to flow increases. Damage can be caused by fibrosis progression, nodule formation and intrahepatic vasoconstriction.

PH can be categorised into 3 types; prehepatic, intrahepatic (sinusoidal and presinusoidal) and posthepatic. These are dependent on the location of the underlying cause of PH.

Prehepatic PH is caused by issues affecting the splanchnic, portal and mesenteric veins i.e. before blood flow reaches the liver e.g. portal vein thrombosis.

Intrahepatic PH is dependent on increased resistance to blood flow, usually due to progression of fibrosis/cirrhosis. Furthermore, when the location of increased resistance within the liver is taken into account, intrahepatic PH is further divided into sinusoidal and presinusoidal components, with the former mainly caused by cirrhosis and the latter by idiopathic PH and schistosomiasis for example.

Any obstruction to blood flow out of the liver is defined as posthepatic PH. The location of this obstruction can be anywhere from the outer part of the liver up to the inferior vena cava (IVC). Posthepatic PH can be caused by Budd-Chiari syndrome for example.

One of the most common complications of PH is the formation of gastro-oesophageal varices (present in approx. 50% of patients with a new diagnosis of cirrhosis[284]). PH leads to the development of collateral vessels that lead blood flow away from the liver into the systemic circulation. These collateral vessels develop mainly at the stomach (gastric varices) and the oesophagus (oesophageal varices). When the PPG becomes too high, they are unable to contend with this increased pressure and become fragile and prone to bleeding, with serious and sometimes fatal results. The 6-week mortality rate in patients with variceal bleeding is 20%[285].

The definition of ascites is the accumulation of free fluid in the peritoneal cavity. PH is the commonest cause of ascites. A consequence of cirrhosis and ascites is the risk of developing infections, termed spontaneous bacterial peritonitis (SBP).

Hepatic Encephalopathy (HE) in the context of cirrhotic PH is a neuropsychiatric disorder stemming from complications in CLD. Its main symptoms include fine motor and cognitive impairment, which lead to emotional and behavioural disturbances. In certain cases, HE can also lead to a state of coma, brain oedema, permanent damage to the nervous system and death. PH and increased pressure opens up venous pathways for intestinal toxins to bypass the liver and enter general blood circulation, with ammonia believed to be the main culprit in HE. This excess ammonia passes through the blood-brain barrier and increases the level of ammonia in the central nervous system (CNS) to toxic levels. Normally the liver would detoxify and regulate the concentration of ammonia, but in HE, due to blood flow bypassing the liver, the brain is affected by it.

In practice it is very difficult to measure the PPG as this would need direct puncture of the portal vein; access of the portal vein is through the spleen, or through the liver, all of which have risk of complications, in particular bleeding. There is no communication between the systemic and portal circulations. The gold standard measure for assessing the severity of PH is the hepatic venous pressure gradient (HVPG) measurement, obtained via transjugular catheterisation, which closely correlates with the PPG[286].

$$HVPG = WHVP - FHVP,$$

where WHVP is the wedged hepatic venous pressure and FHVP is the free hepatic venous pressure.

HVPG measurement has been shown repeatedly in numerous studies to be an excellent prognostic tool[26] and both single and repeat measures yield clinically important information (Table 16).

Table 16: Value of HVPG measurements in prognosis.

Single HVPG measurements	
≥ 6 mmHg	definition of PH
≥ 10 mmHg	Defines CSPH- risk of developing ascites, varices, decompensation[24], HCC[25].
≥ 12 mmHg	Variceal bleeding[287]
≥ 16 mmHg	Increased chance of mortality[23]
> 20 mmHg	Treatment failure and mortality in acute variceal bleeding[288,289]
≥ 22 mmHg	Mortality in acute alcoholic hepatitis[290]
Repeat measurements	
Reduction to ≤ 12 mmHg	Prevention of variceal bleeding and reduction in size of varices[291,292].
Reduction of ≥ 20% from baseline	Lower risk of first variceal bleeding event[292,293]

Despite the utility of the HVPG, it is invasive, expensive and only carried out in specialist tertiary centres[22] due to the requisite level of expertise.

In clinical practice, patients are screened for oesophageal varices by gastroscopy; this procedure is also (minimally) invasive, and not tolerated in all patients due to pharyngeal reflex. In a patients perspective study, a fifth of patients experienced moderate to severe pain during the procedure[294].

The Baveno VI criteria (platelet count > 150,000 mm^{-3} and liver stiffness assessed by transient elastography (TE) < 20 kPa) have been shown to perform well in ruling out patients without varices needing treatment (VNT) for unnecessary endoscopy, but a recent systematic review and meta-analysis of studies validating these criteria showed that application of the criteria only avoided upper endoscopy (EGD) in about 33% of patients[295], and still had high rates (440-

560 per 1000 EGDs) of unnecessary EGDs. If a surrogate non-invasive marker of PH or predictor of presence of varices and/or VNTs could be found, it would have significant clinical utility as identified by the Baveno VI Consensus Workshop on Methodology of Diagnosis and Therapy in PH[27].

PH causes congestion and fibrosis in the spleen, which increases its water content. Consequently, many encouraging biomarkers of PH involve the spleen[50,104,296–298]. In a proof-of-concept study carried out at our centre, in 19 patients, splenic cT_1 correlated with HVPG measurements ($r = 0.69$, $p = 0.001$) and could identify patients with clinically significant PH (CSPH) defined as HVPG > 10 mmHg with high accuracy (AUROC 0.92[193]). This technique could therefore act as a screening tool to rule out those who do not need HVPG or endoscopy – saving money that could be spent on other health service resources.

This promising pilot data required further validation; therefore, the overarching aim of this chapter was to gather further data from multiple centres and investigate whether spleen cT_1 could predict the presence of PH as defined by either HVPG ≥ 10 mmHg or radiological evidence of PH. Since the liver-specific iron-correction has not been validated in the spleen and therefore lacks physiological rationale, spleen T_1 was also assessed for the same outcomes.

Numerous studies have shown TE of the liver (reviewed by Berzigotti[47]) and magnetic resonance elastography (MRE) of the liver to correlate with HVPG[299]. A study has also shown uncorrected liver T_1 to correlate with HVPG measurements[117]. Liver cT_1 was therefore also investigated for the same aims.

Through further data acquired at the John Radcliffe Hospital, Oxford and a collaboration between Perspectum Diagnostics (PD) and a centre at Singapore General Hospital (SGH) we were able to further investigate T_1 and cT_1 parameters of abdominal organs as surrogate biomarkers of PH.

5.2 Aims

The specific aims of this chapter were to:

1) Update the pilot data obtained at OCMR with additional scans using spleen T_1 and the updated and standardised cT_1 algorithm[188] to:
    a) Assess liver cT_1, spleen T_1, and spleen cT_1 for:
        i) correlation with HVPG values
        ii) diagnostic accuracy for HVPG $\geq$ 10 mmHg and > 16 mmHg
        iii) diagnostic accuracy for presence of varices
    b) perform interobserver variability analysis
    c) conduct multivariate analysis for diagnostic accuracy of spleen T_1/cT_1 for HVPG $\geq$ 10 mmHg

2) Examine data from SGH to:
    a. Assess liver cT_1, spleen T_1, and spleen cT_1 for:
        i. correlation with HVPG values
        ii. diagnostic accuracy for HVPG $\geq$ 10 mmHg and > 16 mmHg
    b. Conduct multivariate analysis for diagnostic accuracy of HVPG $\geq$ 10 mmHg and 16 mmHg.

c. Evaluate the ability of spleen cT_1 to predict patients with future episodes of variceal bleeding
d. Examine a) and c) excluding patients on beta blocker treatment

This chapter has been presented as a collation of separate sub-studies and are discussed both as standalone studies and in conclusion the results are summarised as a whole.

5.3 OCMR cohort

5.3.1 Introduction

Previous work as described in the Introduction has shown spleen cT_1 to have a good correlation with HVPG[193]. Since publication more patients have been recruited to the study and the cT_1 algorithm has been updated and refined (Chapter 2).

5.3.2 Aims

The aim of this section therefore was to:

a) extend and update the analysis using the newly standardised cT_1 technique (standardised to field strength and vendor) of both liver and spleen cT_1 and also assess spleen uncorrected T_1 for:
   i) correlation with HVPG
   ii) diagnostic accuracy of HVPG $\geq$ 10 mmHg and > 16 mmHg
   iii) diagnostic accuracy for presence of varices
b) perform interobserver variability analysis for spleen cT_1
c) conduct multivariate analysis of non-invasive markers for diagnostic accuracy of HVPG $\geq$ 10 mmHg.

5.3.3 Methods

Patient selection

Consecutive patients referred for liver biopsy (with/or without portal pressure measurements) for suspected cirrhosis were recruited from the John Radcliffe Hospital

(Oxford, UK) between August 2013 and June 2018[b]. Those referred for percutaneous biopsy consented to change their procedure to transjugular biopsy with portal pressure measurements. Patients were adults and exclusion criteria were contraindication to MR scans. Patients were also excluded from the analysis if their portal pressures could not be calculated or they did not attend an MR assessment. All patients gave written, informed consent and the study was approved by the local Research Ethics Committee (ref 13/SC/0243).

Study visit protocol

Study visits included MR assessment, blood sampling and liver stiffness measurements by transient elastography (TE). Study visits occurred prior to transjugular biopsy, which was performed by an interventional radiologist blinded to the MR data. Serum-based scores of fibrosis (aspartate aminotransferase/alanine aminotransferase (AST/ALT) ratio, AST to platelet ratio index (APRI) and Fibrosis-4 (FIB-4)) were calculated.

MR protocol and MR analysis

Transverse abdominal shMOLLI images were acquired with parameters: 5(1)1(1)1 scheme, TI: 110 ms, 110 ms + RR, 110 ms + 2RR, 110 ms + 3RR, 110 ms + 4RR, 190 ms and 270 ms, TR/TE: 234.97/1.05 ms, TR of individual bSSFP readouts 2.5 ms, slice thickness 8.0 mm, GRAPPA acceleration factor 2, flip angle: 35°. FOV: patient dependent, matrix size: 144 x 192.

[b] Recruited by Drs Michael Pavlides and Christina Levick. Data from 19 of these patients have been reported elsewhere in Dr Michael Pavlides' DPhil thesis and in a recent publication by Levick et al.[193] I do not claim this as original work, rather as a reanalysis with updated MRI methods and increase in numbers (5, 26%). This section also serves a reference point for the remaining studies.

T_2* maps were acquired with a multi-echo GRE acquisition was used with parameters: TR: 500 ms, TEs: 1.25, 2.46, 3.69, 4.92, 6.15, 7.38, 8.61, 9.84 ms, 6 mm slice thickness, GRAPPA acceleration factor 2, flip angle: 20°. DIXON method used for fat fraction analysis. FOV: patient dependent, matrix size 104 x 128.

Magnetic resonance spectroscopy (MRS) was performed using a STimulated Echo Acquisition Method (STEAM)[204] sequence with the following parameters: voxel size 20 x 20 x 20 mm^3 (shim volume 40 x 40 x 40 mm^3), TE = 10 ms, mixing time = 7 ms, TR: 2 s for water-suppressed and TR: 4s for unsuppressed data with 1024 points acquired at bandwidth 50 Hz. 3 scans of 5 repetitions with water suppression were performed and 1 scan of 3 repetitions with water suppression turned off was performed.

Transjugular biopsy and portal pressure measurements

The HVPG/transjugular biopsy procedure was carried out on patients under conscious sedation with constant monitoring of vital signs after an overnight fast and 72-hour pause in any antiplatelet/anticoagulant medication. Under local anaesthetic, with the patient in the supine position, the right internal jugular vein was catheterised after which a straight vascular catheter was introduced to measure the wedged and free hepatic venous pressures (WHVP, FHVP). The catheter was advanced until it achieved a wedged position that was confirmed radiologically, thus obtaining the WHVP. The catheter measured the FHVP 2-4 cm from the confluence of the hepatic vein with the inferior vena cava. Pressures were measured by a digital pressure measurement device connected to the catheter via a pressure transducer. Pressure readings were allowed to stabilise for 30-40 seconds before being recorded. HVPG calculated as WHVP – FHVP. Biopsies were then obtained using a 19G transjugular biopsy needle and reported centrally, blinded to any MR data.

Statistical analysis

Boxplots were generated to compare spleen cT_1 values of patients with HVPG < 10 mmHg and HVPG ≥ 10 mmHg. Patients were also stratified into groups of increasing severity (HVPG < 10 mmHg, 10 mmHg ≤ HVPG <16 mmHg and HVPG ≥ 16 mmHg) and their spleen cT_1 values compared. Wilcoxon non-parametric tests were performed to assess significance between these groups. Correlation between spleen cT_1 and HVPG were calculated using Pearson's correlation coefficient. Receiver operating curves (ROC) were generated for spleen cT_1's diagnostic ability to identify patients with HVPG ≥ 10 mmHg and ≥ 16 mmHg. The best performing cut-offs (Youden index) for prediction of severities of PH were generated. ROC analysis was also performed for presence of varices.

Multivariate models using a combination of up to 2 clinical variables (to avoid overfitting) were investigated to assess prediction of HVPG ≥ 10 mmHg.

Interobserver analysis was independently performed blinded to the HVPG results by two analysts (myself and a senior clinician Dr Michael Pavlides) on 19 datasets. The Pearson's correlation coefficient was calculated, intraclass coefficient calculated, and Bland-Altman plots were generated.

5.3.4 Results

33 patients were recruited and 24 were included in the analysis. One patient missed the planned HVPG/transjugular biopsy procedure due to development of severe alcoholic hepatitis post study visit. Five only had their WHVP measured and 3 did not attend their study visit for MR assessment (Figure 25). Baseline characteristics for the whole cohort and those

with HVPG < 10 mmHg and HVPG ≥ 10 mmHg are described in Table 17, with only Ishak fibrosis stage, APRI score and spleen cT_1 being significantly different in those with and without CSPH.

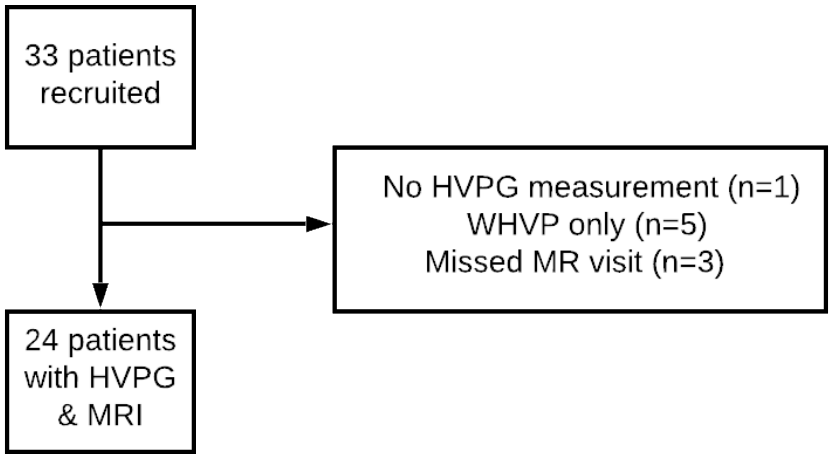


Figure 25: Study flow diagram (OCMR).

Most patients were male (n = 18, 75%), with main disease aetiologies being NAFLD, ArLD and viral hepatitis. Median age was 58 (46-67) and BMI was 28 (24-32) kg/m^2. 14 had liver cirrhosis defined as Ishak fibrosis stage F5-6.

An outlying patient was identified who had only mild fibrosis on biopsy and no other features of cirrhosis or PH. Liver synthetic function tests (albumin, prothrombin time and platelet count) were normal. It is possible the HVPG measurement in this case was inaccurate.

Spleen T_1 and cT_1 by severity of PH

Spleen cT_1 was higher in patients with HVPG ≥ 10 mmHg than in those with HVPG < 10 mmHg (1358 ms vs 1231 ms, p = 0.0024, n = 10 vs 14, Figure 26b, Table 18), and excluding the outlying patient the difference increased further (1370 ms vs 1231 ms, p = 0.0002, n = 9 vs 14). Spleen

T_1 was higher in patients with HVPG $\geq$ 10 mmHg than in those with HVPG < 10 mmHg (1351 ms vs 1204 ms, p = 0.0009, n = 10 vs 14, Figure 26a), and excluding the outlying patient the difference increased further (1366 ms vs 1204 ms, p < 0.0001, n = 9 vs 14).

When grouped by severity of PH, patients with HVPG 10-15 mmHg, but not HVPG ≥ 16 mmHg had significantly higher spleen cT_1 and T_1 than those with HVPG < 10mmHg. There was no difference in patients with 10 ≤ HVPG <16 and HVPG ≥ 16 mmHg (Figure 26). When excluding the outlier, both spleen cT_1 and T_1 were higher in both groups of patients with CSPH than in those with HVPG < 10 mmHg, but there was still no difference between the severities of CSPH.

Table 17*:* OCMR baseline characteristics. Characteristics provided for the whole cohort, those with HVPG < 10 mmHg and HVPG ≥ 10 mmHg.

	All (n = 24)	**HVPG < 10 mmHg (n = 14)**	**HVPG ≥ 10 mmHg (n = 10)**	**P value**
Age	58 (46-67)	63 (48-72)	57 (47-58)	0.278
Male (%)	18 (75%)	11 (79%)	7 (70%)	0.668
BMI (kg/m^2)	27.7 (24.07-32.03)	25.8 (23.3-30.4)	30.6 (25.6-32.2)	0.403
Cirrhosis (%)	14 (58%)	6 (43%)	8 (80%)	0.08
Ascites (%)	4 (20%)	3 (21%)	1 (10%)	0.497
Ishak stage	5 (2 – 6)	2 (1-5)	6 (5-6)	**0.047**
Bilirubin (μmol)	13.5 (10.8-22.5)	13.0 (8.8-18.0)	17.0 (11.3-25.5)	0.428
ALT (IU/l)	42 (22.8-70.3)	37 (23.5-82.0)	48.0 (24.0-65.3)	0.815
ALP (IU/l)	175 (116-224)	160 (106-208)	197 (150-233)	0.305
Albumin (g/l)	38 (33.8-45)	38 (34-45)	39 (34-41)	0.792
AST (IU/l)	40 (35-74)	37 (35-64)	45 (37-75)	0.422
GGT (IU/l)	113 (50-275)	105 (48-275)	231 (78-273)	0.483
Platelets (x 10^9)	132 (110-171)	135 (120-163)	107 (97-178)	0.151
PT (s)	14.4 (13.1-16.1)	13.5 (12.6-15.4)	14.9 (14.3-17.0)	0.120
AST/ALT	1.09 (0.88-1.73)	1.05 (0.88-1.30)	1.27 (1.10-1.80)	0.262
APRI	0.742 (0.610-1.20)	0.667 (0.512-0.744)	1.06 (0.849-1.580)	**0.002**
FIB-4	3.41 (2.16-4.83)	2.59 (1.98-3.68)	3.72 (3.16-5.84)	0.154
Liver TE (kPa)	17.2 (9.4-36.1)	12.0 (9.0-53.3)	32.5 (10-34.3)	0.948
MR data				
Spleen cT_1 (ms)	1289 (1188-1351)	1231 (1173-1267)	1358 (1321-1392)	**0.002**
Spleen T_1 (ms)	1254 (1189-1337)	1204 (1177-1252)	1351 (1317-1387)	**0.0009**
Liver cT_1 (ms)	913 (858-1028)	894 (856-967)	989 (885-1225)	0.154
Liver fat (%)	3.65 (1.41-7.58)	5.00 (1.30-7.50)	2.3 (1.74-13.7)	0.931

p values given for difference between the two subgroups. Only Ishak fibrosis stage, APRI and spleen cT_1 were significantly different between those with and without CSPH. Significance set at $p < 0.05$. No Bonferroni correction applied. BMI (body mass index), ALP (alkaline phosphatase), GGT (gamma-glutamyl transferase, PT (prothrombin time). Data presented as median and (interquartile) ranges or number and %.

Table 18: Spleen cT_1 and T_1 values for different severities of PH (OCMR).

	HVPG < 10 mmHg median (IQR)	HVPG 10 – 15 mmHg median (IQR)	HVPG ≥ 16 mmHg median (IQR)
Spleen cT_1	1231 (1173-1267)	1358 (1335-1392)	1345 (1278-1385)
Spleen T_1	1204 (1177-1252)	1351 (1336-1386)	1342 (1266-1382)

Spleen T_1 /cT_1 and liver cT_1 correlation with HVPG

Spleen cT_1 showed a moderate correlation with HVPG measurement (r = 0.47, p = 0.022, Figure 27b), which increased when the outlier was excluded (r = 0.61, p = 0.002, Figure 27d). Spleen T_1 showed a moderate correlation with HVPG of r = 0.54 (p = 0.007, Figure 27a), rising to r = 0.71 (p = 0.0002, Figure 27c) with exclusion of the outlier. Liver cT_1 did not correlate with HVPG (r = 0.29, p = 0.16) even when excluding the outlying patient (r = 0.35, p = 0.097).

Spleen T_1 and cT_1 ROC analysis for severity of PH and presence of varices

Spleen cT_1 could identify those patients with HVPG ≥ 10 mmHg with AUC 0.8571 (95% CI: 0.6784-1.00, p = 0.01, Table 19), increasing to 0.9365 (95% CI: 0.8441-1.000, p < 0.01) when the outlier was excluded. The identified best cut-off of 1289 ms performed with sensitivity 90%, specificity 79%, positive predictive value (PPV) 75% and negative predictive value (NPV) 92%, improving to 100%, 79% 75% and 100% respectively with exclusion of the outlier.

Spleen T_1 could identify those patients with HVPG ≥ 10 mmHg with AUC 0.8857 (95% CI: 0.7714-1.00, p < 0.01), increasing to 0.9683 (95% CI: 0.9013-1.000, p < 0.01) when the outlier was excluded. The identified best cut-off of 1289 ms (same as cT_1) performed with sensitivity

90%, specificity 93%, positive predictive value (PPV) 90% and negative predictive value (NPV) 93%, improving to 100%, 93% 90% and 100% respectively with exclusion of the outlier.

Spleen cT_1 could not identify those patients with HVPG > 16 mmHg with AUC 0.6625 (95% CI: 0.2578-1, p = 0.34), with best cut-off of 1316 ms for higher sensitivity and 1372 for higher specificity. Spleen T_1 could identify those patients with HVPG > 16 mmHg with AUC 0.6625 (95% CI: 0.266-1, p = 0.32). These had high 95% confidence intervals and therefore should be interpreted with caution. In addition, there were only 4 patients with HVPG > 16 mmHg, so this comparison is unlikely to be meaningful.

17 patients had recent endoscopy data available on presence of varices, of which 6 had varices present. Spleen cT_1 could not identify presence of varices (AUC 0.5909; 95% CI: 0.2423-0.9395, p = 0.74) and spleen T_1 could not either (AUC 0.7121; 95% CI: 0.3912-1, p = 0.22). The large 95% CIs associated with these results mean they should be interpreted with caution.

Spleen cT_1/T_1 was unable to distinguish groups for both prediction of HVPG > 16 mmHg and prediction of varices because AUC CIs included 0.5.

Multivariate analysis

In multivariate analysis, no other non-invasive test except spleen cT_1 and spleen T_1 were independently predictive of HVPG $\geq$ 10 mmHg (data not shown).

In logistic regression, when spleen cT_1 was combined with platelet count, the AUC for prediction of HVPG $\geq$ 10 mmHg rises from 0.86 to 0.90 (95% CI: 0.72-1.00, p = 0.007), with the correlation with HVPG rising from 0.47 to 0.61 (95% CI: 0.28-0.82, p = 0.0014). When spleen

T_1 was combined with platelet count, the AUC for prediction of HVPG $\geq$ 10 mmHg rose from 0.8857 to 0.8929 (95% CI: 0.7102-1.00, p = 0.004), with the correlation with HVPG rising from 0.51 to 0.63 (95% CI: 0.30-0.82, p = 0.001).

No other combination of parameters with spleen cT_1/T_1 performed better than spleen cT_1/T_1 and platelet count (no more than 2 were included to avoid overfitting, only non-invasive variables in which there were no missing data were included).

When excluding the outlier, the spleen cT_1 + platelet count model had AUC of 1.00 (95% CI: 1-1) and correlation of 0.87 (95% CI: 0.72-0.95, p < 0.0001) with same results for T_1.

Table 19: Diagnostic accuracy of spleen cT_1/T_1 for severity of PH and varices

		Cut-off (ms)	AUC	95% confidence intervals:	P value	Sens %	Spec %	PPV %	NPV %
Spleen cT_1	HVPG $\geq$ 10	1289	0.8571	0.68-1	0.01	90	79	75	92
	HVPG $\geq$ 16	1316	0.6625	0.26-1	ns	75	65	30	93
	HVPG $\geq$ 16	1372	0.6625	0.26-1		50	90	50	83
	varices	1324	0.5909	0.24-0.94	ns	67	73	57	80
Spleen T_1	HVPG $\geq$ 10	1289	0.8857	0.77-1	< 0.01	90	93	90	93
	HVPG $\geq$ 16	1308	0.6625	0.27-1	ns	75	70	33	93
	varices	1264	0.7121	0.39-1	ns	83	73	63	89

Interobserver variability

Spleen cT_1 interobserver variability analysis yielded a Pearson's correlation coefficient of 0.89 (p < 0.0001) indicating a strong agreement between observers and good intraclass correlation coefficient of 0.83 (0.62-0.93).

Bland-Altman analysis showed a bias of -28 ms (95% CI: -3 – -52 ms) with upper limits of agreements 76.7 ms (95% CI: 121-32 ms) and lower limits of agreement -132 ms (95% CI: -88 – -177 ms).

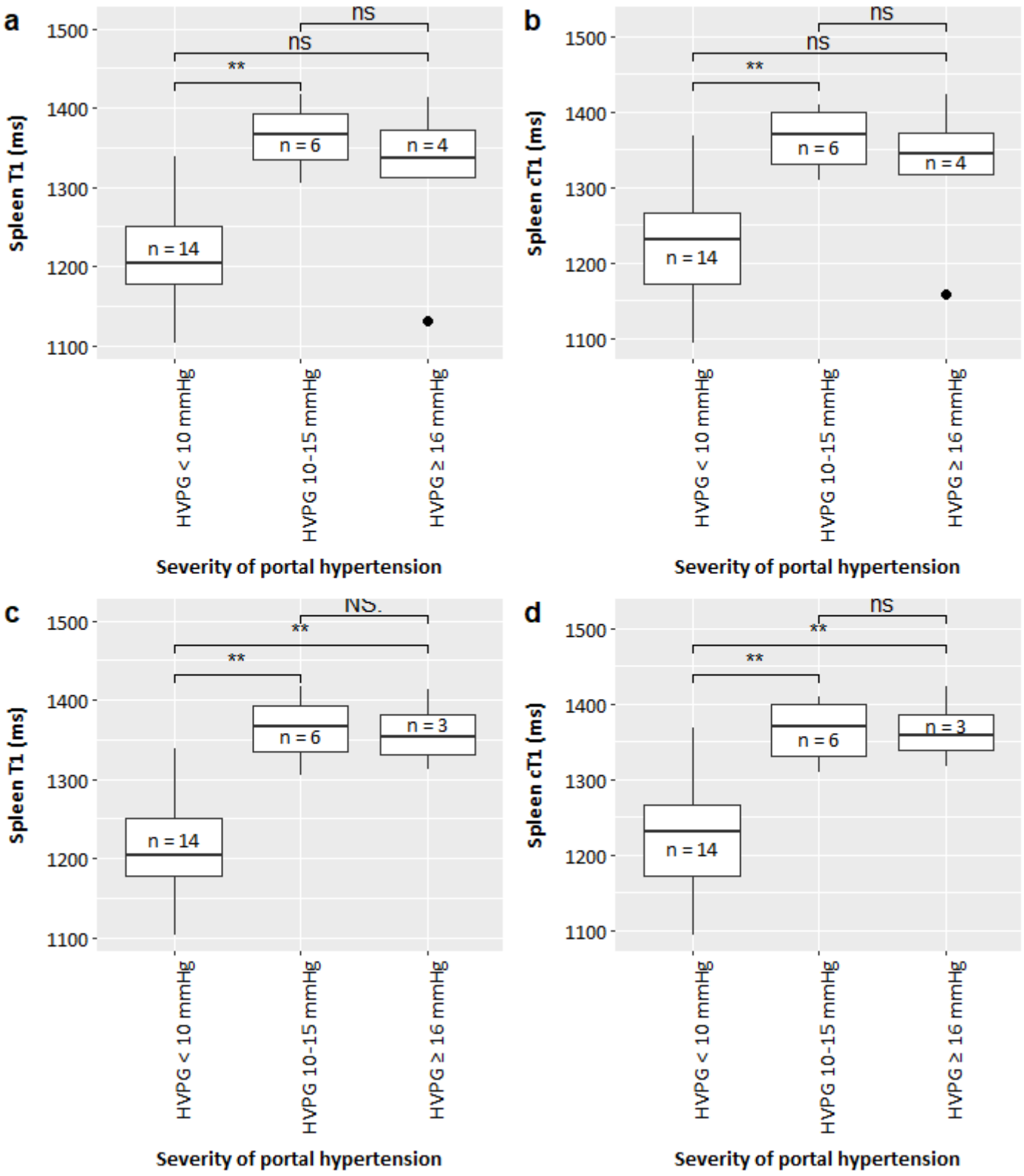


Figure 26: **Stratification of severity of portal hypertension by spleen T_1/cT_1 in patients with chronic liver disease.** * denotes $p < 0.05$, ** denotes $p < 0.01$. HVPG – hepatic venous pressure gradient. a) spleen T_1, b) spleen cT_1, c) spleen T_1 excluding outlier, d) spleen cT_1 excluding outlier.

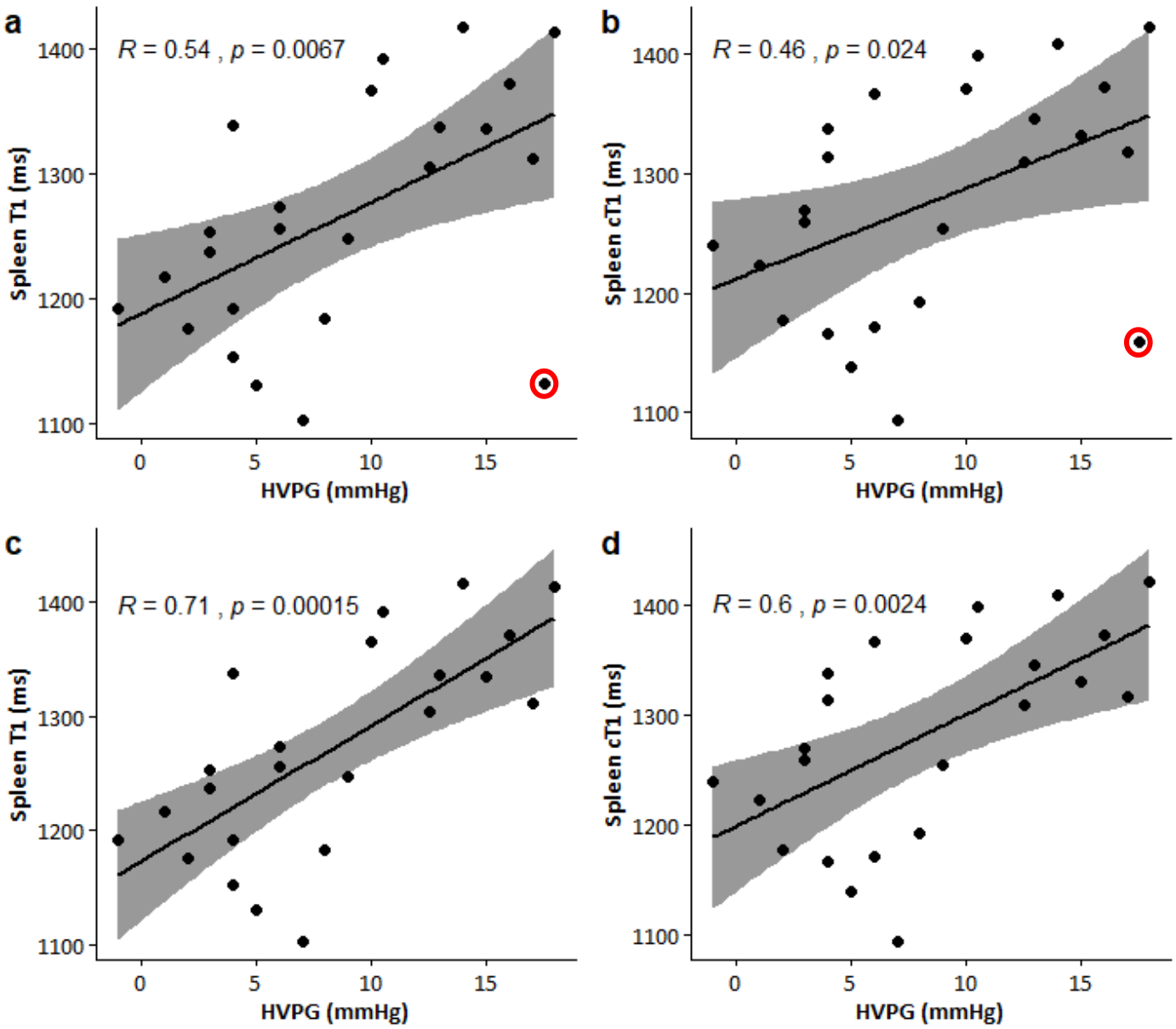


Figure 27: **Correlations between HVPG and spleen T_1/cT_1.** a) spleen T_1 and HVPG including all 24 patients, b) spleen cT_1 and HVPG including all patients, c) spleen T_1 and HVPG excluding the outlying patient with normal clinical tests but high HVPG, d) spleen cT_1 and HVPG excluding the outlying patient. Shaded area denotes 95% confidence interval. R represents the Pearson's correlation coefficient. Outlying patient circled in red in a) and b).

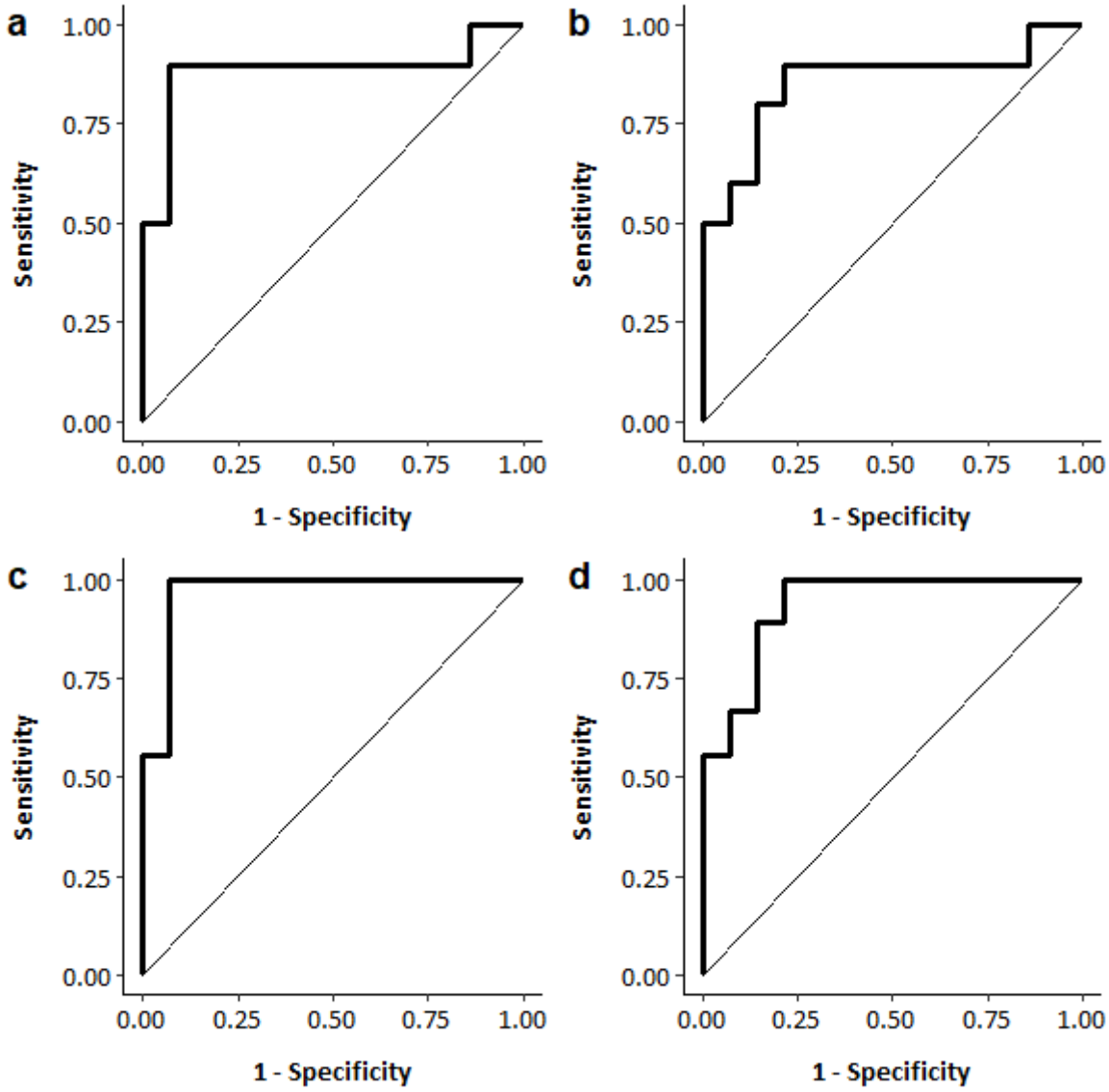


Figure 28: **Receiver operating curves (ROC) to show classification of patients with HVPG ≥ 10 mmHg** by a) spleen T_1 (AUC 0.8857, 95% CI: 0.7714-1.00, $p < 0.01$), b) spleen cT_1 (AUC 0.8571, 95% CI: 0.6784-1.00, $p = 0.01$), c) spleen T_1 excluding the outlying patient (AUC 0.9683, 95% CI: 0.9013-1.00, $p < 0.01$), d) spleen cT_1 excluding outlying patient (AUC: 0.9365, 95% CI: 0.8441-1.00, $p < 0.01$).

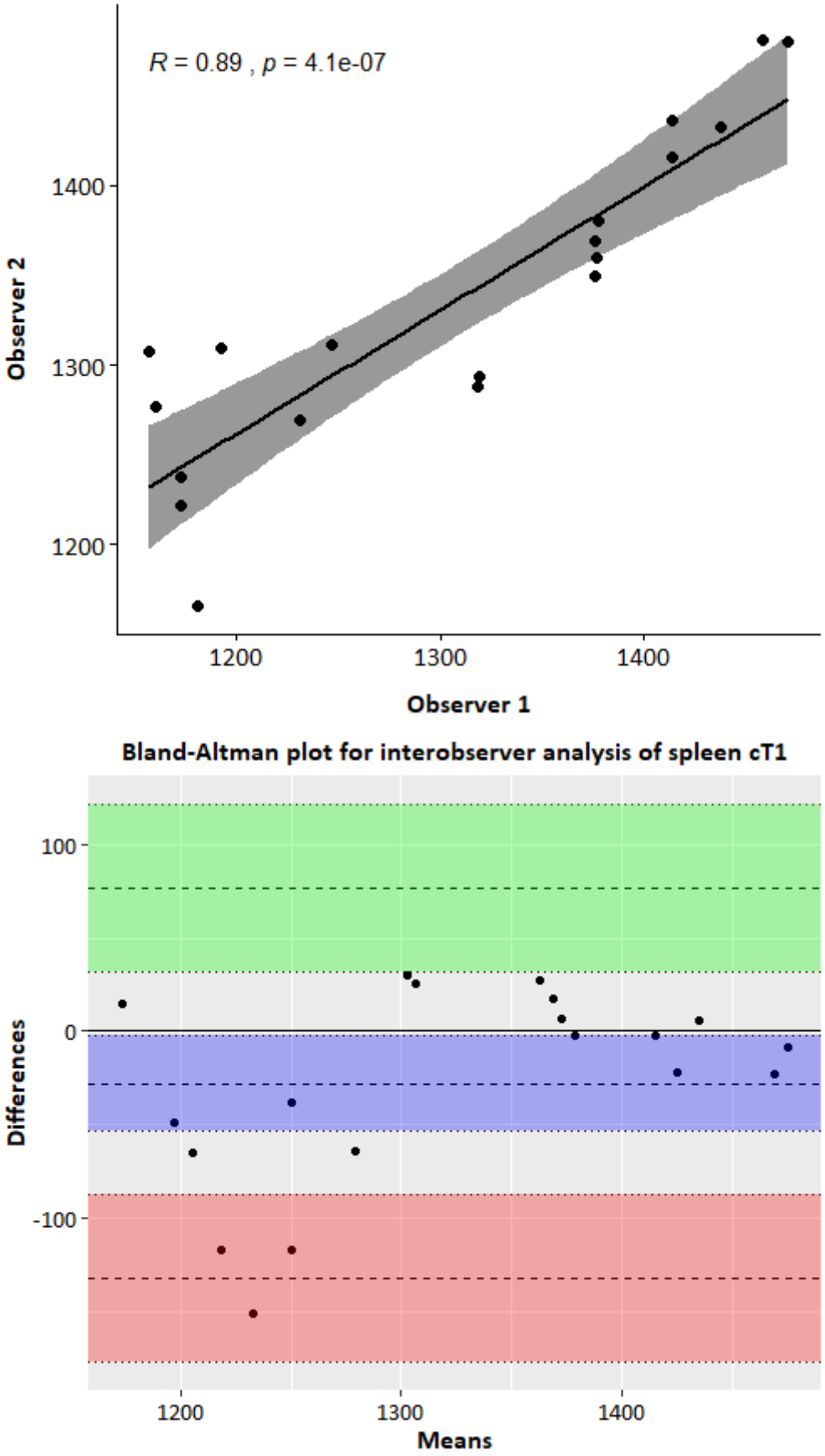


Figure 29: Interobserver variability analysis. (Top) Correlation between two independent observer's analysis of spleen cT_1 values in n = 19 patients. R represents Pearson's correlation coefficient. (Bottom) Bland-Altman plot for the same analysis with shaded areas outlining upper (green) and lower (red) limits of agreement and bias (purple).

5.3.5 Discussion

Discrimination of patients with HVPG ≥ 10 mmHg vs those with HVPG < 10 mmHg was good, and improved with the exclusion of the outlying patient, as reported previously by Levick et al[193]. The cut-offs derived in this study using the new standardised cT_1 algorithm were lower (1289 vs 1376 ms) than reported previously, but this is not unexpected, as the algorithm generally maps the 'old' cT_1 to a 'new standardised' lower cT_1, as shown by the updated liver cT_1 cut-off for prediction of clinical outcomes (825 ms vs 875 ms (reported as fibrosis and inflammation (LIF) score of 2 in previous publications[187,203])).

Spleen cT_1 could not distinguish between different severities of CSPH, which is a finding in line with other publications investigating the correlations of spleen TE, liver TE and spleen T_1 with HVPG, in which these parameters correlated well with HVPG up to 12 mmHg, but not above[49,117,300]. Whilst it is postulated that the reason for this lack of association with HVPG ≥ 10 mmHg is due to extrahepatic factors for liver stiffness/liver fibrosis-based measures, we initially hypothesised that these factors would be incorporated in spleen cT_1 measurements. However, with increasing PH the same effect might be encountered with the spleen; increasing congestion and fibrosis may only be correlated with severity of HVPG up until a point, as the opening up of collateral vessels or varices may then reduce the degree of splenic congestion, thereby accounting for the lack of discriminatory ability at greater severity of PH. With larger numbers, sensitivity analysis could be performed to assess whether presence of varices at high HVPG levels affects correlation with spleen T_1/cT_1. Another explanation for the 'flattening out' of spleen T_1/cT_1 values at higher HVPG could be due to reaching the T_1 value of blood, indicating that there may be a ceiling after which spleen T_1/cT_1 shows no further correlation with HVPG.

There was one outlying patient with normal clinical liver function tests and clinical examination, but high (17.5 mmHg) portal pressure measurement. The inclusion/exclusion of this patient significantly altered the correlation coefficients of spleen T_1/cT_1 with HVPG. This highlights not only the high technical ability required to perform HVPG measurements but also the small size of our dataset which is liable to change significantly with one datapoint.

Spleen T_1 showed a moderate AUC for presence of varices, (0.71) but did not reach significance, probably due to small numbers in the group. Additionally, due to the small number, we included presence of all varices, most of which were grade 1. In another cohort of 91 patients with compensated advanced CLD (cACLD) recruited here at OCMR[301], spleen T_1 was shown to have 100% sensitivity for identifying varices needing treatment (VNTs), which is a more clinically useful measurement. Larger cohorts would allow us to investigate the ability to identify patients with VNTs.

Interobserver agreement was good, with $r = 0.89$ ($p < 0.001$). In comparison to our intraclass correlation coefficient (ICC, 0.83), ICCs for spleen stiffness assessed by point shear-wave elastography was 0.96[302] and 0.76 for acoustic radiation force imaging (ARFI), (albeit with ARFI measurements only available in those with larger spleens)[303], and 0.98 by MRE[304]. Caution must be taken in that our analysis was performed over only 19 patients, and larger numbers are required to obtain a more accurate value.

In the correlation plot of spleen cT_1 vs HVPG (Figure 27B) the correlation seems tighter amongst the high HVPG/spleen cT_1 values, and less so at the lower values. The reason for this is unclear having examined our data set fully. Initial thoughts were that the iron correction would have affected the lower spleen T_1 values since they can overlap with the upper limits

of liver T_1 values, a region in which the iron correction would have a more significant effect. However, both the T_1 and cT_1 vs HVPG plots are very similar, so it is likely the main contributing factor is the greater interobserver variability at lower values as shown in Figure 29A. We have not been able to postulate a particular reason for this and examination of the data has not yielded any explanations such as high spleen T_1/cT_1 heterogeneity. However, this could be an impact of the Liver*MultiScan*™ colourmap if differences in cT_1 at the lower range map to a smaller discernible range of colour on the screen.

We postulated that using spleen T_1 rather than cT_1 may yield a weaker correlation, but in fact spleen T_1 performed similarly for all tests performed, and even outperformed cT_1 consistently but not significantly. This is a difference to previous analysis within our group, in which uncorrected spleen T_1 performed inferiorly to spleen cT_1 using the original correction algorithm. The reason for this is likely two-fold. The new iron correction method has a different range of utility, still being designed for the liver and not the spleen – in previous versions the correction extended to the spleen range of T_1 values, which are generally higher than those found in the liver. It is now not deemed operable in that extended range since there would not be any effect within it. Secondly, the older correction did not have the field strength correction included, which will also contribute to the differences.

Overall, examining our data more closely, in many of the abdominal slices, the T_2* images yielded lower T_2* R-squared values (a measure of how reliable the T_2* quantification is) within the spleen than in the liver (0.971 vs 0.983, p = 0.015). This may have contributed to the larger variation in overall spleen T_2* values (IQR: 15.1 – 31.9 ms vs 13.9 – 22.5) and variability within the spleen T_2* ROIs (IQR: 8.41 vs 3.38 ms, p = 0.03), which could manifest itself in less accurate spleen cT_1 values. Further technical development work in this area could reduce interobserver

variability. In general, it has often been more difficult to obtain reliable T_2* values due to poor quality R^2 maps as the gradient echo (GRE) sequence is very susceptible to patients' motion. Whether T_2* introduces variation should be addressed by looking at the interobserver agreement between spleen T_1 which we did not do here.

No other variable other than spleen T_1 and spleen cT_1 were independently associated with presence of CSPH, but the combination of spleen T_1/cT_1 and platelet count performed excellently. Thrombocytopaenia is associated with PH[305] and combinations of platelet count with other non-invasive markers are used in many of the composite scores in hepatology; the APRI score, LSPS (liver stiffness to platelet ratio score), Baveno VI criteria for screening for VNTs and others, so it's inclusion as an adjunct variable is not surprising. One of the desired characteristics of a non-invasive tool for the assessment of PH is the detection of varices. A Cochrane database systematic review also showed platelet count to spleen length to have sensitivity of 85% for diagnosis of oesophageal varices[306]. A composite biomarker including spleen cT_1/T_1 could therefore have great clinical utility for both diagnosis of CSPH and varices.

Liver cT_1 did not correlate with HVPG, which is in accordance with the results published by our group using the older version of the iron correction method[193]. These results are in contradiction to many of the results relating to liver T_1[117], TE and MRE[99]'s correlations with HVPG. Our cohort contained some unexpectedly low liver cT_1 values in patients with histologically confirmed cirrhosis, which is likely to have impacted on the correlation with HVPG. It is unclear what this reason is, but we postulate that in some patients with advanced fibrosis/cirrhosis, just as liver volumes are observed to decrease[307], water content and thus liver T_1 may also decrease. Secondly, there may be advanced fibrosis, but reduced inflammation in some of these patients, which would also result in reduced liver cT_1. In these

cases, spleen T_1/cT_1 would be a useful additional marker of disease severity. Another factor that could be causing low liver T_1 values is the interaction with liver fat. We did not look at fat fraction as a modulator of low liver T_1 values in multiple regression in this work but is a suggested next step.

5.4 Singapore General Hospital cohort

5.4.1 Introduction

Following on from the extension to the pilot study at our centre, a collaboration was undertaken between PD and A/Prof Jason Chang, Head & Senior Consultant, Department of Gastroenterology & Hepatology, Singapore General Hospital (SGH). This collaboration involved recruitment of patients undergoing clinically indicated HVPG measurements and abdominal MRI scans at SGH, with image and statistical analysis performed by myself. The collaboration was made in the attempt to obtain similar data and perform similar analysis as in Section 1, but in a larger cohort and in a different centre.

5.4.2 Aims

The aims therefore of this section were: to assess in all patients and excluding those on non-selective beta-blocker (NSBB) treatment the ability of liver cT_1 and spleen T_1 and cT_1 either on univariate or multivariate analysis for:

a) correlation with HVPG

b) ability to stratify patients with significant portal hypertension (CSPH) as defined by HVPG ≥ 10 mmHg

c) ability to stratify patients with CSPH at clinically relevant thresholds of HVPG > 12 and > 16 mmHg

d) prediction of clinical events associated with development of CSPH including episodes of variceal bleeding

5.4.3 Methods

Patient selection

This study was designed as a prospective observational single-centre study to evaluate the diagnostic performance of cT_1 to predict CSPH in a cohort of patients undergoing HVPG measurements.

Patients seen in the Department of Gastroenterology and Hepatology, SGH who were scheduled to undergo HVPG measurement for evaluation of PH were prospectively recruited to participate in this study. Inclusion criteria for participation included (1) age >21 years, (2) CLD with clinical indication for HVPG measurements, (3) ability to provide informed consent. Exclusion criteria included severe hepatic decompensation with refractory ascites, active HE or liver failure, contraindications to MRI including indwelling metallic devices, claustrophobia or creatinine clearance <30 ml/min, pregnancy, inability or unwillingness to provide informed consent. Recruitment for the study was performed between August 2017 and Mar 2019. The study protocol complied with the Declarations of Helsinki and was approved by the SingHealth Centralized Institutional Review Board (CIRB reference number: 2017/2425). Patient demographic and clinical data, including aetiology of liver disease, severity of cirrhosis, endoscopic findings and clinical events including variceal bleeding, were recorded in a standardized data form[c].

[c] Study protocol provided by Jason Chang, SGH. I did not carry out this study and not claim this work as my own, I performed all statistical analysis and spleen image analysis blinded to clinical data and HVPG measurements.

MR protocol

T_1 mapping: shMOLLI 5(1)1(1)1, 192x144 acquisition matrix, segments=84, bandwidth=246 Hz/Pixel, parallel imaging factor=2, 8mm slice thickness, flip angle=35°, FOV=440mm x 330mm, echo spacing (SSFP TR) =4.76ms, TE=1.93ms, initial TI=145ms, TI increment = 80ms.

T_2* mapping: 128x104 acquisition matrix, 6mm slices thickness, flip angle=20°, FOV=400mmx325mm, TE = 2.38ms, echo spacing=2.38ms, number of echoes = 8.

MRI scans were performed using a Siemens Magnetom Avanto-FIT, 1.5T scanner (Erlangen, Germany) All MRI scans were performed within 8 weeks of HVPG measurement.

HVPG measurement

HVPG measurements were performed by qualified interventional radiologists who are trained to perform HVPG, using a standardized HVPG protocol[308]. Briefly, ultrasonography-guided puncture of the internal jugular vein is performed using a strict aseptic technique. Position of the guidewire in the vein is confirmed by fluoroscopy. A 6 French (Fr) sheath is inserted over the guidewire and is secured in the internal jugular vein, through which a 4 Fr Multipurpose A catheter (Cordis, Johnson and Johnson, Florida, USA) is advanced into the hepatic vein under fluoroscopic guidance. The middle hepatic vein is preferentially cannulated, failing which the right hepatic vein is cannulated. The catheter is positioned no more than 2 cm into the hepatic vein for measurement of the free hepatic venous pressure (FHVP). An occlusion 5.5 Fr ‘over-the-wire’ Fogarty balloon catheter (Edwards Lifesciences, California, USA) is advanced distally into the appropriate position in the hepatic vein and the balloon is inflated for measurement of wedge hepatic venous pressure (WHVP). After inflation of the balloon, 5 mL of contrast is injected and a venogram obtained to ensure total occlusion of the hepatic

vein. This step is essential to exclude the presence of intrahepatic portal venous shunts that would result in underestimation of the true WHVP. If total occlusion of the hepatic vein is not satisfactorily obtained, the balloon is deflated, and the catheter is repositioned to ensure adequate occlusion. Pressure readings are appropriately calibrated for measurement of venous pressures and adequate time (at least 45–60 seconds) is allowed for stabilisation of readings before a final measurement is recorded. FHVP and WHVP readings are repeated three times to obtain triplicate set of readings. Readings that differ by >2 mmHg are discarded, the position of the catheter and balloon is re-checked, and the measurement is repeated.

The major difference between this and the approach taken here at the John Radcliffe is the use of a balloon catheter rather than a wedge catheter, which is considered to have less variability and is more accurate than straight catheters[309].

Statistical analysis

Patients were stratified into those with and without HVPG ≥ 10 mmHg and non-invasive scores compared. The Wilcoxon non-parametric test were performed to assess significance between these groups. Correlations between each of spleen cT_1 and liver cT_1 and HVPG were calculated using Pearson's correlation coefficient. No correction for multiple comparisons were made.

Area under the receiver operating curve (AUROC) analysis was performed for the classification of HVPG ≥ 10 mmHg, > 12 mmHg (associated with increased risk of variceal bleeding) and > 16 mmHg (associated with increased risk of mortality), and prediction of

episodes of variceal bleeding by MR parameters and blood-based tests (albumin, ALT, AST, bilirubin, creatinine, GGT, haemoglobin (Hb) prothrombin time (PT) and platelet count).

Subgroup analysis was performed excluding those on nonselective beta-blocker (NSBB) treatment as changes in cardiac output may affect the spleen cT_1 values.

The best performing variables on univariate analysis (up to a maximum of 3 to avoid overfitting) were then assessed in general linear modelling/logistic regression analysis. Combinations of MR-derived parameters only (liver and spleen cT_1) were also assessed for classification of severity of CSPH and prediction of variceal bleeding.

Associations of MR parameters with HVPG were also assessed after excluding patients with NCPH, in whom HVPG may not reflect the true portal pressure.

5.4.4 Results

During the study period, a total of 56 subjects fulfilled the inclusion criteria and provided informed consent to participate in the study. 9 subjects did not complete the study – 2 withdrew consent as they were unwilling to proceed with MRI, 2 developed medical complications and were too ill to proceed with the study and 5 subjects missed their MRI appointments within the 8-week window period. In total, 47 subjects completed the study and underwent paired HVPG and MRI according to the study protocol (Figure 30).

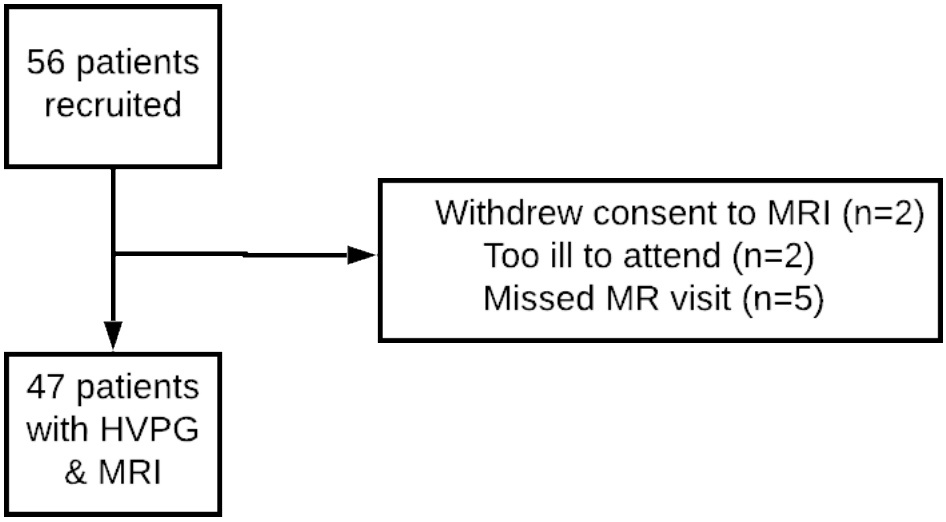


Figure 30: Study flow diagram (SGH)

Patients had median age 63 (54-67), were evenly split between male and female (53% male), had cirrhosis in 37/47 (79%) and main aetiologies were NASH (21%), hepatitis B virus (HBV, 15%) and alcohol-related liver disease (ArLD) 15%. 3 patients did not have cirrhosis on biopsy (two had perisinusoidal fibrosis not amounting to cirrhosis and one had nodular regenerative hyperplasia (NRH)), and 7 subjects had NCPH (3 with post-hepatic portal hypertension due to

cardiac disease, 2 with idiopathic noncirrhotic portal hypertension, 1 with myeloproliferative disease with left sided portal hypertension and 1 with Budd-Chiari syndrome).

Forty-six out of 47 patients had valid HVPG data. These 46 patients had paired multiparametric MRI and HVPG data, with spleen cT_1 available in all 46, and liver cT_1 available in 45 (1 failing quality control check). Only these 46 patients were included in analysis.

There were no statistically significant differences in any of the blood serum-based biomarkers except bilirubin ($p = 0.014$, Table 20) between those with HVPG ≥ 10 mmHg and < 10 mmHg. There were no statistically significant differences between those with HVPG ≥ 10 mmHg and < 10 mmHg in any of the MR-based biomarkers, with a trend towards significance shown for liver cT_1 ($p = 0.066$).

Median follow-up duration for variceal bleeding was 19 months (IQR 12-24). During this period, 14 subjects had episodes of variceal bleeding (30%).

Correlation of MR parameters with HVPG

Liver cT_1 ($r = 0.411$, $p = 0.005$) was correlated with HVPG, but not spleen cT_1 ($r = -0.035$, $p = 0.817$, Figure 31) nor spleen T_1 ($r = 0.04$, $p = 0.795$). Liver and spleen cT_1 were not significantly correlated ($r = 0.030$, $p = 0.844$).

Classification of severity of PH based on HVPG thresholds

Univariate analysis for the prediction of HVPG ≥ 10 mmHg, HVPG > 12 mmHg and > 16 mmHg are detailed in Table 21:

Table 20: SGH baseline characteristics. Characteristics provided for all patients, those with HVPG ≥ 10 mmHg and < 10 mmHg,

	All (n = 46)	**HVPG < 10 mmHg (n = 8)**	**HVPG ≥ 10 mmHg (n = 38)**	**P value**
Age	63 (54-67)	54 (51-58)	64 (57-67)	0.042
Male (n, (%))	24 (53%)	5 (63%)	19 (50%)	ns
Cirrhosis (%)	37 (79%)	2 (25%)	35 (92%)	<0.001
Varices	33 (72%)	5 (63%)	28 (74%)	ns
Variceal bleed	15 (34%)	4 (50%)	11 (29%)	<0.01
Aetiology				<0.01
NASH	10 (21%)	0 (0%)	10 (26%)	
ArLD	7 (15%)	1 (12.5%)	6 (16%)	
HBV	7 (15%)	1 (12.5%)	6 (16%)	
Cryptogenic	6 (13%)	0 (0%)	6 (16%)	
Other*	17 (36%)	7 (75%)	10 (26%)	
Bilirubin (μmol)	21.5 (14.3-35.0)	9 (6.5-15.0)	22.0 (18.3-35.0)	0.014
ALT (IU/l)	35 (25-46)	31 (24-42)	37 (26-47)	0.434
ALP (IU/l)	118 (90-170)	77 (63-126)	120 (92-163)	0.147
Albumin (g/l)	34 (30-40)	37 (31-40)	32 (30-40)	0.459
AST (IU/l)	51 (40-74)	52 (39-58)	52 (41-75)	0.562
GGT (IU/l)	114 (51-189)	66 (32-101)	131 (63-197)	0.063
Platelets (x 10^9)	127 (75-153)	161 (139-186)	97 (72-149)	0.054
PTT (s)	29.2 (26.0-32.5)	26.7 (25.2-28.9)	29.7 (26.1-32.7)	0.095
MR data				
Spleen cT_1 (ms)	1352 (1278-1428)	1372 (1330-1428)	1358 (1276-1428)	0.643
Spleen T_1 (ms)	1097 (1030-1182)	1140 (1056-1194)	1087 (1030-1176)	0.817
Liver cT_1 (ms)	805 (767-928)	787 (728-798)	824 (768-948)	0.066

Significance between the groups assessed by the Wilcoxon non-parametric test reported in the far-right column. Alanine aminotransferase (ALT), aspartate aminotransferase (AST), non-alcoholic steatohepatitis (NASH), alcohol-related liver disease (ArLD), hepatitis B virus (HBV), gamma glutamyl-transferase (GGT), prothrombin time (PT) and cT_1 (corrected T_1). *Other included HCV, Wilson's disease, NCPH, AIH/PBC, Caroli disease.

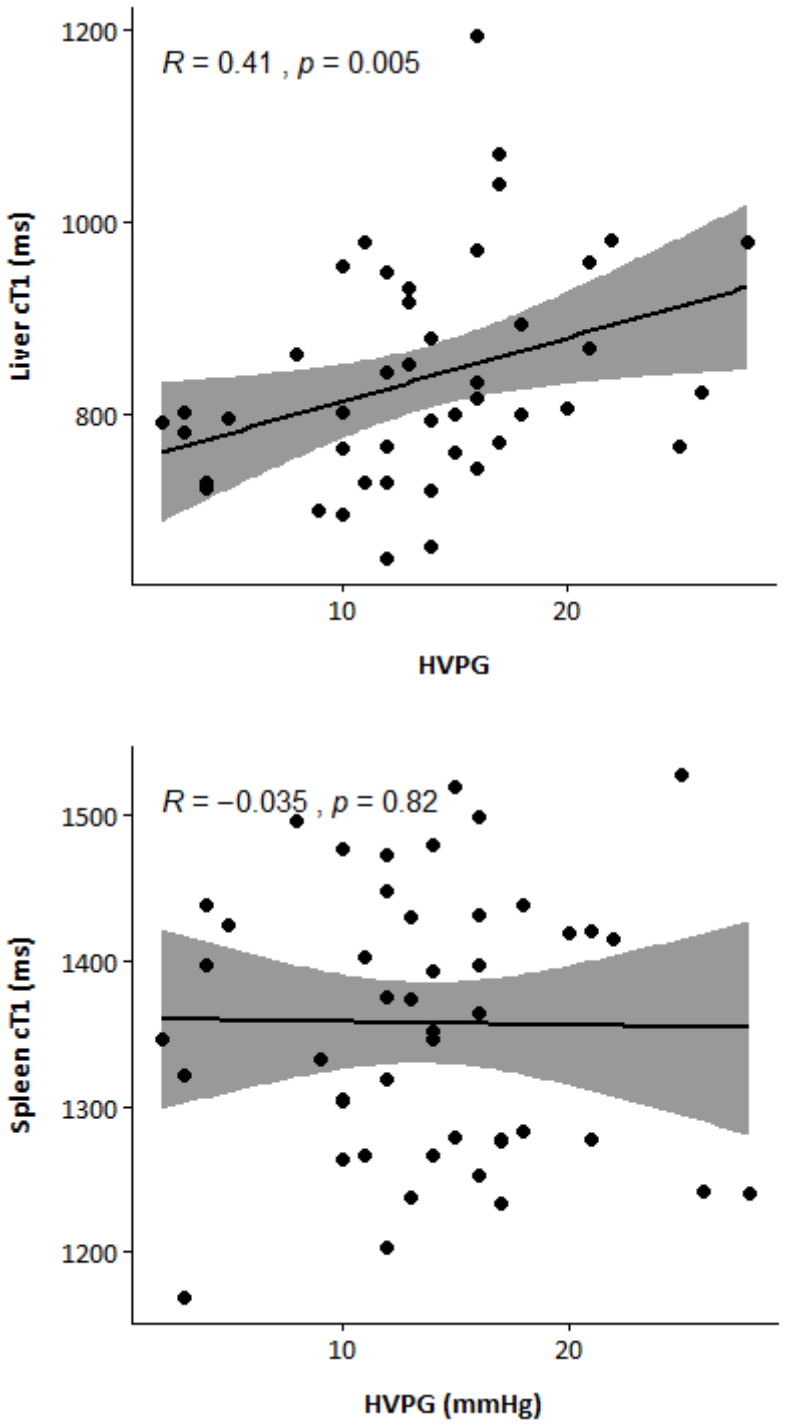


Figure 31: Correlation between hepatic venous pressure gradient (HVPG) measurements and abdominal cT_1. a) liver cT_1. Shaded areas represent the 95% confidence intervals. Significance was set at $p < 0.05$.

Table 21: Prediction of severity of portal hypertension based on clinically used HVPG thresholds. Biomarkers with highest consistent AUCs highlighted in bold to be assessed for multivariate analysis.

	Test	AUC	95% CI	p value	n with vs n without	AUC	95% CI	p value	n with vs n without	AUC	95% CI	p value	n with vs n without
	Spleen cT_1	0.5543	0.3309-0.7776	ns	38 vs 8	0.4912	0.3171-0.6654	ns	27 vs 19	0.5919	0.3851-0.7987	ns	34 vs 12
	Spleen T_1	0.5280	0.2745-0.7814	ns	38 vs 8	0.5039	0.3246-0.6832	ns	27 vs 19	0.6091	0.4192-0.7989	ns	34 vs 12
	Liver cT_1	**0.7111**	0.5451-0.8722	ns	37 vs 8	**0.7126**	0.5577-0.6832	**0.03**	26 vs 19	**0.7399**	0.5825-0.8973	**0.05**	33 vs 12
	Spleen length	0.7039	0.4976-0.9103	ns	38 vs 8	0.5175	0.4976-0.9103	ns	27 vs 19	0.6176	0.4337-0.8016	ns	34 vs 12
	Platelets	**0.7204**	0.4835-0.9572	ns	38 vs 8	**0.7183**	0.5517-0.8849	**0.01**	27 vs 19	**0.6961**	0.5395-0.8526	**0.05**	34 vs 12
	PTT	0.6929	0.4888-0.8969	ns	35 vs 8	0.5984	0.4104-0.7864	ns	26 vs 17	0.5284	0.3298-0.7270	ns	32 vs 11
	Bilirubin	**0.7970**	0.5619-1.000	ns	38 vs 7	**0.7819**	0.6290-0.9348	ns	27 vs 18	**0.7525**	0.6099-0.8952	ns	33 vs 12
	GGT	0.7583	0.5619-0.9550	ns	20 vs 6	0.5294	0.2778-0.7810	ns	17 vs 9	0.5799	0.3373-0.8224	ns	18 vs 8
	ALT	0.5905	0.3608-0.8201	ns	38 vs 8	0.5487	0.3605-0.7370	ns	27 vs 19	0.5662	0.387-0.7454	ns	34 vs 12
	Hb	0.5099	0.3358-0.6840	ns	38 vs 8	0.5039	0.3326-0.6751	ns	27 vs 19	0.5208	0.318-0.7236	ns	34 vs 12
	ALP	0.6661	0.3956-0.9366	ns	38 vs 8	0.5273	0.3386-0.7160	ns	27 vs 19	0.6348	0.4555-0.8141	ns	34 vs 12
	Albumin	0.5855	0.3779-0.7931	ns	38 vs 8	0.6365	0.4644-0.8085	ns	27 vs 19	0.5650	0.3826-0.7473	ns	34 vs 12
	Creatinine	0.7089	0.4994-0.9184	ns	38 vs 8	0.6501	0.4894-0.8108	ns	27 vs 19	0.5343	0.3294-0.7392	ns	34 vs 12

Liver cT_1, bilirubin and platelet count consistently had the highest AUCs for stratification of severity of PH, but still had high 95% CIs, which accounted for the non-significance of bilirubin at all severities and the non-significance of liver cT_1 and platelet count at HVPG $\geq$ 10 mmHg. Logistic regressions of liver cT_1, and either bilirubin or platelet count (maximum 3 variables to avoid overfitting) for each severity of PH (HVPG ≥ 10, > 12 and > 16 mmHg) had AUCs and 95% CIs described in Table 22. All of the combinations had large 95% CIs and again these results must be taken with caution.

Table 22: Logistic regressions of best performing tests on univariate analysis for prediction of severity of PH.

Test	AUC	95% CI	P value	> threshold vs < threshold
HVPG ≥ 10 mmHg				
Liver cT_1 + platelets + bilirubin	0.8031	0.5851-1	0.01	37 vs 7
Bilirubin + platelets	0.797	0.5603-1	0.01	38 vs 7
Liver cT_1 + bilirubin	0.7992	0.5815-1	0.014	37 vs 7
Liver cT_1 + platelets	0.7095	0.5404-0.8785	ns	37 vs 8
HVPG > 12 mmHg				
Liver cT_1 + platelets + bilirubin	0.7842	0.6436-0.9248	0.003	26 vs 18
Bilirubin + platelets	0.749	0.5915-0.9064	0.007	27 vs 18
Liver cT_1 + bilirubin	0.735	0.5794-0.8907	0.01	26 vs 18
Liver cT_1 + platelets	0.7814	0.6454-0.9173	0.003	26 vs 19
HVPG > 16 mmHg				
Liver cT_1 + platelets + bilirubin	0.750	0.5969-0.9031	0.02	32 vs 12
Bilirubin + platelets	0.6894	0.5266-0.8521	ns	33 vs 12
Liver cT_1 + bilirubin	0.7214	0.5558-0.8869	ns	32 vs 12
Liver cT_1 + platelets	0.7551	0.6051-0.905	0.02	33 vs 12

Prediction of episodes of variceal bleeding

45 patients had available data on episodes of variceal bleeding. For prediction of episodes of variceal bleeding, Hb, PTT and spleen cT_1/T_1 performed best (Table 23). A combination of these parameters did not yield any significantly superior predictive value (0.7395 (95% CI: 0.5765-0.9025)) to spleen cT_1/T_1 or PTT on their own.

Table 23: Prediction of future episodes of variceal bleeding by biomarker

Test	AUC	95% CI	P value	N with and without CSPH
Spleen cT_1	0.7362	0.563-0.9093	0.02	14 vs 31
Spleen T_1	0.7385	0.5742-0.9027	0.01	14 vs 31
Liver cT_1	0.5762	0.3988-0.7536	ns	14 vs 30
Spleen length	0.4793	0.2902-0.6683	ns	14 vs 31
Platelets	0.5288	0.3466-0.711	ns	14 vs 31
PTT	0.7759	0.6228-0.9289	0.02	13 vs 29
Bilirubin	0.6286	0.4494-0.8078	ns	14 vs 30
GGT	0.565	0.2348-0.8952	ns	5 vs 20
ALT	0.5795	0.4109-0.7481	ns	14 vs 31
Hb	0.7143	0.5487-0.8798	0.02	14 vs 31
ALP	0.6118	0.4306-0.793	ns	14 vs 31
Albumin	0.5092	0.3283-0.6901	ns	14 vs 30
Creatinine	0.5806	0.4128-0.7485	ns	14 vs 31
HVPG	0.4833	0.2848-0.6819	ns	14 vs 30

Multivariate analysis of MR parameters only, for assessment of all HVPG thresholds and episodes of variceal bleeding

Combinations of liver cT_1 and spleen cT_1 could not predict HVPG $\geq$ 10 mmHg (AUC 0.7264; 95% CI: 0.5649-0.8878, n = 37 vs 8, p = ns), but could predict HVPG > 12 mmHg with AUC

0.7186 (95% CI: 0.5641-0.8732, n = 26 vs 19, p = 0.03), HVPG > 16 mmHg with AUC 0.7449 (95% CI: 0.5668-0.9321, n = 12 vs 33, p = 0.03) and variceal bleeding with AUC 0.739 (95% CI: 0.5702-0.9076, n = 30 vs 14, p = 0.01).

Combinations of spleen length and spleen cT_1 predicted HVPG $\geq$ 10 mmHg with AUC 0.7434 (95% CI: 0.5413-0.9455, n = 38 vs 8, p = 0.02), could not predict HVPG > 12 mmHg (AUC 0.5478; 95% CI: 0.36-0.7355, n = 27 vs 19, p = ns) or HVPG > 16 mmHg (AUC 0.6569; 95% CI: 0.4797-0.834, n = 12 vs 34, p = ns) and could predict variceal bleeding with AUC 0.7327 (95% CI: 0.5566-0.9089, n = 31 vs 14, p = 0.01).

Combinations of spleen length and liver cT_1 predicted HVPG $\geq$ 10 mmHg with AUC 0.7973 (95% CI: 0.6183-0.9763, n = 37 vs 8, p = 0.01), HVPG > 12 mmHg with AUC 0.7146 (95% CI: 0.5573-0.8719, n = 26 vs 19, p = 0.02), HVPG > 16 mmHg with AUC 0.6995 (95% CI: 0.516-0.883, n = 12 vs 33, p = 0.04) but could not predict variceal bleeding with AUC 0.5786 (95% CI: 0.4044-0.7527, n = 30 vs 14, p = ns).

Using spleen T_1 in place of spleen cT_1 yielded similar results for the above multivariate analysis, data not shown for brevity.

Excluding those on beta-blocker treatment

When including only the patients on beta-blockers (n=19), spleen cT_1 performed excellently in identifying patients who went on to have variceal bleeds 0.9688 (95% CI: 0.898-1, p < 0.01) in 16 without vs 4 with episodes of bleeding. No other biomarker on univariate analysis performed as well (spleen T_1 0.9531 (95% CI: 0.8563-1, p < 0.01)), data not shown.

However, in this subgroup, performance of spleen for prediction of CSPH was not improved in comparison to the whole cohort (Table 24). Combinations of markers with so few patients

were decided against as there was significant likelihood the models may overfit, with the univariate AUCs all being non-significant and having large confidence intervals (Table 24). These results should therefore be taken with caution.

Table 24: Prediction of clinically significant portal hypertension (CSPH) by biomarker including only patients on beta-blocker treatment.

Test	AUC	95% CI	P value	n with and without CSPH
Spleen cT_1	0.6019	0.145-1	NS	18 vs 3
Spleen T_1	0.463	0-0.9578	NS	18 vs 3
Liver cT_1	0.549	0.3086-0.7894	NS	17 vs 3
Spleen length	0.5463	0.03221-1	NS	18 vs 3
Platelets	0.5	0-1	NS	18 vs 8
PT	0.5556	0.2595-0.8516	NS	18 vs 3
Bilirubin	0.6667	0.1135-1	NS	18 vs 3
GGT	0.6364	0.3151-0.9576	NS	11 vs 2
INR	0.5455	0.01423-1	NS	11 vs 3
Urea	0.9259	0.7665-1	NS	18 vs 3
AST	0.713	0.3093-1	NS	18 vs 3
Hb	0.6111	0.3734-0.8488	NS	18 vs 3
ALP	0.537	0-1	NS	18 vs 3
Albumin	0.6296	0.1579-1	NS	18 vs 3
Creatinine	0.7963	0.5396-1	NS	18 vs 3

Exclusion of patients with NCPH

When excluding those with NCPH, the correlation between spleen cT_1 and HVPG was r = -0.11 (p = 0.52); and between spleen T_1 and HVPG was r = -0.18 (p = 0.28). Neither spleen cT_1 nor T_1 were able to predict CSPH with any great degree of accuracy (AUCs 0.5766 and 0.6929 respectively, T_1 data not shown for brevity). A summary of the results from the exclusion of

the 7 patients with NCPH is detailed below. There were 37 patients with HVPG $\geq$ 10 mmHg vs 3 with < 10 mmHg. There was no significant differences in results in comparison to the whole cohort, except liver cT_1 performed better for diagnosis of CSPH (HVPG $\geq$ 10 mmHg) in this subgroup (AUC 0.82, Table 25). In particular the combination of liver cT_1 and bilirubin on logistic regression yielded an AUC of 0.92 (p = 0.05) for HVPG $\geq$ 10 mmHg.

The subgroup of patients on NSBB were exactly the same as in the full cohort, therefore with identical prognostic value of spleen cT_1 for variceal bleeding as described above. HVPG was unable to predict future episodes of variceal bleeding in this subgroup of patients (AUC: 0.64, 95% CI: 0.24-1.00, p = 0.4).

Table 25: Summary of diagnostic and prognostic accuracy of non-invasive tests excluding patients with NCPH.

Test	AUC	95% CI	P value	N with and without
HVPG ≥ 10 mmHg				
Univariate analysis				
Liver cT_1	0.8148	0.64-0.99	0.073	36 vs 3
Spleen cT_1	0.5766	0.39-0.79	0.601	37 vs 3
Bilirubin	0.8716	0.73-1.00	0.080	37 vs 2
Multivariate analysis				
Liver cT_1 + bilirubin	0.9167	0.82-1.00	**0.049**	36 vs 2
Liver cT_1 + spleen cT_1	0.8241	0.70-0.96	0.058	36 vs 3
HVPG > 12 mmHg				
Univariate analysis				
Liver cT_1	0.7160	0.53-0.90	**0.030**	26 vs 13
Spleen cT_1	0.4986	0.33-0.71	0.874	27 vs 13
Bilirubin	0.7577	0.58-0.94	**0.011**	27 vs 12
Platelets	0.6453	0.43-0.86	0.141	27 vs 13
Multivariate analysis				
Liver cT_1 + platelets	0.7426	0.57-0.91	**0.016**	26 vs 13
Liver cT_1 + bilirubin	0.6987	0.51-0.89	**0.041**	26 vs 12
Liver cT_1 + spleen cT_1	0.7189	0.53-0.90	**0.045**	26 vs 13
HVPG > 16 mmHg				
Univariate analysis				
Liver cT_1	0.7284	0.57-0.89	**0.020**	12 vs 27
Spleen cT_1	0.5923	0.39-0.80	0.360	12 vs 28
Bilirubin	0.7284	0.57-0.98	**0.024**	12 vs 27
Platelets	0.6429	0.47-0.82	0.157	12 vs 28
Multivariable analysis				
Liver cT_1 + bilirubin	0.7147	0.54-0.89	0.059	12 vs 26
Liver cT_1 + platelets	0.7222	0.55-0.89	**0.047**	12 vs 27
Liver cT_1 + spleen cT_1	0.7438	0.56-0.93	**0.044**	12 vs 27
Future episodes of variceal bleeding				
Univariate analysis				
Liver cT_1	0.5741	0.39-0.76	0.465	12 vs 27
Spleen cT_1	0.7173	0.54-0.90	**0.031**	12 vs 28
PTT	0.7325	0.56-0.90	0.027	11 vs 26
Haemoglobin (Hb)	0.7396	0.57-0.91	**0.018**	12 vs 28
Multivariable analysis				
Spleen cT_1 + Hb	0.7679	0.59-0.94	**0.012**	12 vs 28

5.4.5 Discussion

HVPG correlated with liver cT_1, but not spleen cT_1. Combinations of liver cT_1, bilirubin and platelet count performed best for discrimination of severity of CSPH but had wide 95% CIs. Spleen cT_1 performed excellently for the prediction of variceal bleeding in patients on NSBB, with no combination of variables on logistic regression improving the performance.

It is not immediately clear the reason for the poor diagnostic ability or lack of correlation with HVPG shown by spleen cT_1 in this study. The spleen has been shown to be an important organ in the pathophysiology of PH and it has been reported in another study of 30 patients with NAFLD, ArLD and autoimmune liver disease that, although spleen T_1 correlated overall with HVPG ($r = 0.40$, $p = 0.028$), the relationship did not hold in patients with HVPG above 10 mmHg[117]. In this cohort from SGH 38/46 (83%) of patients had HVPG ≥ 10 mmHg, potentially skewing our results unfavourably. Nevertheless, the lack of correlation of spleen cT_1 with HVPG was surprising given the widely held assumption that spleen measurements, in some form or other show fairly consistent relationships with increasing portal pressure.

On closer inspection, I examined the ability of the HVPG itself for the prediction of future episodes of variceal bleeding, returning an AUC of 0.48, indicating poor prognostic ability. This result casts doubt on the utility of comparing our MR measures against the HVPG measurements made during this study, as HVPG should be the gold-standard diagnostic measure of these complications of PH. However, this centre has previously published on a standardised protocol, which is also used here for HVPG measurement[308] and the senior investigator has considerable experience in this field.

This study included a group of 7 patients with non-cirrhotic portal hypertension (NCPH) who presented with clinical features of PH. (All patients underwent liver biopsy, confirming absence of histological cirrhosis). The median HVPG in this subset of patients was only 4 mmHg (IQR 3-11 mmHg). Patients with pre-hepatic NCPH often present with normal or only mildly elevated HVPG levels, regardless of severity of their disease, underestimating the true portal pressure. This is because the HVPG is an accurate measure of cirrhosis-related PH (as it is a good reflection of sinusoidal pressure), but not necessarily PH caused by pre-sinusoidal factors. HVPG does not necessarily predict clinical outcomes in patients with NCPH as it does in patients with cirrhotic PH. This was recently shown by follow up over 3.6 years in a cohort of 53 patients with NCPH[310], in which the authors questioned the utility of HVPG in these patients. As a result, there is an under-representation of 'true normal' HVPG values in this cohort. This could have therefore affected the accuracy of the HVPG measurements as mentioned in the previous paragraph. However, when excluding those with NCPH there was still no significant correlation between spleen cT_1 and HVPG, a caveat being that only 3/37 remaining patients had HVPG < 10 mmHg, significantly reducing the range of HVPG values. This is likely to have negatively impacted on the correlation between spleen cT_1, especially at the lower end; since studies have shown spleen T_1[117] and spleen TE[49] to correlate better with HVPG at the lower end of the spectrum, and worse or not at all at higher HVPG values. This is also a contrasting aspect of the patient cohort when compared to the OCMR patients, which had even groups of HVPG < 10 and ≥ 10 mmHg and contained only one patient with NCPH. The only other major difference between this study and the OCMR study was the ethnicities of the cohort, since the main aetiologies were similar (NAFLD/NASH, viral hepatitis and ArLD).

When subgroup analysis was performed excluding those with NCPH, overall results did not change significantly except for increased diagnostic performance of liver cT_1 for CSPH. This makes sense, as now the only cause of PH in these patients was related to cirrhosis/advanced fibrosis.

It was initially thought that poor quality MR scans could have been affecting spleen cT_1's correlation with HVPG, but all these scans were passed through a quality control process by PD as part of its Liver*MultiScan*™ service. The iron correction did not significantly alter spleen T_1's ability to prognosticate, (even though it did significantly alter the absolute values due to the linearity of the field strength correction), with similar statistics for all results obtained with cT_1, similarly to Section 1.

The surprising lack of correlation between spleen cT_1 and HVPG may be explained by the hypothesis that there may be an upper limit for which spleen cT_1/T_1 remains corelated to portal pressure. After a certain point, it may be that compensatory effects such as opening of of collateral vessels halts the increasing resistance to flow from the gut. This may cause the cT_1 to plateau off. This combined with the very low proportion of patients with HVPG < 10 mmHg may have significantly affected any correlation. At high spleen cT_1/T_1, it is also possible that blood cT_1/T_1 values begin to overlap and therefore it is not possible to obtain longer cT_1/T_1 values above a certain limit.

When including only the patients on NSBB, neither spleen cT_1 nor T_1 improved their diagnostic statistics except in their excellent ability to discriminate future episodes of variceal bleeding, a serious complication of PH that can lead to death. However, this was in a small cohort (n = 19) with skewed groups (4 vs 16), so must be validated further.

Liver cT_1 correlated moderately with HVPG measurements, which has been shown to be the case with liver stiffness measured by TE[47] and MRE[99] and also liver T_1[117]; liver cT_1 also performed better than spleen cT_1 and T_1 in the discrimination of severities of PH < 12 mmHg, again indicating as in section 1 that it may have a role to play in the diagnosis of CSPH, probably more so in patients with cirrhosis-related PH. Bilirubin and platelet count however performed similarly in classifying those with CSPH, are currently utilised in clinical practice, and are cheaper. Combinations of these and liver cT_1 performed best for predicting severity of PH, but the AUCs had large 95% CIs and must be taken with caution. Furthermore, the increase in diagnostic accuracy from the individual biomarker's accuracies was very small, except in the case of liver cT_1 + bilirubin prediction of CSPH (AUC 0.92) in the subgroup excluding those with NCPH.

Spleen cT_1 however was shown to be a good predictor of future episodes of variceal bleeding, both in the whole cohort (0.7362 (95%CI:0.563-0.9093)), and especially when including those patients on NSBB treatment (0.9688 (95% CI: 0.898-1). Variceal haemorrhage is an extremely serious consequence of PH and can lead to death; spleen cT_1 may be able to identify patients who would fail primary prophylaxis for variceal bleeding.

When we included those patients on NSBB, the results of ROC analysis for classification of patients with HVPG $\geq$ 10mmHg contained very large 95% confidence intervals, and so any of the results produced should be taken with caution.

Spleen cT_1 has not shown to correlate with HVPG in this study but may be due to the skewed proportions of the cohort with NCPH and very few with HVPG < 10 mmHg. However, even when excluding those patients with NCPH, the results did not improve. Spleen cT_1 has shown

utility in predicting future episodes of variceal bleeding, especially in patients on NSBB treatment, and still remains a potentially useful tool that requires further investigation.

Chapter 6. T_1 and cT_1 of abdominal organs for the assessment of portal hypertension in a paediatric population of autoimmune liver disease

The work in this Chapter has been accepted as a paper in:

Dillman J, Tkach J, Ganghi D, Singh R, *Jayaswal A*, Miethke A, Trout A. Relationship between Magnetic Resonance Imaging Spleen T_1 Relaxation and Other Radiologic and Clinical Biomarkers of Liver Fibrosis in Children and Young Adults with Autoimmune Liver Disease.

6.1 Abstract

Background & Aims: Chronic liver disease (CLD) can lead to portal hypertension (PH), which is associated with development of varices and hepatic decompensation. Non-invasive markers of the spleen and liver have been shown to be associated with PH and liver and spleen cT_1 have had varying success in diagnosing portal hypertension in adult cohorts. For T_1/cT_1 to be of clinical utility it is preferable that is can diagnose PH accurately in children/young adults also. We therefore aimed to assess spleen T_1/cT_1 and liver cT_1 for the assessment of PH in a paediatric population.

Patients & Methods: Paediatric patients with autoimmune liver diseases (AILD) and radiological assessment of PH were recruited from Cincinnati. They underwent imaging and clinical assessments. The primary aim was diagnosis of clinically significant PH (CSPH) defined by radiological evidence (including ascites, varices, splenomegaly) of PH.

Results: Spleen cT_1 was able to predict presence of CSPH with AUC 0.81, with spleen and liver cT_1 on logistic regression able to predict CSPH with AUC 0.92.

Conclusions: T_1/cT_1 mapping of the liver and spleen can diagnose portal hypertension in patients with autoimmune liver disease.

6.2 Introduction

To maximise the potential utility of splenic T_1 or cT_1 to assess PH it needs to be assessed in a paediatric setting in addition to the adult population. Changes in specifics such as hormone levels and cell growth as well as general lifestyle all occur in the transition from childhood, through adolescence and into adulthood. It has been shown that, although smaller than that of adults in absolute terms, the size of the spleen increases with age in children until adulthood[311]. There may therefore also be differences between spleen T_1/cT_1 in adults vs children, which may need to be taken into account in clinical application of spleen T_1/cT_1.

The main aetiologies of CLD are different in children than adults. In the UK, ArLD, viral hepatitis and NAFLD are the three main causes of adult CLD, followed by autoimmune liver disease (AILD) and biliary diseases. Paediatric liver disease includes biliary atresia, familial intrahepatic cholestasis, metabolic liver disease, AILD and others. AILD therefore, is amongst the numerous causes of CLD in both children/young adults and an adult setting. AILD includes primary sclerosing cholangitis (PSC), autoimmune hepatitis (AIH) and autoimmune sclerosing cholangitis (ASC). In PSC, where autoimmune injury affects the intra and extrahepatic medium sized bile ducts; this can cause multifocal bile duct stricturing and cause concentric and obliterative fibrosis. AIH on the other hand affects the hepatocytes causing interface hepatitis. ASC encompasses characteristics from both PSC and AIH. These can cause development of fibrosis, cirrhosis and subsequently PH.

Diagnosis of PH can be confirmed by radiological evidence of splenomegaly, varices or ascites. This method lacks a quantitative continuous variable like HVPG against which to assess

severity of disease, limiting the comparison of any non-invasive diagnostic tool to classification of a binary outcome, but still assesses a clinically important endpoint.

Initial published data from this cohort had shown liver cT_1 to be able to predict the presence of PH as defined by radiological assessment with an AUC of 0.86 (95% CI: 0.75-0.97)[312]. Spleen T_1 and cT_1 had not been assessed as part of this original study, but I was invited to perform the blinded image analysis for the quantification of spleen T_1/cT_1 as part of an additional collaboration between Cincinnati Children's hospital (CCH) and PD.

6.3 Aims

The aims of this section were to:

1. assess spleen T_1 and cT_1's ability to classify those with radiological evidence of PH in a population of children and young adults.
2. assess liver cT_1's ability to classify those with evidence of PH
3. assess linear combinations of liver cT_1 and spleen T_1/cT_1

6.4 Methods[d]

Patient selection

Children and young adults with AILD were enrolled in a prospective registry at the Center for Autoimmune Liver Disease at CCH. Children and young adults (ages 0-25) with a diagnosis of AILD (PSC, AIH and ASC) were approached to enrol in the registry, starting in autumn 2016. The registry allows for the collection of pertinent radiologic, histopathologic, clinical, laboratory and demographic data. Two radiologists (9 and 5 years of post-fellowship experience respectively) reviewed coronal and axial T_2-weighted anatomical images separately to document the presence of splenomegaly, varices and ascites. Reviewing radiologists knew patients had a diagnosis of AILD but were blinded to specific diagnoses as well as all clinical, laboratory and quantitative MRI data. Consensus review was used to manage any discrepancies. Splenomegaly was defined as craniocaudal splenic length that was more than two standard deviations above the mean length for participant age and gender[313]. For participants older than 17 years, the 17-year-old mean age and standard deviation was used. Varices were defined as visible portosystemic collateral vessels. Ascites were defined as any volume of free fluid in the peritoneal cavity, including adjacent to the liver or spleen. A diagnosis of radiologic PH was defined as the presence of any of the following:

(1) varices alone; (2) splenomegaly and ascites; (3) splenomegaly and varices; or (4) splenomegaly, ascites and varices.

[d] Methods section is the same as published by Dillman et al.[312] – I did not contribute to this – I only contributed to the image analysis for the spleen cT_1 values. The text is duplicated for maximum accuracy and background information – I do not claim it as my own work.

MR protocol

As part of participation in this registry, participants underwent multiparametric MRI assessment at 1.5T (Ingenia, Philips Healthcare, Best, Netherlands). For T_1 mapping, a MOLLI 5(1)1(1)1 acquisition scheme was used with the following parameters: TR/TE=4.76/2.36 ms, flip angle=35°, FOV=440 mm, matrix=192×192, section thickness=8 mm, slice gap=7 mm, half Fourier=0.75, acceleration=2, receiver bandwidth=312 Hz/pixel, and number of averages=1. A multi-echo gradient echo (GRE) sequence was used (TE = 2.37-18.96 ms) to acquire T_2* maps. All MR images were analysed by me as described in Chapter 2, blinded to the clinical data. Example images shown in Figure 33.

Statistical analysis

Boxplots comparing spleen T_1/cT_1 between patients with and without PH were generated. Comparisons were carried out using the Wilcoxon non-parametric test. Tests for correlations were performed using the Spearman's rank correlation coefficient. ROC analysis was carried out for the prediction of radiological PH and varices. Best cut-offs using the Youden index (best combination of sensitivity and specificity) were calculated for spleen T_1/cT_1.

Logistic regression of liver cT_1 and spleen T_1/cT_1 was performed for prediction of PH. No analysis comparing MRI parameters to clinical data or combining clinical data in a logistic regression model was possible since the study was ongoing and full unblinding could not take place yet.

Liver cT_1 had already been assessed in univariate analysis by Dillman et al[312], so was not repeated here. Liver stiffness measurements were assessed by MRE but data sharing of the

MRE results was not covered by the agreement between CCH and Perspectum, therefore stiffness measurements were unavailable for this Chapter.

6.5 Results

40 patients were recruited, of which all had paired MR data and radiological diagnoses of PH. Patients were of median (IQR) age 16 (13-17), 21 (53%) were male, and autoimmune hepatitis (AIH), primary sclerosing cholangitis (PSC) autoimmune sclerosing cholangitis (ASC) were present in 19 (47.5%), 7 (17.5%) and 11 (27.5%) respectively. Full data described in Table 26**.**

Table 26: Baseline characteristics (CCH)

Disease aetiology	
	N = 40
Autoimmune hepatitis, n, (%)	19 (47.5)
Primary sclerosing cholangitis, n, (%)	7 (17.5)
Autoimmune sclerosing cholangitis, n, (%)	11 (27.5)
Other*, n (%)	3 (7.5)
Demographics	
Age, median (IQR)	16.2 (13.0-17.3)
Male sex, n (%)	21 (53)
Time since diagnosis (years), median, (IQR)	2.9
Alanine aminotransferase (U/L), median (IQR)	48 (25.5-74.5)
Aspartate aminotransferase (U/L), median (IQR)	31 (19.8-54.5)
Albumin, mean $\pm$ standard deviation	3.77 $\pm$ 0.55
Alkaline phosphatase, median (IQR)	164 (89-258)
Gamma-glutamyl transferase, median (IQR)	42 (17-196)
Platelet count (x 10^9/L), median (IQR)	247 (113-299)
Presence of varices, n (%)	11 (27.5)
Presence of ascites, n (%)	11 (27.5)

Radiological evidence of PH, n (%)	12 (30)

*IgG4 = 1, choledochal cyst = 1, undetermined = 1

Spleen T_1/cT_1 in patients with and without PH

Spleen cT_1 was higher in patients with PH than in those without (1478 ms vs 1401 ms, p = 0.0014, n = 12 vs 28, Figure 34), higher in those with splenomegaly and ascites than those without (1483 ms vs 1402 ms, n = 11 vs 29, p = 0.0011), and higher in those with splenomegaly and varices than those without (1483 ms vs 1402 ms, n = 11 vs 29, p = 0.0011). Exactly the same patients had varices as had ascites. Spleen T_1 was higher in patients with PH than in those without (1205 ms vs 1124 ms, p = 0.0014, n = 12 vs 28, Figure 34), higher in those with splenomegaly and ascites than those without (1211 ms vs 1125 ms, n = 11 vs 29, p = 0.0011).

Spleen T_1/$cT_1$for classification of PH

Spleen cT_1 could predict presence of PH with AUC 0.8125 (95% CI: 0.6515-0.9735) (n = 12 with and n = 28 without PH, Figure 35). Identical results were obtained for spleen T_1 AUC: 0.8125 (95% CI: 0.6515-0.9735). The best cut-off (Youden index) was 1488 and 1216 ms respectively with identical sensitivity 58%, specificity 100%, PPV 100% and NPV 88% for both.

Spleen T_1/cT_1 for classification of varices

Spleen cT_1 could predict presence of varices with AUC 0.8245 (95% CI: 0.6531-0.9958), (n = 11 with and n = 29 without). The best cut-off was also 1488 ms with sensitivity 64%, specificity 62%, PPV 100% and NPV 90%. Identical results were obtained for spleen T_1 with cut-off 1216.

Liver and spleen cT_1 by aetiology

Liver cT_1 was significantly higher in patients with AIH as compared to those with PSC (p = 0.05, Figure 32A). There were no significant differences between the other aetiologies. There were no significant differences observed in spleen cT_1 (or T_1) between any aetiologies (Figure 32B).

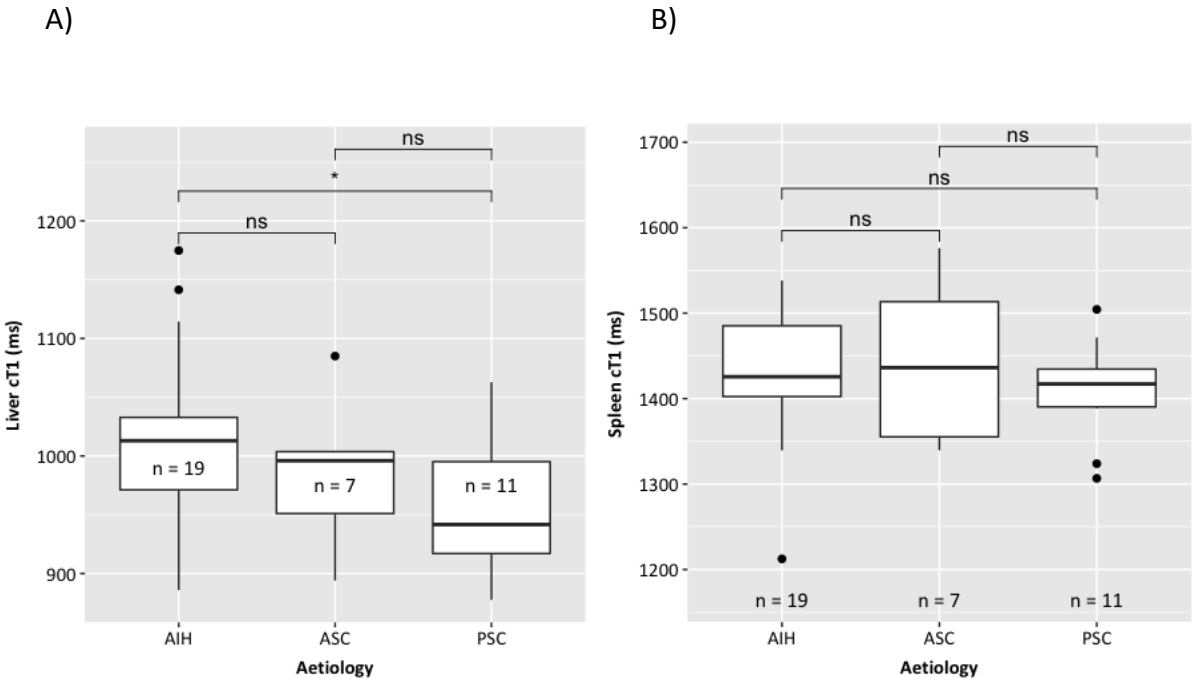


Figure 32: Boxplots showing liver and spleen cT_1 by disease aetiology. A) Liver cT_1 in AIH, ASC and PSC B) Spleen cT_1 values in AIH, ASC and PSC. * denotes $p < 0.05$

Logistic regression of liver and spleen MR

Liver and spleen cT_1 were moderately correlated (R = 0.43, p = 0.0065, data not shown).

When assessed in a logistic regression model, liver cT_1 and spleen cT_1 could predict PH with AUC 0.932 (95% CI: 0.8518-1) and liver cT_1 and spleen T_1 could predict PH with AUC 0.926 (95%CI: 0.8426 – 1, data not shown).

Spleen T_2* values

Spleen T_2* values ranged from 35.6-106.4. This is a broad range with a very high average but is unlikely to be indicating abnormally low levels of liver iron. R-squared average value was 0.905, well below the expected 0.95 level. In this particular study, since the T_2* values were so high, they would not have brought the effects of the iron correction into play in transforming T_1 to cT_1. Rather the only adjustment would be the (almost linear) standardisation for field strength – this is the reason why the diagnostic accuracy for spleen T_1 and cT_1 are almost identical.

A) B)

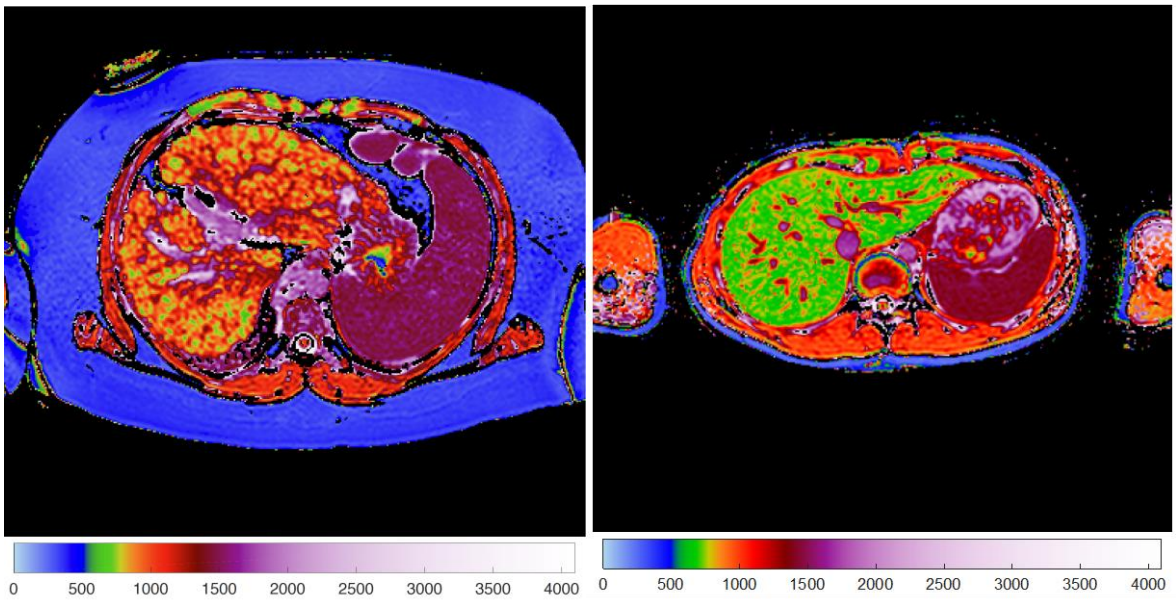


Figure 33: **Example abdominal cT_1 images of paediatric patients with PH and without PH.** A) Patient (7yrs old) has ASC with spleen cT_1 of 1566 ms, spleen T_1 1301 ms and liver cT_1 of 1003 ms B) Patients (18 yrs old) has PSC with spleen cT_1 of 1416 ms and liver cT_1 of 852 ms.

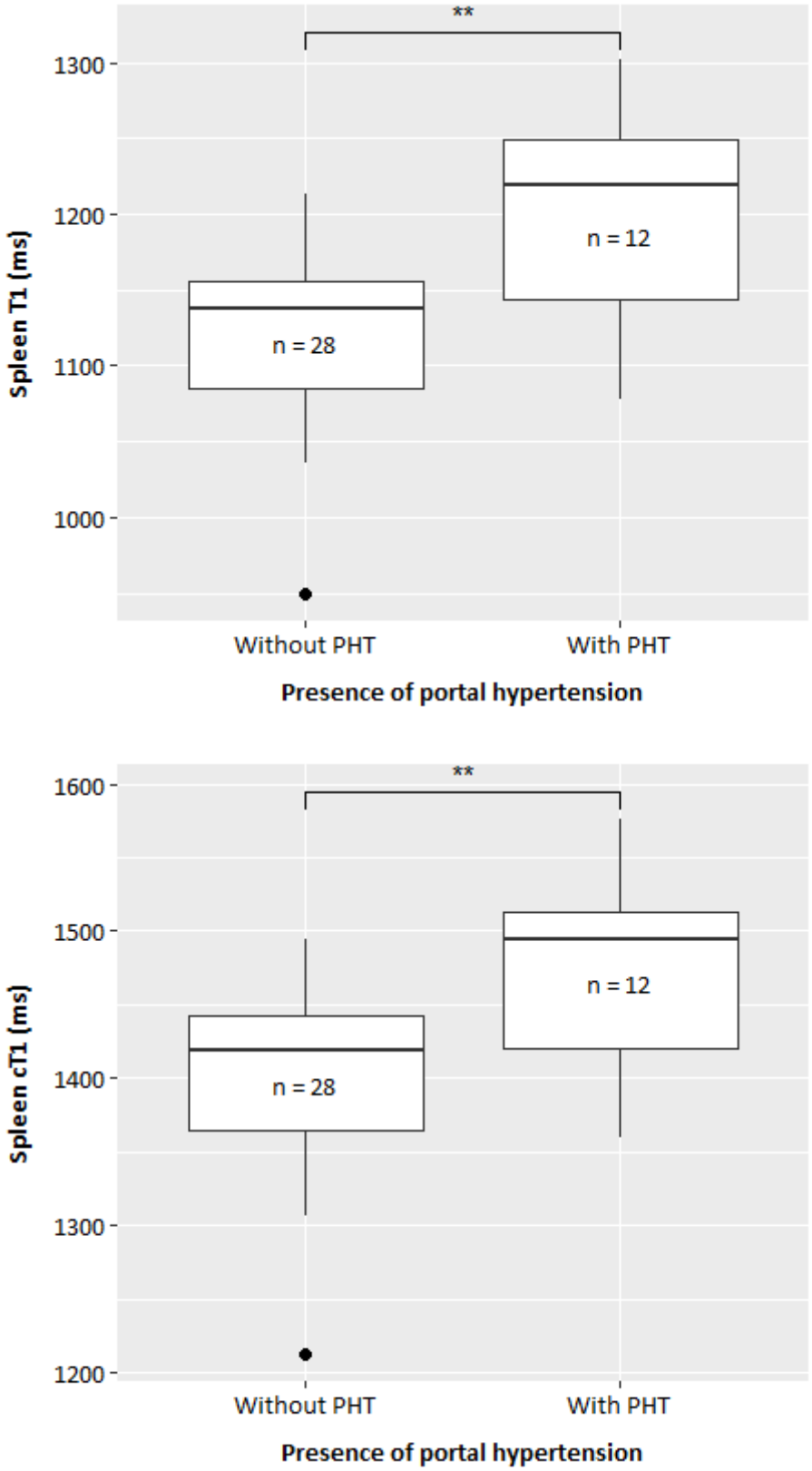


Figure 34*:* Boxplots to show differences in spleen cT_1 and T_1 in patients with and without radiological evidence of portal hypertension. This was a paediatric population with autoimmune liver disease. Radiological evidence of portal hypertension was based on presence of (1) varices alone; (2) splenomegaly and ascites; (3) splenomegaly and varices; or (4) splenomegaly, ascites and varices. ** denotes $p < 0.01$.

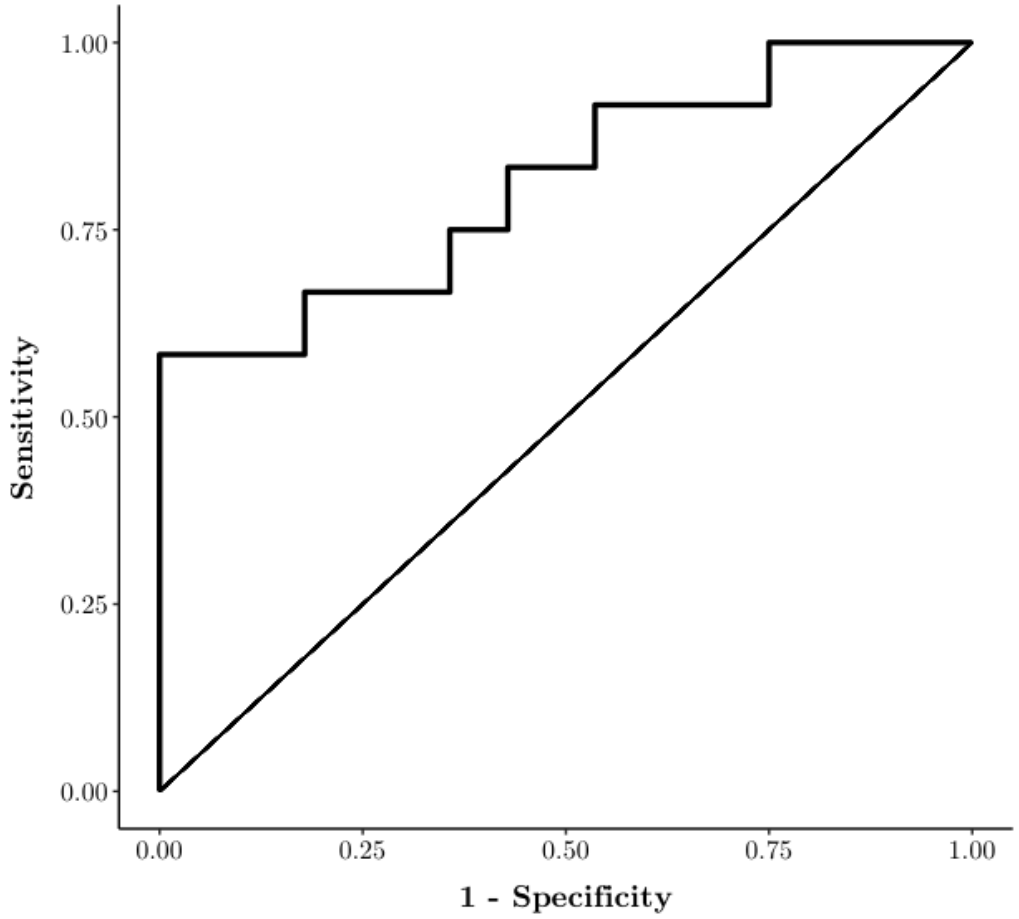


Figure 35*:* Receiver operating curve (ROC) for the assessment of portal hypertension (PH) by spleen cT_1. Population was a paediatric population of biliary disease (n = 40, 12 with PHT, 28 without PHT). AUC 0.8125 (95% CI: 0.6515-0.9735).

6.6 Discussion

Spleen cT_1 has shown good accuracy (AUC 0.81) for the prediction of PH as assessed by radiological means, and logistic regression of the two showed excellent diagnostic accuracy (AUC: 0.93). Multiple studies have assessed liver and spleen stiffness for presence of varices using MRE, TE and ARFI, but other than this work, there is a lack of literature in the assessment of paediatric PH using non-invasive methods against HVPG.

The mean time after diagnosis was 2.9 years at MRI assessment. It would be of interest to see whether imaging earlier in the timeline after disease diagnosis can yield an earlier diagnosis of PH and therefore better inform patient management.

In the same study liver and spleen stiffness by magnetic resonance elastography (MRE) performed exceptionally well for the prediction of radiological PH with AUCs of 0.98 and 0.96 respectively. This is at odds with the current literature which reports more modest performances for organ stiffness by TE, MRE and point shear wave elastography (p-SWE) for assessment of PH. There are few other studies examining abdominal non-invasive methods measuring splenic inflammation/fibrosis/congestion for assessment of PH in children/young adults, but in a paediatric cohort of 22 patients with biliary atresia having undergone Kasai operations, splenic MRE values (but not liver MRE values) were significantly different in those with and without clinical diagnosis of PH, with/without presence of varices, and could predict presence of gastroesophageal varices with AUC 0.844. However, a low APRI cut-off (0.65) was able to predict varices with the same AUC[101]. Another study investigating TE of the liver and spleen showed spleen stiffness to have an AUC of 0.87 for detection of varices in 51 children with biliary atresia, with a specificity of 93% if both liver and spleen stiffness best performing cut-offs were combined[314]. In an observational cohort study of 67 children selected for

endoscopy, spleen stiffness assessed by TE had AUC 0.92 for clinically significant varices[315]. Ozkan et al.[316] assessed p-SWE in 38 healthy children and 56 paediatric patients with biopsy-proven liver disease. All underwent splenic and liver SWE, with reliable measurements of splenic stiffness found in 76/94 (81%). This study however found liver stiffness to have a higher AUC for prediction of PH than splenic stiffness (0.906 vs 0.746, p = 0.024). Two-dimension SWE (2D-SWE) of the spleen was assessed against liver stiffness, APRI score, spleen diameter in 34 children with biliary atresia. Spleen stiffness had the highest AUC (0.900 with sensitivity 93% and specificity 90%.

It is unclear why MRE performs so accurately in the cohort we have described here, but it may be due to the different disease aetiology in our study, as most of the other studies summarised here have been in children with biliary atresia specifically. Further work with larger numbers in different centres is therefore needed to yield more robust MRE and MRI results.

Liver cT_1 also performed better than spleen cT_1 for the same classification with AUC 0.86 vs 0.81 (although the 95% CIs overlapped significantly) and was moderately correlated with spleen cT_1, indicating that the two tests may have similar diagnostic ability in paediatric AILD-related PH. However, when combined in a logistic regression model they were able to classify those with PH with AUC 0.93. This could be due to a significant contribution to PH from both interface hepatitis due to AIH manifesting as raised liver cT_1 as well as periductal fibrosis from PSC that may have more of an effect on the splenic cT_1 values than liver cT_1 values. This highlights the utility of being able to obtain both liver and spleen measurements in a single scan, unlike MRE and TE which require manual application of the driver/probe to the different organ of interest.

The range of spleen cT_1 values in patients with CSPH were noticeably higher in this paediatric population than in either of the OCMR or SGH groups (CCH: 1426 $\pm$ 72 ms, SGH 1357 $\pm$ 91 ms, OCMR: 1277 $\pm$ 95 ms). This is an interesting finding, and further studies should be done to see whether this holds across centres, countries and specific populations.

Spleen T_1 values were significantly lower than cT_1 values and correlated almost perfectly. This is likely due to the high T_2* values, which indicate low iron and therefore no iron-specific correction, as well as the contribution from the field strength correction which is linear. The lack of iron-correction algorithm being applied due to being out of the liver T_1 range is likely why there is such a large difference between the T_1 and cT_1 values here. This is also partially shown in Figure 33 and Chapter 8.

The spleen is one of the major organs of the lymphatic system and in children, the constituents (white pulp vs red pulp) are of different proportions than in adulthood as the immune system is generally more active in children. It has been shown that MR signal intensity on T_1-weighted imaging decreases and T_2-weighted increases with changing ratio of red pulp-to white pulp[317]. Another potentially contributing issue is that of disease aetiology. This cohort was entirely AILD, and the OCMR and SGH cohorts were dominated by NAFLD, ArLD and viral hepatitis. In patients with PSC and ASC, of which there were a significant number in this cohort, the disease is focused around the ducts (periductal fibrosis) unlike in diseases with more diffuse fibrosis (NAFLD, ArLD, viral hepatitis). This may have accounted for the higher spleen values compared to the OCMR and SGH studies.

Splenomegaly is also a frequent symptom of PH and used in composite scores, and accurate measures of spleen size can be obtained using MRI (as well other imaging methods). There

are still significant levels of cell development at younger ages which can sometimes cause certain organs to be relatively larger than in adulthood. Spleen size increases through childhood[318] and some older children/young adults' spleens exhibit larger size than grown adults, especially in older age since it is reported spleen size diminishes in later life[319]. A larger spleen would normally indicate risk of PH, but this variation associated with adolescence may need to be taken into account in the setting of the radiological index test.

Patients had unusually high T_2*; however, the R-squared values were also extremely low, potentially resulting in poor fitting of the T_2* relaxation curve. Regardless of the quality of the T_2* maps, this indicates no effect on the T_1 values from the iron correction on the differences in T_1 and cT_1. Rather the field strength standardisation is having the only effect, evident in the almost identical diagnostic performance of T_1 and cT_1 despite their differing absolute values.

Liver cT_1 was higher in patients with AIH than in those with PSC, but spleen cT_1 was no different across all groups. As mentioned earlier AIH is characterised by interface hepatitis and PSC is characterised by 'onion-ring' or periductal fibrosis. It is likely that localised, rather than diffuse regions of fibrosis will result in lower overall liver cT_1. The difference only just reached significance, so further numbers are needed to validate this finding.

Our results for classification of presence of varices/ascites (AUC: 0.8125) have a good, if slightly lower diagnostic accuracy here, but without the additional need for the hardware associated with MRE, and without the high technical failure rates associated with spleen (and liver) TE[320]. However, the combination of spleen and liver cT_1 performed best for diagnosis of PH with an AUC of 0.932 (95% CI: 0.8518-1), performing favourably to all the described results.

A major limitation of this study is that we have used a radiological reference standard and compared it to a new radiological index test. If one is able to obtain the diagnosis of PH through a separate scan, the utility of T_1/cT_1 may be reduced.

Chapter 7. T_1 and cT_1 of abdominal organs' summary for assessment of portal hypertension

6.1 Summary

In this and the previous Chapter, through collaboration with partners in different countries and in varying populations, by ethnicity, age and disease aetiology, we have investigated further utility of spleen cT_1 for the assessment of PH.

For the discrimination of CSPH (Table 27), spleen cT_1 performed well in the extended cohort of varying CLD aetiologies from our centre (OCMR) and within a paediatric cohort of AILD from CCH. However, spleen cT_1 was unable to discriminate for presence of CSPH in a cohort of mixed liver disease aetiologies from SGH but had excellent prognostic value in predicting future episode of variceal bleeding in patients on NSBB treatment in the same cohort.

Table 27: Summary of diagnostic accuracy of abdominal T_1/cT_1 for PH

		AUC	95% CI	P value
OCMR	Spleen cT_1	0.86	0.68-1.00	0.01
	Spleen T_1	0.89	0.77-1.00	< 0.01
	Liver cT_1	0.68	0.43-0.93	ns
SGH	Spleen cT_1	0.55	0.33-0.78	ns
	Spleen T_1	0.53	0.27-0.78	ns
	Liver cT_1	0.71	0.55-0.88	ns
CCH	Spleen cT_1	0.81	0.65-0.97	< 0.01
	Liver cT_1	0.86	0.75-0.97	< 0.01
	Spleen + Liver cT_1	0.93	0.85-1.00	< 0.01

As a surrogate marker of PH, it would be desirable for spleen T_1/cT_1 to be able to measure a response to treatment as we showed liver cT_1 could in Chapter 4. In another study that we collaborated on[116], liver and spleen T_1 showed a varied response shown to treatment by

NSBBs. Numerous trials using carvedilol either as primary or secondary prophylaxis for variceal bleeding have shown mixed results, so the variable clinical efficacy of the NSBBs could be a reason for the lack of response. Ideally to assess a response in spleen T_1, one needs a measurable response that is directly related to the HVPG, unlike heart rate response which is not[321]. A more suitable study would be to assess patients before and after undergoing transjugular intrahepatic shunts (TIPS) procedures, thereby obtaining pre- and post- TIPS portal pressure measurements and MRI. These studies should be carried out in multiple centres with uniform HVPG and MRI protocol. Studies using 4D flow MRI have done this in small cohorts, and other studies have shown decreases in spleen stiffness pre and post TIPS. For example, in a pilot study by Guo et al[100], 10 patients with cirrhosis had MRE before and after transjugular intrahepatic portosystemic shunt (TIPS) procedure, with HVPG also being measured. The magnitude of the viscoelastic modulus decreased significantly after surgery in both liver and spleen, but only the spleen modulus was correlated with relative change of HVPG.

The fact that liver cT_1 has either performed as well or better than spleen cT_1 for the assessment of PH in the studies from SGH and CCH indicated that the two may be best used in conjunction, for example liver cT_1 on its own may be accurate enough to diagnose patients with cirrhosis-related PH, but spleen cT_1 may be a useful independent predictor in NCPH (shown in one patient from OCMR dataset by Levick et al[193]) or in predicting those at risk of variceal bleeding . Liver T_1 was also shown to correlate with HVPG in 30 patients undergoing portal pressure measurements ($r = 0.835$, $p < 0.001$)[117]. There is a strong argument for combining liver and spleen measures, regardless of whether they are MR or TE-based. For example, in the detection of oesophageal varices, combining the Baveno VI criteria with spleen stiffness measurements (with a cut-off of spleen stiffness ≤ 46 kPa) resulted in a spared

endoscopy rate (SER) of 43.8% with <5% missed VNT rate[54]. However, we must account for the failure/unreliability rate of both liver TE, and the high rate of unsuccessful spleen TE measurements (15–20%)[47]. This is a recurring theme over which MRI exhibits its greater applicability. Many of the most utilised scores in hepatology are formed by utilising more than one variable, and from the work in this thesis, combinations of liver cT_1, spleen cT_1 and platelet count should be examined further.

PH is a complex pathophysiological problem with effects not only restricted to the liver-gut axis. Corrections have been made to the liver T_1 for excess iron, but we know that T_1 also rises with inflammation. In patients with systemic inflammation, liver and spleen T_1 values may be artificially elevated, a confounding factor that we cannot yet correct for/separate. A prospective study by Jansen et al.[322] showed in 83 patients undergoing TIPS procedures that despite a reduction in spleen stiffness in all patients after TIPS, 27 exhibited an increase in liver stiffness. These patients who showed increased liver stiffness also showed higher overall and/or hepatic venous levels of proinflammatory cytokines and worse survival after TIPS. This is another example showing that spleen T_1/cT_1 alone may not be sufficient for assessment of PH in all scenarios.

Studies should also investigate further confounding effects such as effect of immune response on spleen T_1 in the same way that iron and fat have been studied in the liver. If there is a way of determining/quantifying the effect of immune response on the measured T_1 that may be a factor that is necessary to correct for or take into account as well. This could have implications for patients who are ill and taking antibiotics or whose immune systems are under duress.

Studies have shown both liver and spleen stiffness to be associated with severity of PH. A study showed liver stiffness as assessed by MRE to be correlated with HVPG measurements ($r = 0.64$, $p = 0.01$) in 15 adults with advanced liver disease caused by HCV, HCV/HIV or HIV/NASH. In this study only 4/15 had HVPG $\geq$ 10 mmHg[99]. 3D multifrequency MRE is able to decompose the viscoelastic parameters of the organ under observation into different components, namely the stiffness, elasticity and viscosity. The multifrequency approach also allows the generation of the frequency power law γ, which represents the architectural structure of the tissue. This technique was assessed in 36 patients with cirrhosis being evaluated for liver transplantation and showed that the splenic loss modulus, which represents the viscosity of the spleen, was best correlated with HVPG ($r = 0.57$, $p = 0.002$) and could identify those with severe PH (HVPG $\geq$ 12 mmHg) with AUC = 0.81 and those with high-risk varices with AUC = 0.93 but had a technical success rate of only 86%[104]. Another study by Shin et al.[105] showed in 139 patients with liver cirrhosis screened for varices spleen stiffness by MRE to correlate with grade of oesophageal varices, but failed in 17/156 (11%) of patients. A similar study showed liver stiffness by MRE to have a good AUC for presence of varices AUC: 0.859 and high grade varices AUC: 0.810[106].

This summary shows that there are conflicting results relating to whether elevated spleen or liver stiffness as assessed by MRE is best associated with HVPG values and clinical symptoms of PH.

Liver and spleen stiffness assessed by TE is currently leading the way in the non-invasive assessment of PH with excellent results for both prediction of CSPH and correlation with HVPG values. In a review by Berzigotti[47], in 13 studies assessing liver TE correlations with HVPG, 4 studies showed $r > 0.8$, 6 showed r between 0.7-0.8 and 3 showed $r < 0.7$. In the

same review, 15 studies (2 in addition to the studies reporting correlation metrics) were evaluated for diagnosis of CSPH. Eight studies had AUC > 0.9, the remaining 7 all having AUC ≥ 0.8; however, all these studies used varying cut-offs for definition of CSPH. One systematic review and meta-analysis showed spleen stiffness to have summary correlation coefficient of 0.72 (95% CI: 0.63-0.8) and summary AUC of 0.84 (0.72-0.92). Spleen stiffness has also been used as an adjunct marker to liver stiffness for the detection of varices as mentioned above[54]. However, a criticism was that spleen stiffness values only reached 75 kPa, the same upper limit for liver. Recently a new spleen stiffness probe has been developed[55] and described in a paper by Stefanescu et al.[56] to increase the spared endoscopy rate when used in combination with the Baveno VI criteria to 26.5%.

Nevertheless, TE still has high rates of failure, coefficients of variation[230] and unreliability especially in NAFLD/NASH, and MRE signal is still degraded in patients with high iron, the prevalence of which is considerable[86]. There is still therefore an unmet clinical for spleen T_1/cT_1 to fulfil.

Spleen T_1 has performed similarly to cT_1 and on its own may end up being of utility, especially if the liver-specific iron correction method is found not to perform accurately on possible further validation and no alternative model is formulated (see Figure 33, Chapter 8). Standardisation of T_1 to field strength and vendor would still be a powerful tool in this case. Indeed, in a conference abstract presented by Christina Levick[301], assessing non-invasive methods of screening for varices, in 91 patients with compensated advanced CLD (cACLD), endoscopies could be avoided in 66% of patients, missing no patients with VNTs if patients had either spleen T_1 < 1037ms or spleen diameter < 137 mm. By comparison, application of the expanded Baveno criteria avoided endoscopies in 48% and missed 2 patients (5%) with

VNTs. On their own both spleen T_1 < 1037 and cT_1 < 1252 were able to avoid 45% and 43% of endoscopies missing no VNTs. However, a study by Wagner et al.[102] showed that liver stiffness but not spleen stiffness assessed by MRE, spleen T_1 or liver T_1 was correlated with HVPG. In this study, spleen stiffness was only successful in 27/33 patients (81%). In another paper by Palaniyappan et al.[117], spleen T_1 was correlated with HVPG ($r = 0.4$, $p = 0.028$) but above 10 mmHg specifically, the relationship was lost, the potential reasons for which we have briefly discussed earlier.

Our results for spleen cT_1 indicate spleen cT_1 requires a stronger evidence base to be considered as a *promising* biomarker due to the inconsistent results obtained in these three studies. However, if the positive results from these studies can be replicated, it is worth considering where in the clinical pathway abdominal multiparametric MRI might fall. As a surrogate marker of PH specifically, it could be used in the assessment of patients with cACLD as a screening test to avoid unnecessary biopsies/HVPGs and endoscopies and its diagnostic accuracy could be complemented by simple blood tests such as platelet count. It could also be used as routine follow-up in patients considered at-risk of decompensation as from our results it may be able to predict those at future risk of variceal bleeding. For this, dedicated high throughput imaging would be desirable, which could reduce associated costs.

It is likely that in order to gain the optimum results from spleen cT_1, further work is required on both validating the iron correction algorithm in the spleen and improving the T_2* variability. The former is difficult to do as there is rarely an availability of splenic histology due to reasons for splenic explant other than splenic rupture (unsuitable for obvious reasons). Technical improvement of the T_2* variability is outside the scope of discussion in this project, but the inherent susceptibility to motion of the GRE sequence used may be a limiting factor

– another option may be to investigate an entirely different iron-quantification sequence such as magnetic susceptibility weighting for example[323].

We did not perform a meta-analysis of the three studies due to the heterogeneities of the studies, (different endpoints, paediatric vs adult populations, proportion of NCPH patients), but aim to compare aspects of these studies in the next Chapter.

Chapter 8. Spleen T_1 and cT_1 in healthy volunteers and the UK Biobank imaging study

8.1 Abstract:

Background & Aims: To utilise spleen T_1/cT_1 as a surrogate biomarker of portal hypertension (PH), a working normal range of T_1/cT_1 values is required and applicability in all populations is desired. This Chapter aims to collate paired spleen T_1/cT_1 and demographic data in healthy subjects to develop a 'normal' range and investigate variables that affect spleen T_1/cT_1.

Patients and Methods: Healthy volunteers recruited from OCMR and participants from the United Kingdom Biobank (UKBB) underwent MRI for T_1 and T_2* mapping. Spleen T_1 and cT_1 were investigated for correlations with continuous demographic and clinical variables and for differences in categorical variables. The healthy volunteers from this Chapter and the patient data from Chapter 5 were collated and placed into three categories of increasing disease severity ('healthy', 'assessed for but without clinically significant PH (CSPH)' and 'with CSPH'). Differences in both spleen T_1 and cT_1 between disease severity were assessed.

Results: Spleen cT_1 had ranges of 1270 ± 93 ms and 1325 ± 61 ms in the OCMR healthy and UKBB healthy cohort respectively. Spleen cT_1 increased with disease severity category (1325 vs 1343 vs 1415 ms, $p < 0.001$). Significantly higher spleen cT_1 values were observed in women vs men (UKBB cohort: 1336 vs 1304 ms, $p < 0.0001$ and OCMR cohort: 1307 vs 1206 ms, $p < 0.0001$) and in children vs adults assessed for PH (1426 vs 1363-1277 ms, $p < 0.0001$, (CCH vs SGH-OCMR)). Spleen T_1/cT_1 correlated with liver cT_1 in the OCMR cohort ($r = 0.61$, $p < 0.0001$) and in the UKBB ($r = 0.23$, $p < 0.0001$).

Conclusions: Spleen T_1/cT_1 increases with disease severity, but differences in spleen T_1/cT_1 are comparable to differences observed between men and women and between adults and children. Further work is required to normalise/compensate for these differences.

8.2 Introduction

Having built upon previous studies highlighting the utility of spleen T_1/cT_1 relaxometry in the assessment of portal hypertension (PH)[117,193], it was shown in Chapter 5 that both spleen T_1 and cT_1 correlated well with the gold standard measure of PH (the hepatic venous pressure gradient (HVPG)) in an adult cohort recruited from Oxford, but not in an independent adult cohort recruited from Singapore. Spleen T_1/cT_1 could identify the presence of PH assessed by radiological means in a paediatric cohort (presence of splenomegaly, ascites and varices) and predict future episodes of variceal bleeding in patients who had clinically indicated portal pressure measurements. However, neither spleen T_1 nor cT_1 performed uniformly well across the studies in Chapter 5, and as such can only be described as promising surrogate biomarker of PH, similarly to magnetic resonance elastography (MRE) of the spleen which has also reported variable results in the diagnosis of PH from the literature. These are early data and refinement of the spleen T_1/cT_1 measurements may be required, as has been carried out for liver cT_1 since its inception.

In order to utilise spleen T_1/cT_1 as a surrogate biomarker of PH, it is necessary to establish a working normal range of T_1/cT_1 values, in much the same way as has been done for liver cT_1[189]. The requirement for a range of normal values is important to ensure correct cut-offs can be applied for accurate diagnostic and/or prognostic performance. Additionally, one needs to ascertain whether this biomarker is applicable in all populations, and whether there are any further variables affecting spleen T_1/cT_1 that require consideration.

There is a paucity of data relating to ranges of 'normal' spleen T_1 values in the literature. Liu et al.[324] reported ranges of 1102 ± 66 ms at 1.5T and 1352 ± 114 at 3T in 22 healthy controls in a study investigating the 'splenic switch-off' mechanism in a cardiac imaging study cohort.

In another study, similar ranges were reported of 1057 ± 42 ms at 1.5 T and 1328 ± 31 ms at 3 T in 6 healthy volunteers[325]. Contrastingly Wagner et al.[102] reported 770 ± 220 ms for patients with HVPG < 5 mmHg. There is a clear requirement for more data to define a normal range. In addition to this, there is a dependence on both the type of scanner/vendor and the field strength at which imaging is performed. This dependence can introduce unwanted variation to quantitative T_1 mapping, making it difficult to compare scan results between sites in clinical practice. It is therefore desirable to have some kind of standardisation, which cT_1 can provide.

The UK Biobank (UKBB) population health study gathers information from 500,000 volunteers to improve the prevention, diagnosis and treatment of diseases[326]. Perspectum Diagnostics Ltd (PD) provides liver imaging analysis for the Imaging Enhancement subset of the UKBB, which aims to recruit 100,000 subjects. Most of these images also provide a cross-sectional image of the spleen, which allows measurement of spleen T_1/cT_1. The population base are volunteers with no known serious illness, with an age range of 40-69 at enrolment. As a result, the UKBB population cannot be considered representative of the general population as it varies with a variety of sociodemographic, physical, lifestyle and health-related characteristics. It has been shown to have a 'healthy volunteer' bias[327], which although making it unsuitable for calculating generalisable disease prevalence statistics, provides us with a useful dataset for understanding spleen T_1/cT_1 in a healthy population. The 40-69 age group was studied because it includes people at risk of developing some of the most important diseases (cardiovascular disease, cancer, diabetes amongst others), but who are not likely to not be affected by significant confounding factors associated with old age. This however does restrict us from observing trends in younger participants.

A cohort of healthy volunteers undergoing abdominal MRI was also recruited at our centre previously by Dr Michael Pavlides that provided another dataset with a different range of ages and characteristics. Data from OCMR was collected at 3T, whereas data from the UKBB was collected at 1.5 T field strength; this allowed us to assess any differences between the two.

8.3 Aims

The specific aims of this chapter were to:

1) Establish a range of spleen T_1/cT_1 values in participants from the UKBB population health study and investigate associations between spleen T_1/cT_1 and participant characteristics.
2) Establish a range of spleen T_1/cT_1 values in healthy volunteers from our centre (OCMR) and investigate associations between spleen T_1/cT_1 and volunteer characteristics.
3) Compare established 'normal values' for spleen T_1/cT_1 from UKBB and OCMR with populations at-risk of PH from cohorts introduced in Chapter 5.

The results from these three sections are discussed as a whole at the end of the chapter rather than individually.

8.4 Spleen T_1 and cT_1 in the UKBB

8.4.1 Methods

Participant selection

Details of how participants were recruited, as well as techniques for blood sampling and handling, health questionnaires etc. can be found in the UKBB Protocol[328]. This information has not been detailed in this thesis for brevity. Briefly, participants were contacted via an invitation letter to attend an assessment centre that is convenient to them. Participants were targeted to be between 40 and 69 years of age.

MR protocol and analysis

T_1 mapping: shMOLLI 5(1)1(1)1 sequence with scan parameters: Trigger delay set to 50 ms, first TI: 170 ms, with five increments of 50 ms each, TR = 4.8, TE: 1.93 ms, slice thickness 8.0 mm, flip angle: 35° and FOV: 440 x 330 mm^2.

T_2* mapping: GRE scan with parameters: TR: 27 ms, TEs: 2.38, 4.76, 7.14, 9.52, 11.90, 14.28, 16.66, 19.04, 21.42 and 23.80 ms, 6 mm slice thickness, GRAPPA acceleration factor 2, flip angle 20° and 6°, FOV: 400 x 400 mm^2.

Automated spleen segmentation analysis was developed by researchers from PD, which made it feasible to obtain spleen T_1 and cT_1 values for a large number of participants. This technique performed a 3D segmentation of a VIBE image to identify the spleen. This segmentation mask was eroded slightly and mapped back onto the 2D MOLLI slices, from which spleen MR parameters were measured using an automated system based on the manual segmentation analysis technique described in Chapter 2.

Quality control (QC) based exclusively on liver analysis was included in the initial data sift and excluded images with artefacts and mis-triggering of the acquisition sequence. Data from participants with liver volumes under 0.5 litres had already been excluded as they reported volumes incorrectly, due to failed automated segmentations (from poor imaging data). The same applies for spleen volumes under 0.012 litres. These volume thresholds were decided by the analysts. The subsequent dataset could then be used for analysis in this Chapter.

Data acquisition from UKBB

As an investigator on Perspectum Diagnostics' application (ID: 9914, Determining the Outcomes of People with Liver Disease) to the UKBB, I was able to access and download specific 'baskets' of data. Two datasets were downloaded: one containing patient characteristics of a wide variety (details in statistical analysis) and another containing blood biochemistry and full blood count results.

Statistical analysis

The participants were investigated for the ranges of MRI parameters including spleen T_1 and cT_1. Correlations between spleen cT_1, spleen T_1, and liver cT_1 were assessed.

Exploratory analysis for associations between spleen T_1/cT_1 and both continuous and dichotomous/discrete variables was performed. Continuous variables included: aspartate aminotransferase (AST), alanine aminotransferase (ALT), albumin, alkaline phosphatase (ALP), c-reactive protein (CRP), creatinine, bilirubin, gamma-glutamyl transferase (GGT), glucose, haemoglobin (Hb), oestradiol (in women only), red blood cell count, platelet count, age, body mass index (BMI), hip circumference (HC), waist circumference (WC), height and weight. Discrete/dichotomous variables included: BMI > 30 kg/m^2, alcohol drinker status, sex,

smoking status, medication for cholesterol, blood pressure or diabetes and for women > 52 years old (average age of menopause in the UK; this was assessed based on previous observations at our centre that young women have higher spleen T_1/cT_1).

Correlations were assessed using the Spearman's rank correlation coefficient, and differences in dichotomous/discrete variables were analysed using the t-test (if normally distributed) or the Wilcoxon non-parametric test (if non-normally distributed). Normally distributed variables are presented as mean ± standard deviation, non-normally distributed variables are presented as median and interquartile range (IQR). Correlation plots and boxplots were produced for results considered to be of clinical interest. The Bonferroni correction was applied for multiple comparisons.

8.4.2 Results

After combining the imaging data, the blood results data and the participant characteristics data, information was available on 12,230 participants. Firstly, images were excluded with spleen T_1 R-square value < 0.995, as recommended by PD for accurate fitting of the T_1 relaxation curve, removing 2,341 MR scans, leaving 9,889 valid scans. Scans with liver cT_1 > 825 ms were excluded, (as this corresponds to both significant fibroinflammation and would not therefore count as a 'healthy population'), removing 212 of those participants, leaving 9,677 participants. Scans with high liver fat (indicating fatty liver; > 5%, n = 1,837) and high spleen fat readings (> 5%, n = 216) as there is very little fat in the spleen, were removed, leaving 7,624 valid MR scans. Participants with unhealthy BMI defined as BMI > 25 kg/m^2, (n = 3,074) were then excluded, leaving us with 4,550 valid scans. (At this point a subset of those participants with and without obesity (BMI > 30 kg/m^2) was created as well). Participants with

abnormal spleen T_2* levels (< 0.5 ms, n = 560) were then excluded, leaving 3,990 MRI scans with paired UKBB data (Figure 36).

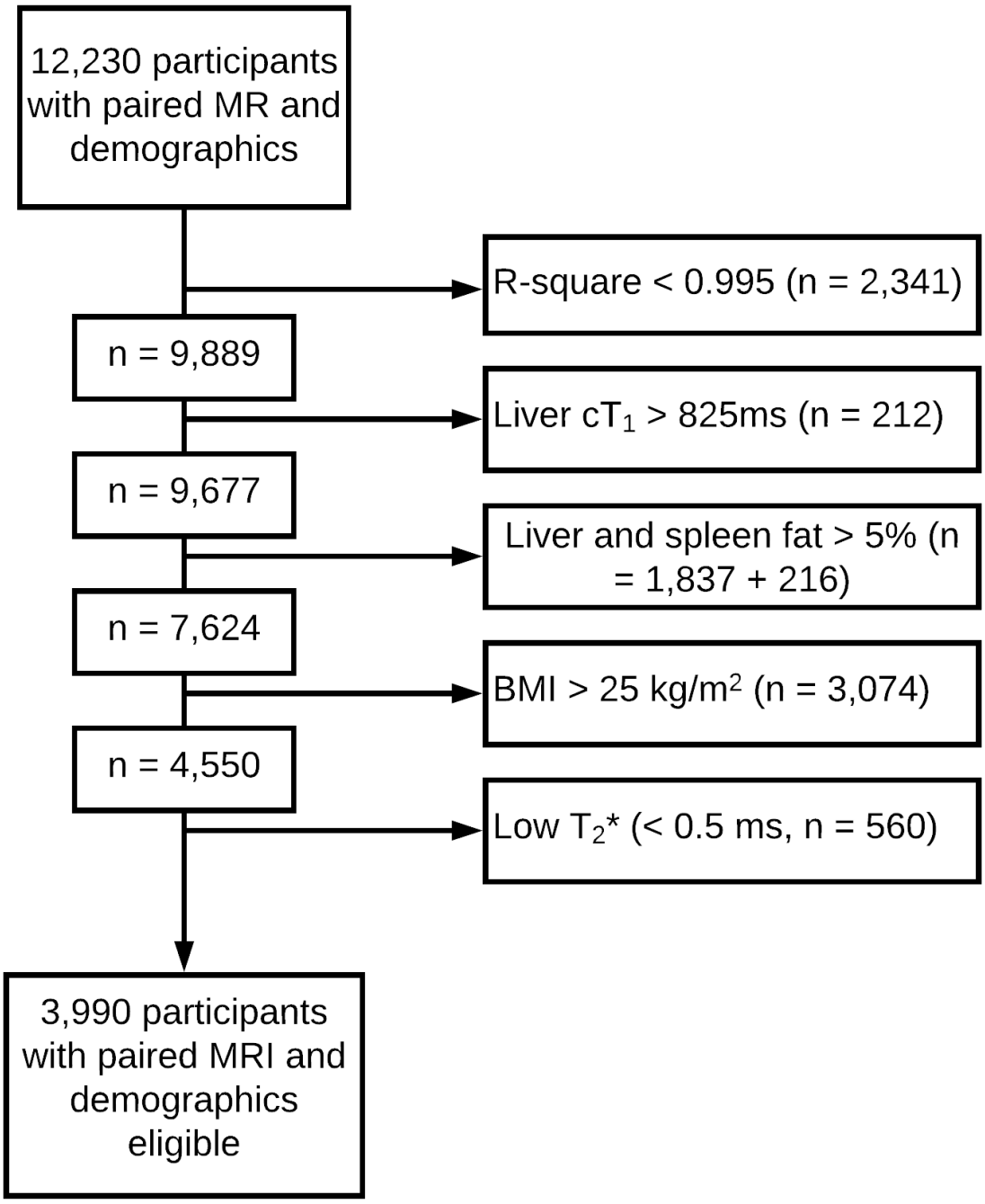


Figure 36: Flow diagram describing selection of UKBB scans suitable for analysis.

Spleen T_1 had mean value 1037 ± 59 ms and spleen cT_1 had mean value 1342 ± 62 ms and median (IQR) value of 1337 (1288-1364 ms). Spleen T_2* had median value 38 ms (25.5-51.8 ms). The distributions of spleen T_1 and spleen cT_1 were markedly different as shown in (Figure 37). This is perhaps explained by the 'correction' mapping from T_1 to cT_1 depicted in Figure 38. It appears that as well as the iron correction functioning within a certain T_2* range, it also only functions up until a maximum T_1 threshold (the upper limit of expected liver T_1), after which only the linear standardisation for field strength mapping (from 1.5T to 3T) functions.

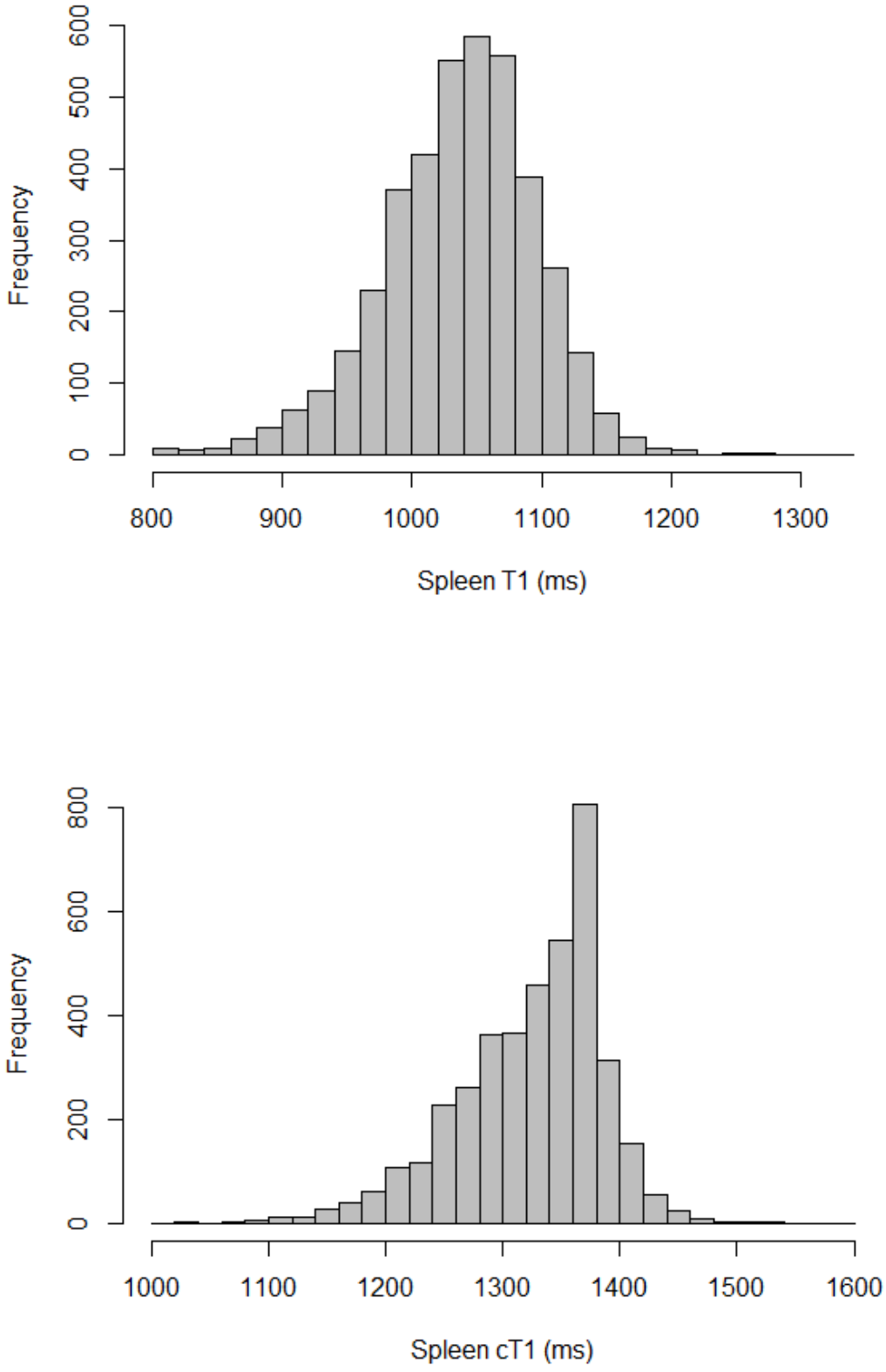


Figure 37: Histograms of spleen T_1 and cT_1 distributions in the UK Biobank.

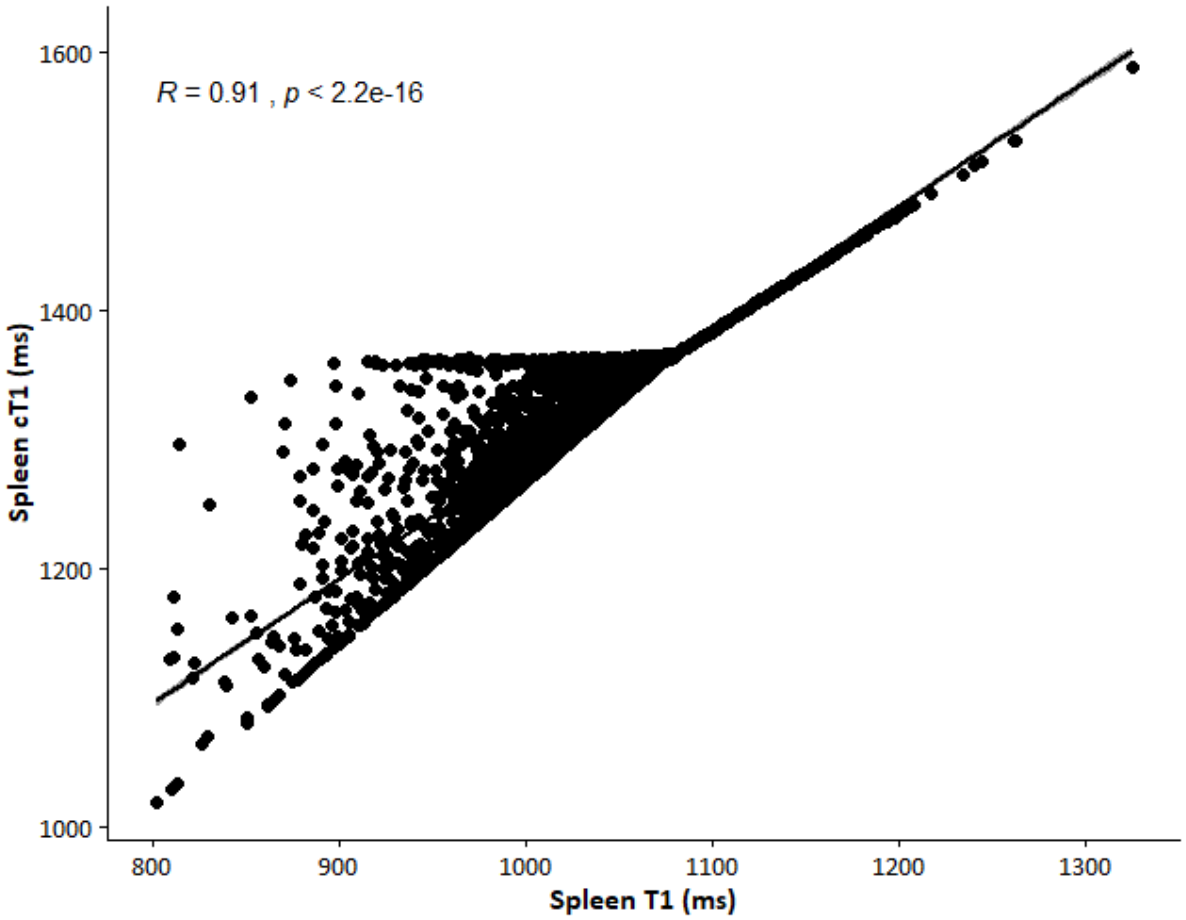


Figure 38: Correlation between spleen T_1 and spleen cT_1 in the UKBB. Correlation coefficient R calculated using the Spearman's rank correlation coefficient.

Spleen MR parameters – correlation

Spleen T_1 correlated very weakly with liver cT_1 (r = 0.202, p = 2.9 x 10^{-16}, n = 3,990). Spleen cT_1 correlated very weakly with liver cT_1 (r = 0.23, p = 2.2 x 10^{-16}, n = 3,990). Spleen T_1 correlated very strongly with spleen cT_1 (r = 0.91, p < 2.2 x 10^{-16}, n = 3,990, Figure 38). Spleen T_2* correlated weakly with liver T_2* (r = 0.14, p < 2.2 x 10^{-16}, n = 3,990).

There was no significant correlation between spleen T_1/cT_1 and any of the following: ALP, blood glucose, white cell count or oestradiol levels (in women) (Table 28) and hip circumference (Table 29). There was no significant correlation between spleen cT_1 and CRP.

Table 28: Correlations between spleen T_1/cT_1 and simple blood tests from UKBB

	Spleen T_1		Spleen cT_1		Number of data points
	Correlation	P value	Correlation	P value	
Albumin (g/L)	-0.075	$1.4x10^{-5}$	-0.065	0.00014	3369
Alkaline phosphatase (IU/L)	-0.016	0.3394	-0.016	0.3409	3699
ALT (IU/L)	-0.159	$< 2.2x10^{-16}$	-0.142	$< 2.2x10^{-16}$	3699
AST (IU/L)	-0.106	$< 9.4x10^{-11}$	-0.088	$< 7.7x10^{-16}$	3684
Bilirubin (µmol)	-0.158	$< 2.2x10^{-16}$	-0.138	$< 2.2x10^{-16}$	3684
Creatinine (g/L)	-0.225	$< 2.2x10^{-16}$	-0.197	$< 2.2x10^{-16}$	3698
CRP (mg/L)	-0.059	0.0004	-0.039	0.017	3690
GGT (IU/L)	-0.200	$< 2.2x10^{-16}$	-0.160	$< 2.2x10^{-16}$	3699
Glucose (mmol/L)	-0.016	0.347	0.0068	0.695	3368
Haemoglobin (g/L)	-0.343	$< 2.2x10^{-16}$	-0.310	$< 2.2x10^{-16}$	3785
Oestradiol* (pmol/L)	0.070	0.047	0.044	0.210	803
Red blood cell count ($x10^{12}$/L)	-0.246	$< 2.2x10^{-16}$	-0.248	$< 2.2x10^{-16}$	3785
Platelet count ($x10^9$/L)	0.103	$< 2.3x10^{-10}$	0.093	$8.4x10^{-9}$	3785
White cell count ($x10^9$/L)	-0.003	0.846	-0.004	0.802	3785

Abbreviations: ALT – alanine aminotransferase, AST – aspartate aminotransferase, CRP – C-reactive protein, GGT – gamma-glutamyl transferase. *Oestradiol sampled in women only. Statistical significance was set at 0.0038 (0.05 divided by 13) to correct for multiple comparisons.

Table 29: Correlations between spleen T_1/cT_1 and continuous variables from UKBB

	Spleen T_1		Spleen cT_1		Number of data points
	Correlation	P value	Correlation	P value	
Age (years)	-0.084	$9.3x10^{-8}$	-0.051	0.0012	3990
BMI (kg/m^2)	-0.114	$< 6.0x10^{-13}$	-0.099	$< 3.5x10^{-10}$	3985
Height (cm)	-0.194	$< 2.2x10^{-16}$	-0.179	$< 2.2x10^{-16}$	3985

Weight (cm)	-0.208	$< 2.2x10^{-16}$	-0.189	$< 2.2x10^{-16}$	3985
Waist circumference (cm)	-0.233	$< 2.2x10^{-16}$	-0.202	$< 2.2x10^{-16}$	3986
Hip circumference (cm)	-0.035	0.026	-0.035	0.028	3986
Spleen volume (litres)	-0.135	$< 2.2x10^{-16}$	-0.116	1.8×10^{-13}	3990
Liver volume (litres)	-0.090	1.35×10^{-8}	-0.096	1.2×10^{-9}	3990

Abbreviations: BMI – body mass index. Statistical significance was set at 0.0083 (0.05 divided by 6) to correct for multiple comparisons.

Weak ($r \leq |0.3|$) correlations were observed between spleen T_1/cT_1 and all remaining variables i.e. albumin, ALT, AST, bilirubin, creatinine, CRP, GGT, oestradiol, red blood cell count, platelet count, white cell count, waist circumference, age, BMI, height, weight, WC, HC, spleen and liver volume (details are provided in Table 28 and Table 29).

It is worth noting that is easy to obtain very low p values with numbers of participants as large as we have in the UKBB and as such, whilst these correlations are statistically significant, given the low correlation coefficients, they are probably not clinically meaningful. This also applies to analysis of discrete/dichotomous variables i.e. there may be significant differences between category even if the effect size is actually very small. These differences therefore require putting in context relative to differences caused/associated with other variables.

Spleen MR parameters – association with dichotomous/discrete variables

Women had significantly higher spleen T_1, cT_1, and T_2* than men (1050 vs 1015 ms, $p < 0.0001$, 1336 vs 1304 ms, $p < 0.0001$ and 40.4 vs 36.6 ms respectively, $p < 0.0001$) as shown in Figure 39. Men had larger spleens by volume than women (0.16 vs 0.10 litres $p < 0.0001$).

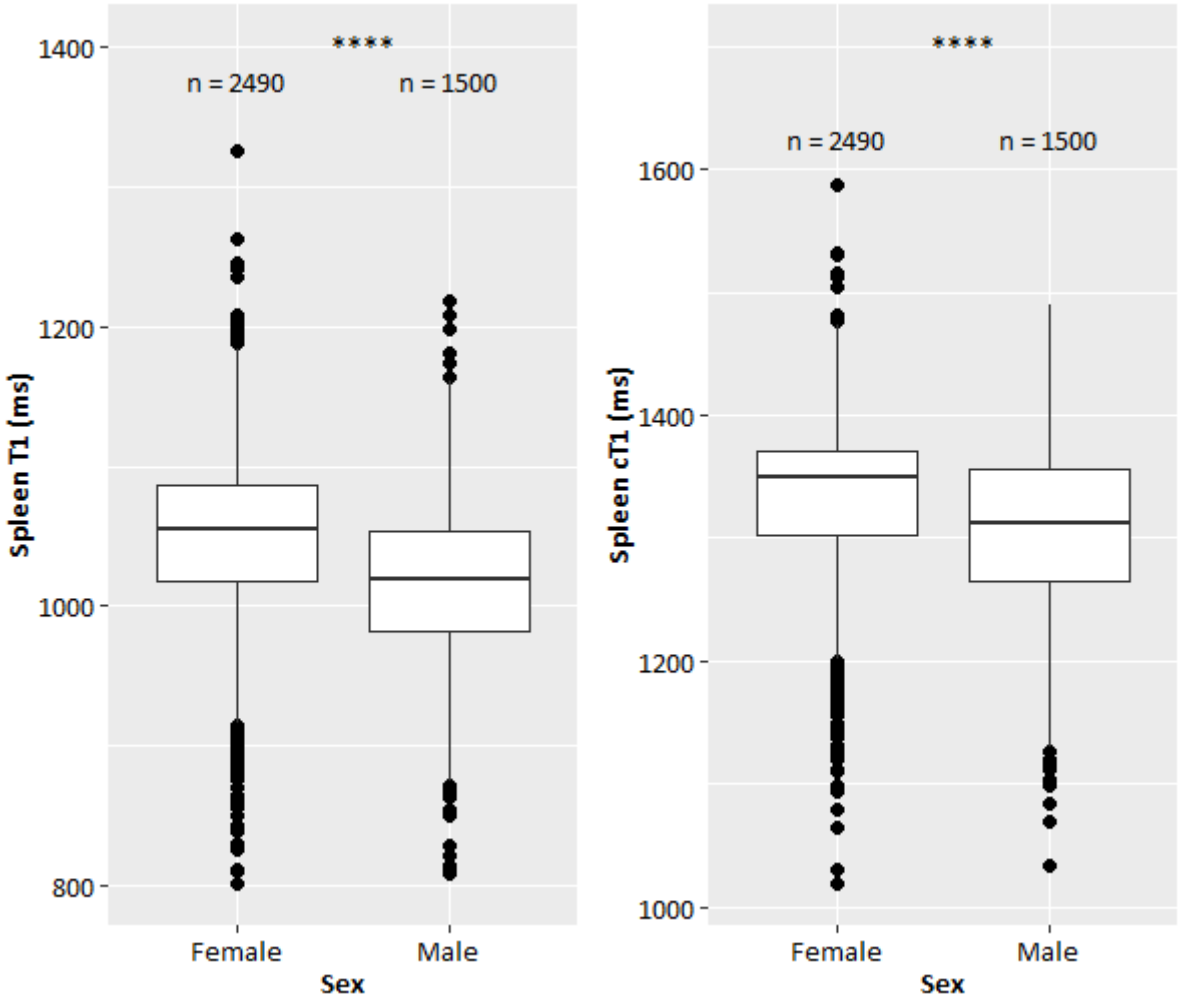


Figure 39: Difference in spleen T_1/cT_1 in UKBB stratified by sex. A) Boxplot showing the difference in spleen T_1 distribution stratified by sex. B) Boxplot showing the difference in spleen cT_1 distribution stratified by sex. **** denotes $p < 0.0001$.

There were no differences in spleen T_1, cT_1 or T_2* across participants' alcohol drinker status (categories being: prefer not to answer (T_1 = 1086, cT_1 = 1370, T_2* = 40.7, n = 2), never (T_1 = 1042, cT_1 = 1322, T_2* = 36.3, n = 107), previous drinker (T_1 = 1035, cT_1 = 1322, T_2* = 40.6, n = 75), current drinker (T_1 = 1037, cT_1 = 1324, T_2* = 38.1, n = 3806).

Spleen T_1 and cT_1 but not T_2* were significantly lower in the participants with liver $cT_1 > 825$ ms compared to those with liver $cT_1 < 825$ ms (T_1: 1025 vs 1042 ms, cT_1: 1322 vs 1337 ms, T_2*: 37.0 vs 38.1 ms, p = 0.0002, 0.0005, and 0.4442 respectively, n = 212 vs 3990).

Participants with BMI $\geq$ 30 kg/m^2 had lower spleen T_1, spleen cT_1 and higher T_2* than participants with BMI < 30 kg/m^2 (1010 vs 1028 ms, 1305 vs 1327 ms and 42.9 vs 38.1 ms respectively, n = 487 vs 7137, $p < 1 \times 10^{-9}$ for all), and had higher liver cT_1 (694 vs 674 ms, $p < 2.2 \times 10^{-16}$).

There were no discernible differences in spleen T_1, cT_1 or T_2* across participants' smoker status (categories being: prefer not to answer (T_1 = 1006, cT_1 = 1312, T_2* = 44.7, n = 3), never (T_1 = 1039, cT_1 = 1324, T_2* = 39.5, n = 2631), previous smoker (T_1 = 1033, cT_1 = 1321, T_2* = 36.9, n = 1157), current smoker (T_1 = 1040, cT_1 = 1328, T_2* = 37.9, n = 199).

There were no discernible differences in spleen T_1, cT_1 or T_2* across medication type (categories being: none (T_1 = 1016, cT_1 = 1305, T_2* = 36.1, n = 1207), do not know (T_1 = 998, cT_1 = 1281, T_2* = 26.9, n = 10), blood pressure medication (T_1 = 1015, cT_1 = 1302, T_2* = 38.9, n = 73), cholesterol-lowering medication (T_1 = 1011, cT_1 = 1299, T_2* = 35.0, n = 178), insulin (T_1 = 1033, cT_1 = 1324, T_2* = 33.2, n = 3). Medication data was only provided for 1471 participants.

There were significant differences in spleen T_1, cT_1 and T_2* in women who were older or younger than 52 years of age; (T_1: 1046 vs 1055 ms, p = 0.0001, cT_1: 1333 vs 1339 ms, p = 0.009, T_2*: 39.1 vs 42.3 ms, p = 0.0001).

After excluding those on cholesterol-lowering medication, insulin and blood pressure medication the final range of 'healthy normal' was observed in 3,736 participants, whose spleen metrics are described in Table 30 and whose participant characteristics are described in Table 31. For the rest of the chapter, this is defined as the 'healthy UKBB' range.

Table 30: Spleen MR-derived metrics for UKBB healthy volunteer subgroup

	Mean or median
Spleen T_1 (ms)	1039 ± 58
Spleen cT_1 (ms)	1338 (1290-1364)
Spleen T_2* (ms)	38.2 (25.5-52.0)
Spleen volume (litres)	0.13 ± 0.06

Table 31: Participant characteristics of UKBB healthy volunteer subgroup

Demographic data	
Age (median (IQR))	55 (48-61)
Sex (male, %)	1372 (36.7%)
Body mass index (kg/m^2)	23.6 ± 1.9
Blood biochemistry and haematology	
Bilirubin (µmol)	9.5 (6.6-10.7)
Alanine aminotransferase (ALT) (IU/L)	17 (13.7-22.0)
Aspartate aminotransferase (AST) (IU/L)	23 (20-27)
Alkaline phosphatase (ALP) (IU/L)	74 (62.3-88.4)
Gamma-glutamyl transferase (GGT) (IU/L)	20.3 (15-28)
Albumin (g/l)	45.5 ± 2.5
White cell count (x10^9/l)	6.1 (5.2-7.1)
Platelets count (x10^9/l)	243 (212-281)

8.5 Spleen T_1 and cT_1 in healthy volunteers in OCMR

8.5.1 Methods

Patient selection

Participants were recruited through posters displayed at the John Radcliffe and Churchill hospitals in Oxford. The inclusion criteria were: ages 16-80 years old and BMI < 25 kg/m^2. The exclusion criteria were: history of liver disease, excess alcohol consumption (14 units a week for women, 21 units a week for men), presence of metabolic syndrome (any one of type 2 diabetes mellitus, hyperlipidaemia, hypertension) and contraindication to MRI.

MRI protocol

Transverse abdominal shMOLLI images were acquired with parameters: 5(1)1(1)1 scheme, TI: 110 ms, 110 ms + RR, 110 ms + 2RR, 110 ms + 3RR, 110 ms + 4RR, 190 ms and 270 ms, TR/TE: 234.97/1.05 ms, TR of individual bSSFP readouts 2.5 ms, slice thickness 8.0 mm, GRAPPA acceleration factor 2, flip angle: 35°. FOV: patient dependent, matrix size: 144 x 192.

T_2* maps were acquired with a multi-echo GRE acquisition was used with parameters: TR: 500 ms, TEs: 1.25, 2.46, 3.69, 4.92, 6.15, 7.38, 8.61, 9.84 ms, 6 mm slice thickness, GRAPPA acceleration factor 2, flip angle: 20°. DIXON method used for fat fraction analysis. FOV: patient dependent, matrix size 104 x 128.

Image analysis was performed using Liver*MultiScan*^{TM} as described in Chapter 2, with the organ of interest being the spleen.

Anthropometric measures and blood sampling

Height and weight and waist and hip circumference were measured.

Blood samples were drawn and analysed for bilirubin, ALT, AST, ALP, GGT, albumin, haemoglobin (Hb) white cell count (WCC), platelet count. Samples from female volunteers, after having been spun and stored at -80°C were sent for oestradiol and progesterone hormone testing.

Statistical analysis

Spleen T_1 and cT_1 ranges were described using the mean ± standard deviation. Spleen T_1 and cT_1 were tested for correlation with simple blood tests, BMI and differences between sex were investigated. Correlation tests were performed using the Spearman's rank correlation coefficient and differences between groups using the Wilcoxon non-parametric test if non-normally distributed or the t-test if normally distributed.

8.5.2 Results

52 healthy volunteers were recruited into the study[4] and all had successful MRI scans. Participants were of median (IQR) age 39 (28-58) and had mean ± standard deviation BMI of 23.0 ± 2.7. Further baseline characteristics are presented in Table 32.

Table 32: Characteristics of healthy volunteers (OCMR)

Demographic data	
Age	39 (28-58)
Sex (male, %)	19 (36.5%)
Body mass index (kg/m^2)	23.0 ± 2.7
Blood biochemistry and haematology	
Bilirubin (µmol)	11 (9-15)

[4] These patients were recruited into the study by Dr Michael Pavlides and 42 of these have had their liver cT_1 values and baseline characteristics described in his DPhil thesis, but spleen T_1 and cT_1 from this group has not been reported there or anywhere else to the best of my knowledge.

Alanine aminotransferase (ALT) (IU/L)	16 (12.5-22.5)
Aspartate aminotransferase (AST) (IU/L)	24 (20-28)
Alkaline phosphatase (ALP) (IU/L)	134 (108-156)
Gamma-glutamyl transferase (GGT) (IU/L)	13 (11-20)
Albumin (g/l)	44.8 ± 2.3
Haemoglobin (Hb) (g/dl)	13.8 ± 1.2
White cell count ($x10^9$/l)	5.4 (4.7-6.5)
Platelet count ($x10^9$/l)	239 ± 47

Median and interquartile range (in brackets) reported for non-normally distributed variables, mean ± standard deviation for normally distributed data.

Spleen MR parameters – range

Spleen T_1 had mean ± standard deviation 1257 ± 96 and spleen cT_1 had mean ± standard deviation 1270 ± 93 and spleen T_2* had median (IQR) of 23.0 (18.9-37.8 ms).

Spleen cT_1 was significantly higher in women than in men (1307 vs 1206 ms, p = 0.0004), with the same observation seen in spleen T_1 (1298 vs 1187 ms, p = 0.0002).

Spleen MR parameters – correlation

Spleen cT_1 correlated significantly with liver cT_1 (r = 0.61, p < 0.0001) and Hb (r = -0.48, p = 0.0005) but not with any of spleen T_2*, ALP, ALT, albumin, AST, GGT or platelet count. Spleen cT_1 showed moderate negative correlations with bilirubin and BMI (r = -0.37, p = 0.008 and r = -0.33, p = 0.015 respectively). Spleen T_2* was weakly correlated with liver T_2* (r = 0.28, p = 0.043). After correcting for multiple comparisons (Bonferroni) only correlations with liver cT_1 and haemoglobin were significant.

In women, spleen T_1/cT_1 neither correlated with levels of oestradiol ($p = 0.584$) nor progesterone ($p = 0.513$).

8.6 Comparison of healthy groups with patients at risk of PH

8.6.1 Aims

The aim of this section was to investigate whether the healthy ranges of spleen T_1/cT_1 found from the participants and healthy volunteers in this Chapter are significantly different from patient ranges of spleen T_1/cT_1 described in Chapter 5. More specifically, the aims were to compare:

1) Spleen T_1 between:
    a. All groups scanned at 1.5T
        i. Broken down into 3 categories, healthy, with diagnosis of CSPH, and assessed for CSPH, but with no diagnosis of CSPH.
    b. All groups scanned at 3T
        i. Broken down into 3 categories, healthy, with diagnosis of CSPH, and assessed for CSPH, but with no diagnosis of CSPH.
2) Spleen cT_1 between:
    a. All groups
        i. Broken down into 3 categories, healthy, with diagnosis of CSPH, and assessed for CSPH, but with no diagnosis of CSPH.

8.6.2 Methods

Spleen T_1 and cT_1 values obtained from all studies across Chapters 5 and 6 were collated.

Participants were divided into three 'categories': healthy (healthy volunteers from OCMR and UKBB participants), with CSPH (defined by either HVPG $\geq$ 10 mmHg (OCMR), variceal bleeding (SGH) or radiological evidence of PH (CCH)) and those assessed for, but without CSPH (all patients assessed for evidence of CSPH, but who either had HVPG < 10 mmHg, no variceal bleeding or no radiological evidence of PH). In the SGH cohort, it was decided to define CSPH and 'assessed for, but without CSPH' by whether they had future episodes of variceal bleeding rather than the HVPG $\geq$ 10 mmHg, given that the HVPG was not predictive of bleeding.

Differences in spleen T_1/cT_1 were assessed separately at 1.5T (cohort from SGH, cohort from CCH and the UKBB) and at 3T (patient cohort from OCMR (OCMR HVPG) and healthy volunteers at the OCMR (OCMR healthy)). Spleen cT_1 was also assessed for differences by category across all the studies as the standardisation corrects for the effect of magnetic field strength.

Boxplots were generated to show differences between groups. Test for differences were performed using the Wilcoxon non-parametric test where data was non-normally distributed and the t-test where data was normally distributed. P values were adjusted for multiple comparisons using the Bonferroni correction.

8.6.3 Results

At 3T (all data from same centre; OCMR) there were significant differences in spleen T_1/cT_1 between those with CSPH and those "assessed for, but without CSPH" as described in Chapter 5. However, there was no significant difference between healthy volunteers and those at risk, but with no CSPH. In fact there was a trend towards significance for higher spleen T_1/cT_1 in the healthy volunteer group (p = 0.02) compared to those at risk, but without CSPH (Figure 40).

Differences in spleen T_1/cT_1 by disease severity were also assessed separately by sex, with boxplots (Figure 41) indicating that the higher spleen values in women significantly affect discrimination of healthy vs any stage of disease severity.

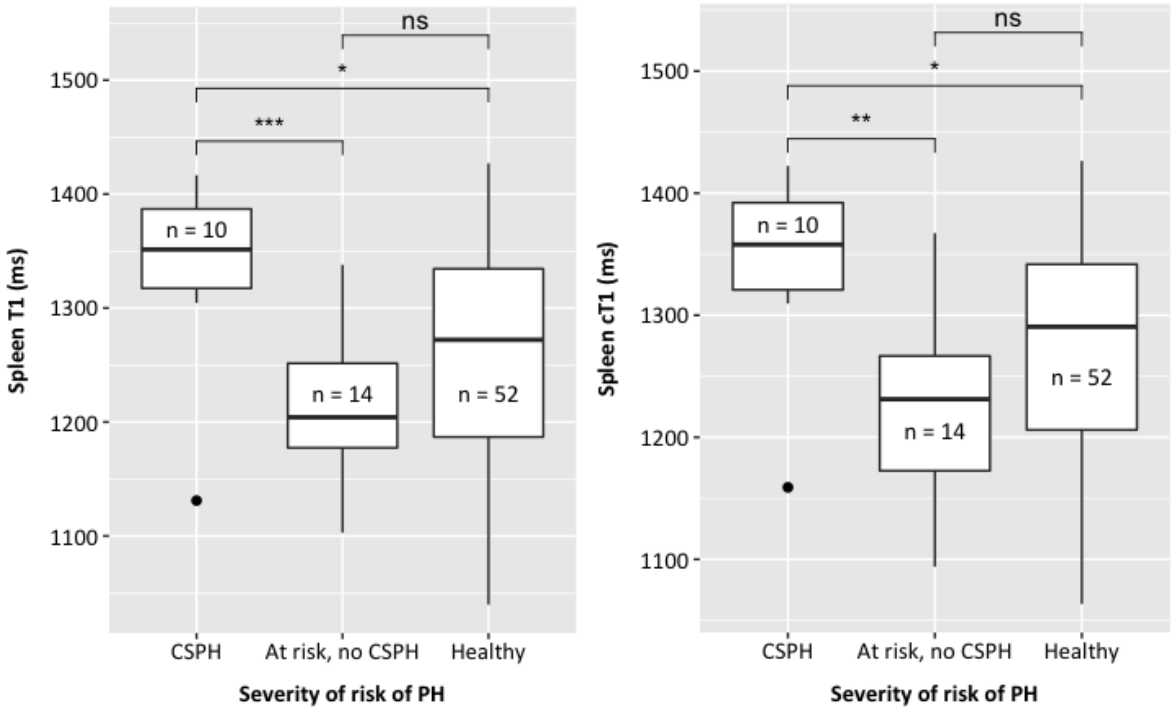


Figure 40: Boxplots showing differences in spleen T_1/cT_1 by disease severity at 3T. ** denotes p < 0.01, * denotes p < 0.05.

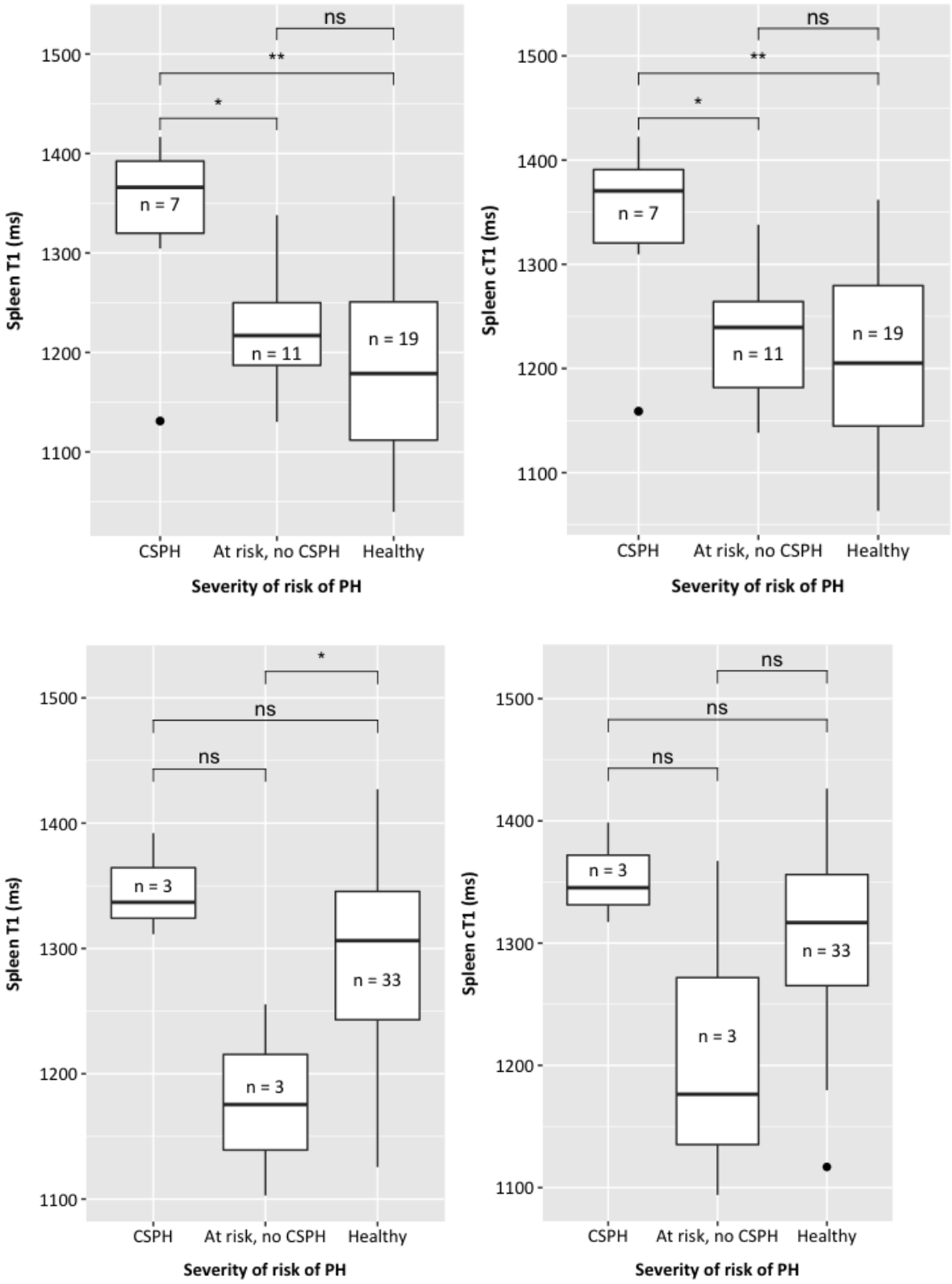


Figure 41: Boxplots showing differences in spleen T_1/cT_1 by disease severity at 3T by sex. Top row: Male, bottom row: female. Black dots represent outlying datapoints.

At 1.5T (SGH, CCH, UKBB) there were significant differences between healthy volunteers and those with CSPH and those with liver disease, but without CSPH (Figure 42).

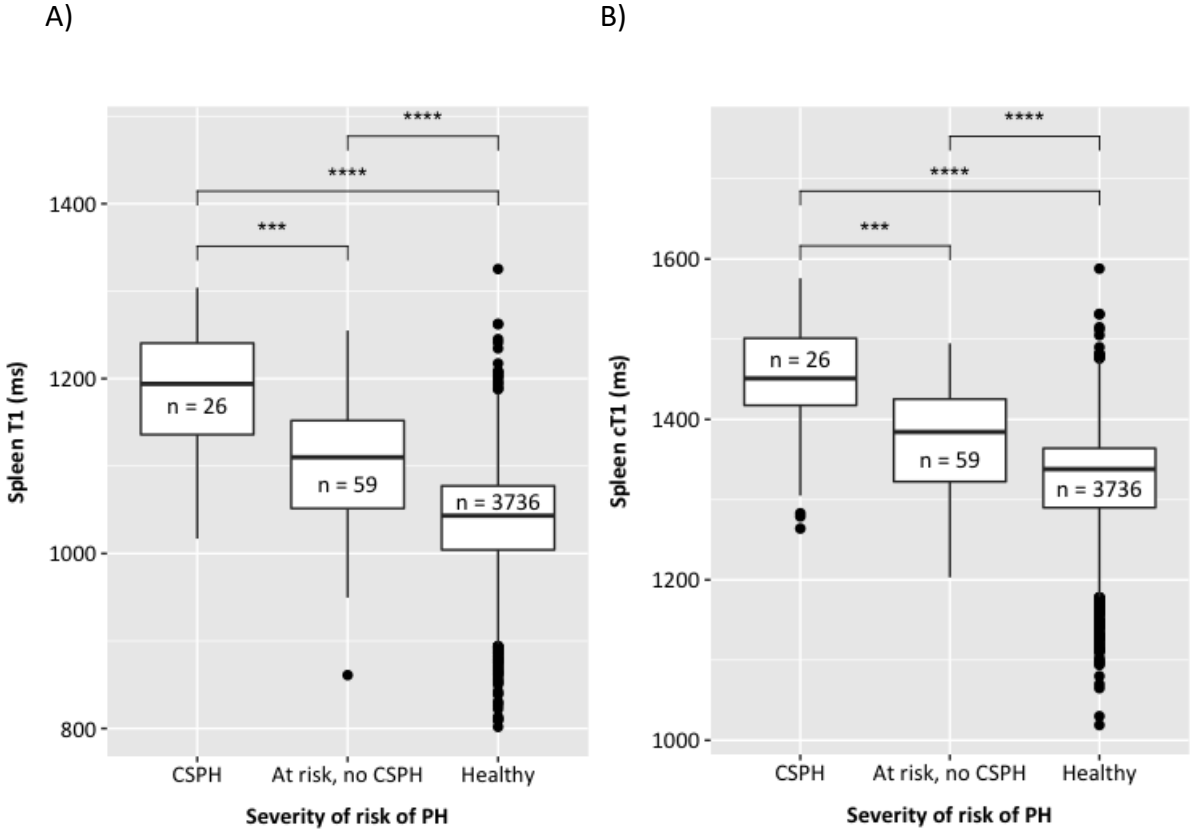


Figure 42: Boxplots showing differences in spleen T_1/cT_1 by disease severity at 1.5T. A) shows spleen T_1 and B) shows spleen cT_1 values. Black dots represent outlying datapoints.

Differences in spleen T_1/cT_1 by disease severity were also assessed separately by sex, with boxplots (Figure 43) indicating that when assessed separately, the discrimination of disease severity is similar between groups as in the whole set of participants.

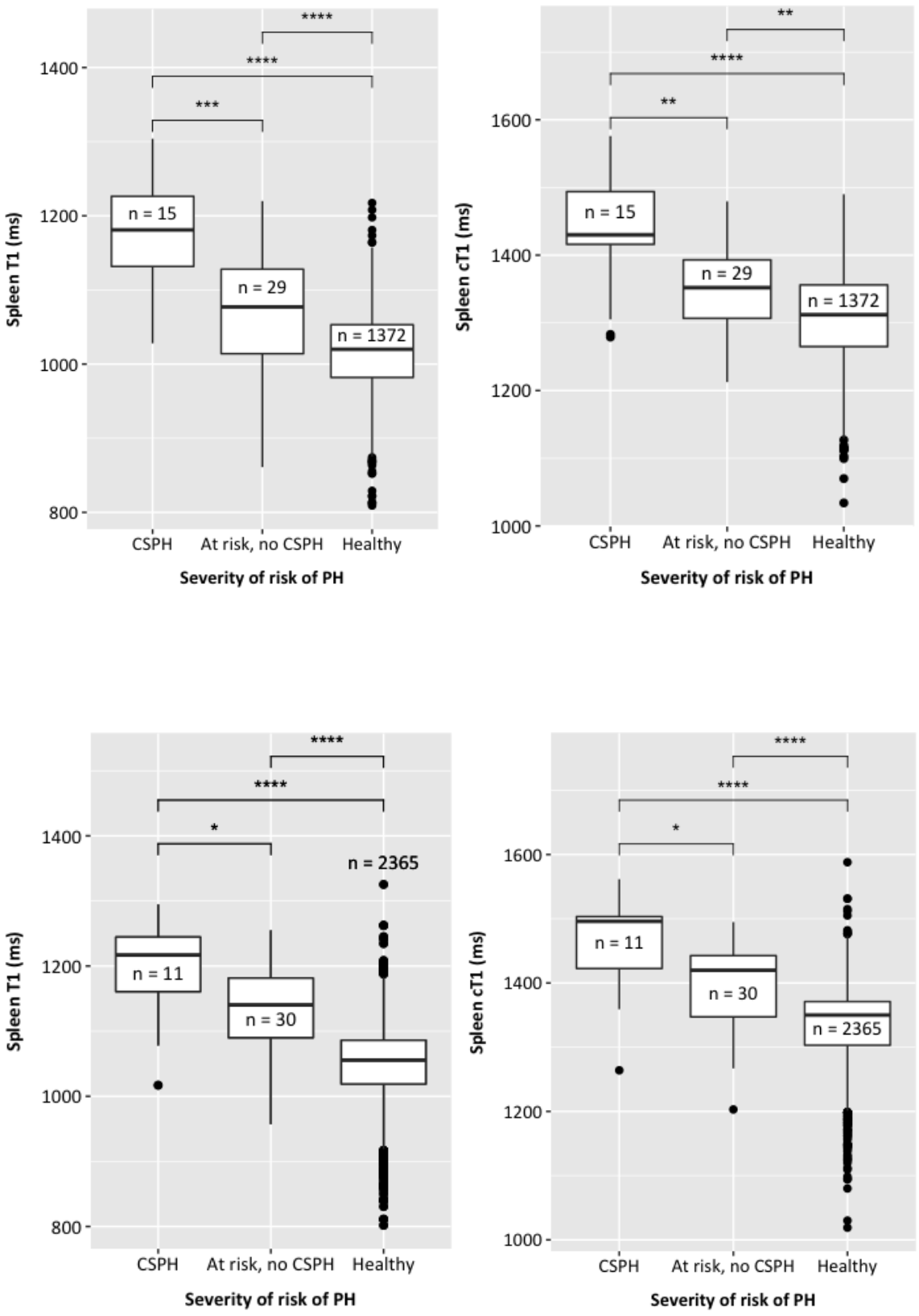


Figure 43: **Boxplots showing differences in spleen T_1/cT_1 by disease severity at 1.5T by sex.** Top row: Male, bottom row: female.

When combining all participants, volunteers and patients across studies, spleen cT_1 ranges were significantly different between all categories of disease severity (1415 vs 1343 vs 1325 ms, Figure 44). The difference between healthy participants and patients at risk, but without CSPH was small but statistically significant, probably because of the large number of healthy participants (mainly from UKBB).

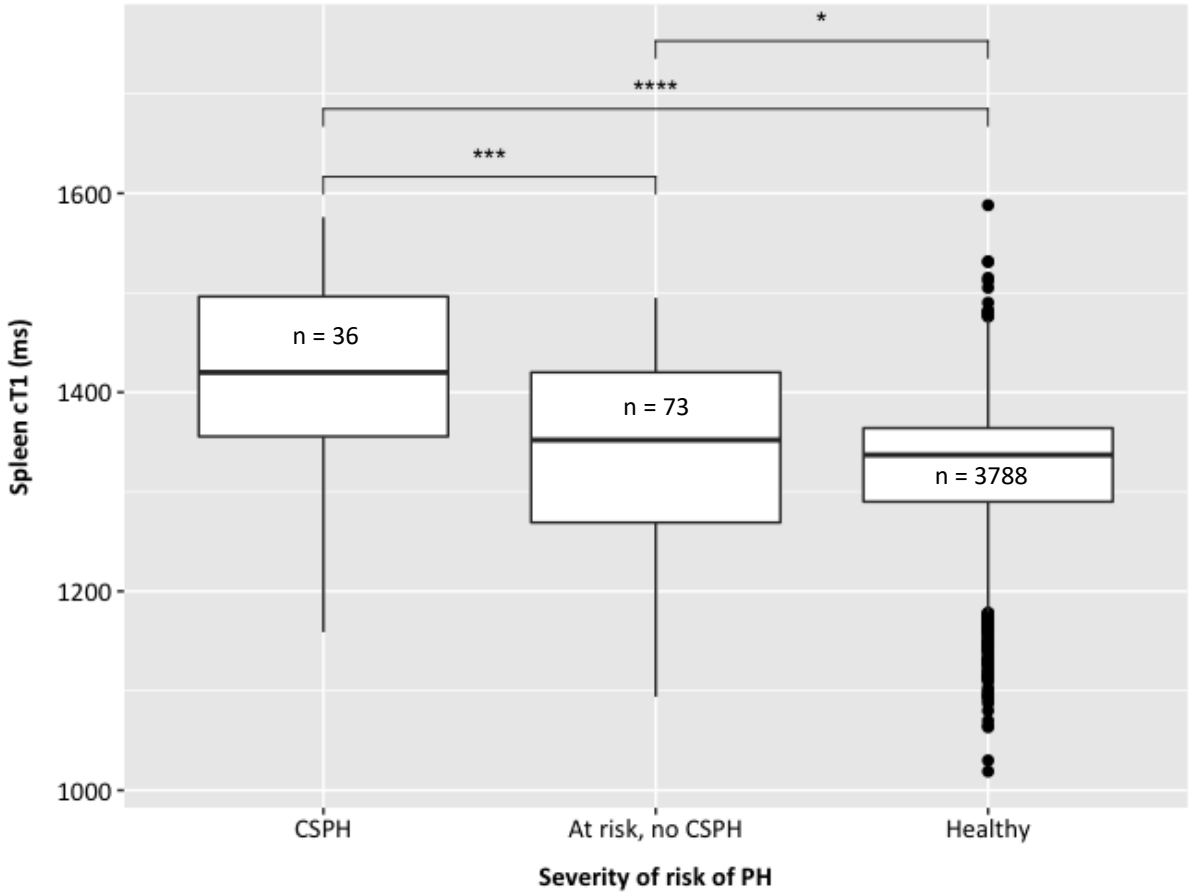


Figure 44: **Boxplots showing increasing spleen cT_1 with increasing disease severity**. Data was collated from all participants, healthy volunteers and patients across studies. Spleen cT_1 data was categorised by disease severity: healthy, assessed for, but with no diagnosis of clinically significant portal hypertension (CSPH), and those with a diagnosis of CSPH.

Spleen T_1 (at 1.5T) and cT_1 (across all groups) was significantly higher in the paediatric population from CCH than any other group (Figure 45). Mean spleen cT_1 from each group were: CCH: 1426 ± 72 ms, SGH 1363 ± 88 ms, OCMR: 1277 ± 95 ms, OCMR healthy volunteers: 1287 ± 92 ms, UKBB: 1325 ± 62 ms.

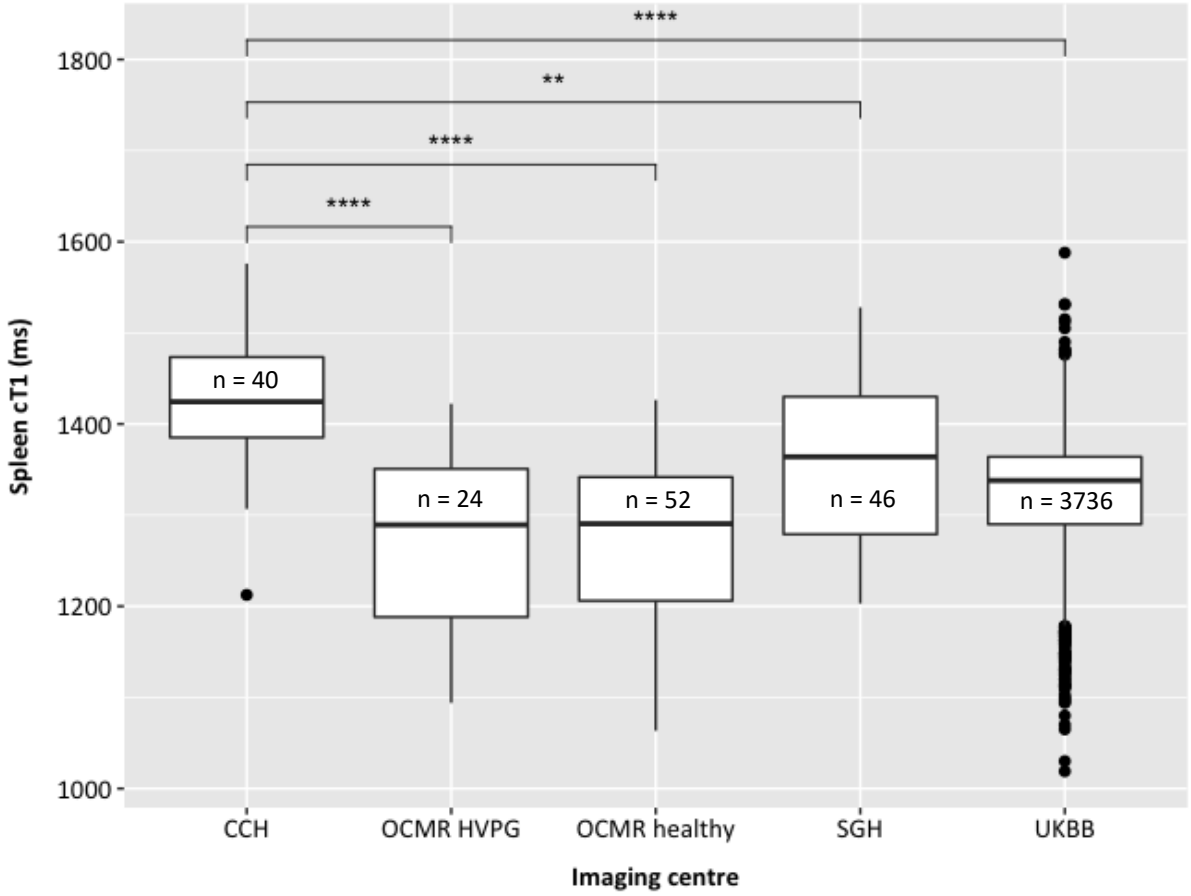


Figure 45: Boxplot showing the difference in paediatric spleen cT_1 from CCH to adult values.

For a greater visual representation of the spleen cT_1 values across study and severity of disease, boxplots were generated comparing spleen cT_1 values across studies and splitting each study into severity of PH as depicted in Figure 46.

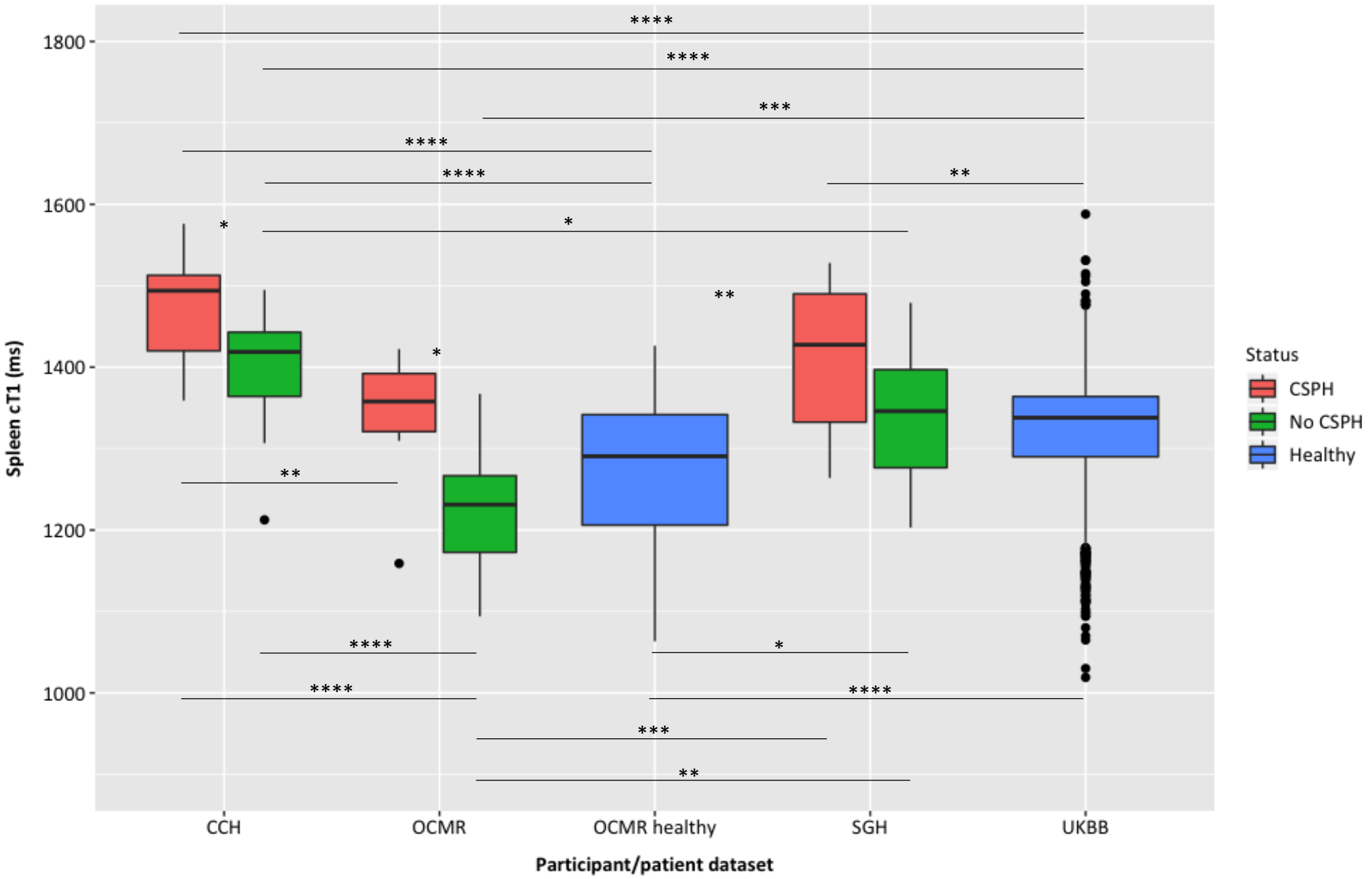


Figure 46: Boxplots showing spleen cT_1 across all centres categorised by severity. Spleen cT_1 values categorised into healthy (blue), assessed for but with no clinically significant portal hypertension (CSPH, green), and CSPH (red). Significance set at 0.05/7 = 0.007 to correct for multiple comparisons. * denotes $p < 0.007$, ** denotes $p < 0.0014$, *** denotes $p < 0.00014$, **** denotes $p < 0.000014$, no line/symbol denotes non-significance. A * with no line indicates relationship between adjacent boxplots.

8.7 Discussion

8.7.1 Main findings

This chapter has provided normal ranges for spleen T_1 at 1.5T and at 3T, and for spleen cT_1. We have further highlighted the utility of these measures for assessment of PH by showing that there are significant differences across disease severity. We have provided further evidence that women have higher spleen T_1 and cT_1 values than their male counterparts, indicating potential adjustments may be required for accurate diagnosis of CSPH. The effect sizes of these male-female differences were clinically relevant, with differences of 100 ms and 40 ms observed in the OCMR healthy volunteer and UKBB cohort respectively. We have observed higher spleen T_1/cT_1 in the paediatric population with autoimmune liver disease (AILD) compared to all other adult populations with CLD (from 63 to 149 ms differences). These male/female and child/adult differences were also clinically relevant in comparison to the spleen cT_1 differences observed between patients' disease severity categories of 'healthy' and CSPH (90 ms difference) and of 'assessed for, but without CSPH' and CSPH (72 ms difference).

8.7.2 UKBB and OCMR healthy volunteers' exploratory analysis

We derived spleen T_1/cT_1 values from two different groups; both prospectively recruited, but with different inclusion/exclusion criteria. The baseline characteristics of both groups (BMI, blood test results) were very similar with the main difference being the UKBB were older (55 vs 39 years old) than the healthy volunteers recruited at OCMR.

There were weak (r = 0.23) correlations between liver cT_1 and spleen cT_1 in the UKBB but a strong correlation between the two (r = 0.61) in the OCMR cohort. One reason we may have observed strong correlations between spleen and liver cT_1 in the OCMR cohort, but a much weaker one in the UKBB cohort is that reduction in liver cT_1 with increasing age was observed in women in the study by Mojtahed et al[189]. This could have affected the overall correlation. In a nominally healthy population, one would not necessarily expect to observe a correlation between variables that are only expected to be elevated in the presence of liver disease. One however might expect some correlation between the two due to factors that affect both liver and spleen T_1 such as hydration levels. Also, given the large number of participants involved in the UKBB, statistically significant; but clinically irrelevant correlations are likely to be observed.

Spleen cT_1 was negatively correlated with haemoglobin (Hb) concentration in both the OCMR and the UKBB cohort (r = -0.48 and r = -0.31 respectively, $p < 0.0001$ for both). Hb is generally higher in men than women, and spleen cT_1 was found to be lower in men than women. When taking sex into account, the correlations were: r = -0.13, $p < 0.0001$ and r = -0.20, $p < 0.0001$ in men and women respectively in the UKBB, and r = -0.17, p = ns and r = -0.47, p = 0.007 in men and women in the OCMR cohort. Haemoglobin levels can vary with menstrual phase[329]. The correlation therefore may be more pronounced in women, in whom we might expect to see a broader range of Hb levels due to this variation. We may also expect to see a broader range of Hb in younger women than in older women due to the onset of menopause, which may explain the stronger correlation coefficient in younger women (i.e. in the OCMR cohort vs the UKBB, r = -0.20, $p < 0.0001$ vs r = -0.47, p = 0.007).

Figure 33 highlighted the practical issues with utilising an iron-correction method that is only validated/functions in the range of liver T_1 values. This figure clearly outlined the inadequacy of the current iron-correction method being applied to spleen T_1 values, as after a certain point only the field strength correction mapping T_1 to cT_1 applied. This may be the reason that certain uncorrected T_1 data from previous chapters showed marginally better performance for diagnosis of CSPH than spleen cT_1.

Participants with BMI > 30 kg/m^2 had lower spleen T_1/cT_1 values than those with BMI $\leq$ 30 kg/m^2, but the difference was too small (17-20 ms) to be considered significant in the context of reproducibility of T_1 times[188]. Spleen T_1/cT_1 was also associated with a weak/modest negative correlation with BMI in both healthy cohorts, but the effect size from the UKBB is likely to be artificially enhanced by the large numbers involved and the significance of the correlation in the OCMR cohort disappeared when corrected for multiple comparisons. Although there have been no studies comparing spleen T_1 to BMI, height or weight, studies have shown that spleen size (both diameter and volume) is larger in taller people, people with higher BMI and in men vs women[330]. Interestingly we have observed the opposite relationships between spleen T_1/cT_1 and sex and spleen T_1/cT_1 and body size/BMI measurements in the UKBB and OCMR healthy cohort. We did not have height and weight measurements for individuals from the SGH or CCH cohort, but this analysis for effect of body size on spleen T_1 in patients with CLD or PH would be a logical next step. It is likely however, that the presence of any splenic congestion would affect the spleen T_1 measurements more so than body size, given the low correlations observed between spleen T_1 and height/weight (-0.19 and -0.20 respectively) in this UKBB cohort. To put the magnitude of these differences into context, the difference in spleen cT_1 between male and female, participants with and without obesity and difference between disease severity are summarised in Table 33 below.

Table 33: Significant differences in spleen cT_1 by dichotomous variables and disease severity.

Disease severity	Healthy **vs** assessed for but no CSPH **vs** CSPH	Differences
	1415 vs 1343 vs 1325 ms	72, 18, 90 ms
Dichotomous variables		
Men vs women (UKBB)	1336 vs 1304 ms	32 ms
Men vs women (OCMR healthy)	1307 vs 1206 ms	101 ms
Children (CCH) vs Adults (OCMR PH and SGH)	CCH: 1426 ms, SGH: 1363 ms, OCMR: 1277 ms	63, 149 ms
BMI > 30 kgm^{-2} vs BMI < 30 kgm^{-2}	1305 vs 1327 ms	22 ms

Much of this work was exploratory in nature, and therefore the associations with multiple comparators from the UKBB must be taken with caution. The significance of correlation between spleen T_1/cT_1 and many of the variables can be explained mostly by the sheer numbers of participants involved in this analysis. Therefore, in this discussion any of the parameters that had a significant Spearman's rank correlation coefficient < 0.3 in the UKBB study are not considered to be of clinical relevance, as with smaller numbers those relationships are likely to disappear. This was shown in the significant correlations observed between spleen T_1/cT_1 and blood tests from the UKBB cohort and the subsequent lack of

correlation between spleen T_1/cT_1 and the same routine blood tests in the heathy volunteers from the OCMR group e.g. in the UKBB spleen T_1 was correlated with GGT ($r = -0.16$, $p < 2.2x10^{-16}$), but not in the OCMR cohort ($r = -0.24$, $p = 0.09$). The means by which we derived these spleen parameters in the UKBB were automated, and although all scans were QC checked for liver metrics and general scan protocol issues (e.g. mistriggering), the same QC was not applied in the spleen – we however excluded all those with low spleen T_1 R-square values, fat swap effects and abnormal T_2* values. This means that although the range of values we have described is likely to be accurate, individual scans may be of suboptimal quality (due to misalignment through patient motion and/or poor segmentation) and further work should be done to carefully, manually sift through the remaining dataset and provide a definitive normal, healthy range.

Women had higher spleen T_1 and cT_1 in both the UKBB cohort and OCMR healthy volunteer group. This has been observed previously in liver cT_1 by researchers in our group and is postulated to be due to differences in water retention due to hormonal changes. The shMOLLI sequence does have a heart-rate dependence on the evaluation of T_1 times that is more pronounced in longer T_1 times. Men and women in general have different heart rates and this may be another contributing reason to the difference in spleen T_1/cT_1 in men and women. Further to this, difference in B_0 field corrections may be present in men vs women due to the difference in spleen size. All these incremental difference between men and women may be contributing to the observed significant differences in spleen cT_1/T_1 in men and women. There was a larger difference between male and female spleen cT_1 values observed in the slightly younger, healthy OCMR cohort than in the healthy UKBB cohort (100 ms vs 40 ms

respectively). We did observe a statistically significant difference in women from the UKBB above and below the ages of 52, (considered to be the average age of the onset of the menopause in the UK). However, the effect size was a difference of only 9 ms on average and therefore cannot be considered clinically relevant. Larger spleen size in men vs women has been reported before, with separate ranges of normal spleen size having been recommended for men and women as shown in a study of 1,200 volunteers[330]. We did not have data on spleen length available but did have liver and spleen volumes available and showed that men had larger spleens by volume than women (0.16 vs 0.10 litres, $p < 0.0001$), in line with the literature[330]. It is worth noting that it may be expected to find a gender difference in spleen volume or organ volume in general since we have not adjusted for body size here.

Within data from OCMR (3.0T), we observed a trend towards higher spleen cT_1 in healthy volunteers than in those assessed for, but without CSPH (Figure 40). On closer inspection, the healthy volunteers contained a much higher percentage of women than the patient group (64% vs 21%), who we have shown to have higher spleen cT_1, driving the observed upward trend. When we separated the cohorts into men and women, we saw a marked effect on the difference in spleen T_1/cT_1 in relation to disease severity in men vs women, due to the higher T_1/cT_1 in the healthy women controls (Figure 41). The healthy women showed a trend towards being younger (35 vs 47 years old, $p = 0.02$) than those with suspected PH and this difference in spleen T_1/cT_1 therefore may be either to do with age or hormonal changes affecting fluid retention. Given the small numbers of female patients assessed for PH (n = 6) in comparison with the healthy volunteers (n = 39), the significance levels should be taken

with caution. In similar analysis carried out with the remaining groups at 3.0T, spleen T_1/cT_1 rose with disease severity in a similar fashion for both women and men.

This indicates that for any future clinical utility, either different cut-offs should be considered for men and women or normalisation for sex should be calculated to avoid misdiagnosis. The difference between healthy and 'assessed for, but without CSPH' was small (18 ms) and may be driven by the large participant numbers in the UKBB – a better analysis would be to take a smaller, random sample of some of the UKBB for this and other analyses.

8.7.3 Healthy datasets comparison to patients with CLD

In the previous Chapter spleen T_1 was shown to perform similarly and in some cases slightly (but not significantly) better than spleen cT_1 for discrimination of CSPH. We could not directly compare spleen T_1 values between different study groups at different field strengths, as 'raw' T_1 values are almost linearly correlated with field strength. However, we observed similar relative differences between ranges of spleen cT_1 across the different studies. This underscores the benefit of the field strength standardisation algorithm if spleen cT_1 is to be used in clinical practice allowing a direct comparison across groups. Moreover, the size of the difference in splenic T_1 between 1.5T and 3T in groups of similar characteristics is greater than 200 ms in comparison to the largest difference between the pooled clinical groups (healthy vs CSPH, 90 ms) indicating a clear requirement for standardisation.

We observed a steady increase in spleen cT_1 with increasing severity of disease (Figure 44). Our criteria for disease severity from SGH involved using future episodes of variceal bleeding as a determinant of whether the patients had CSPH rather than using the HVPG $\geq$ 10 mmHg. We believe this was justified as the HVPG should be the gold-standard measure of severity of

PH, but it did not show prognostic value in Chapter 5; this also allowed us to include the patients with NCPH (as the HVPG is falsely lowered in those patients). This is an encouraging finding, however in almost all of the populations, the ranges of spleen T_1/cT_1 were quite large and in many cases the distributions between severity overlapped significantly (Figure 39). This could indicate that additional imaging (e.g. liver cT_1), clinical, or non-invasive (e.g. platelet count) information may be required in making a correct diagnosis of PH or CSPH and that if used in clinical trials/research studies as an endpoint, then large sample sizes may be needed.

Data was collated from the studies included in this thesis as well as other reported sources to give an overview of spleen T_1/cT_1 and liver cT_1 values in different populations (Table 34). The splenic standard deviations and IQRs in the UKBB and the OCMR healthy cohort are higher than the IQR reported in a reference range for liver cT_1 in the UKBB[189] (643-694 ms, a difference of 51 ms vs 1290-1364, 74 ms difference) or in the OCMR healthy cohort. This is understandable given the higher absolute range, and this difference becomes minimal if its coefficient of variation (CoV) is taken into account (6.5% for spleen cT_1 vs 6% for liver cT_1, data from UKBB healthy cohort). However, in the patient populations, the spleen cT_1 IQRs are similar to their corresponding liver cT_1 values, indicating more variability amongst normal or healthy people in spleen cT_1 compared to liver cT_1.

Table 34: Ranges of spleen cT_1, T_1 and liver cT_1 values

	Spleen cT_1 (ms)	Spleen T_1 (ms)	Liver cT_1 (ms)
UKBB	1325 ± 61 1338 (1290-1364)	1039 ± 67 1043 (1004-1077)	- 666 (643-694)[189]
OCMR healthy	1270 ± 93 1291 (1206-1342)	1257 ± 96 1272 (1187-1334)	765 $\pm$ 70 756 (718-804)

OCMR HVPG	1277 ± 95 1289 (1188-1351)	1264 ± 95 1254 (1189-1337)	954 ± 153 913 (858-1028)
SGH	1363 ± 88 1364 (1279-1430)	1107 ± 100 1104 (1035-1182)	841 ± 121 805 (767-928)
CCH (AILD)	1426 ± 72 1424 (1385-1473)	1148 ± 72 1143 (1106-1202)	985 ± 72 996 (934-1018)
CCH (healthy[331])	-	1172 ± 71 & 1356 ± 87 1.5T & 3T	581 ± 64 & 783 ± 88 1.5T & 3T

All datasets shown as mean ± standard deviation and median (interquartile range, (IQR)) to allow direct comparison with liver cT_1 values which were non-normally distributed.

We initially thought in the previous Chapter that the higher spleen cT_1 in the paediatric population from CCH could be accounted for the by underlying aetiologies of the cohorts as the CCH cohort had markedly different aetiologies of liver disease (all patients being exclusively diagnosed with AILD) than the ones in SGH and OCMR. Alternatively, the difference could be that paediatric spleens have different values to adults. A recent publication by the same group we collaborated with at CCH showed ranges of normal paediatric values for spleen T_1 at 1.5T and 3T field strengths[331]. These were shown to be 1172 ± 71 ms at 1.5T and 1356 ± 87 ms at 3T, both approx. 100 ms higher than their respective values obtained in our studies here in the UKBB (1.5T, 1039 ± 67 ms) and at OCMR (3T, 1257 ± 96). This provides further evidence that children may have higher spleen T_1 values than adults in both healthy volunteers and in CLD. There is of course a chance that CCH is returning systematically high values, but we have found no evidence of that. In the study of healthy children[331], there was no difference in spleen T_1 values between males and females, whereas we have consistently observed differences in the healthy adult cohorts in this Chapter. This

provides a possible further evidence for the explanation that the differences are related to hormonal changes that occur after puberty, which could be causing changes in levels of water retention in women as mentioned above.

We have compared spleen cT_1 values from studies carried out at both 3T and 1.5T field strength. However, since the field strength correction (as well as the iron correction) that transforms T_1 to cT_1 is valid only for liver tissue and liver T_1 ranges, it is unclear how much confidence we should have in directly comparing these spleen cT_1 values across field strengths.

The spleen is split into two main parts, the red pulp (active in mechanical filtration of red blood cells) and white pulp (active in immune response), with an area called the marginal zone straddling the two, and the white pulp being enclosed by the red pulp. The effect of immune response to spleen T_1/cT_1 was briefly touched upon in the previous Chapter's discussion. Heightened levels of immune response may express itself in an enlargement of only the white pulp of the spleen, and this extra contribution could drive overall spleen T_1/cT_1 values upwards. In fact, studies have shown that splenomegaly induced by chronic PH exhibits enlarged white pulp and marginal zone areas[332], whereas spleens enlarged due to congestive splenomegaly exhibit a less prominent white pulp and more distinct red pulp[333]. If there is some way of separating the contributions to/ranges of the T_1 signal from white and red pulp, this may help to account for any confounding effects of the immune system. This would require significant interdisciplinary co-operation to obtain the requisite tissue samples.

The spleen has an important role in storing and recycling iron. Old or damaged red blood cells are filtered by the spleen and broken down by macrophages. Many of these old cells still contain haemoglobin, which is metabolised by the spleen and stored as ferritin until later

released. Now that there is further evidence of spleen T_1 relaxometry's ability to assess severity of PH, future work could focus on the utility of the current iron-correction method, which is not validated in the spleen and therefore lacks physiological rationale. Histology is rarely performed on the *in vivo* spleen due to its propensity to bleed profusely, making validation of iron content difficult. Splenectomy is sometimes indicated for haematological disorders or can be performed together with pancreatic surgery. Splenectomy is also occasionally performed to relieve symptoms of PH. These may present opportunities for paired histologic and non-invasive evaluation of splenic iron content. This is likely to be necessary as it is known that the spleen can experience iron overload, sometimes exhibiting cases of haemosiderin 'clumps' called Gamma-Gandy bodies.

In conclusion, over these last two chapters we have collated and analysed MRI data from multiple centres and shown that spleen T_1 and cT_1 both may have diagnostic and prognostic value in the assessment of PH. If further evidence from larger numbers of patients in more sites shows similar results, these could be used, perhaps in conjunction with liver cT_1 and platelet count for routine assessment of severity of PH in patients with CLD.

Chapter 9: Conclusions and suggested further work

9.1 Summary of main findings

This thesis has further examined the diagnostic and prognostic ability of multiparametric MRI methods in patients with varying severities of liver disease. It was shown for the first time that liver cT_1 can quantify changes in liver fibroinflammation in response to direct acting antiviral treatment (DAA) for hepatitis C virus (HCV). Liver cT_1 exhibited further prognostic significance, especially in patients with non-alcoholic fatty liver disease (NAFLD), alcohol-related liver disease (ArLD) and viral hepatitis (VH), and when failure/unreliability rates were taken into account, could predict clinical outcomes while transient elastography (TE) could not. We provided further evidence that spleen T_1/cT_1 and liver cT_1 have utility in the diagnosis and prognosis of clinically significant portal hypertension (CSPH) and presented a reference range for spleen T_1/cT_1 from the United Kingdom Biobank resource (UKBB).

9.2 Liver cT_1 prediction of clinical outcomes

This thesis has shown further evidence that liver cT_1 is a good predictor of clinical outcomes, with particular value in screening patients with the major causes of liver disease (NAFLD, ArLD and VH in the western world, correctly predicting all liver-related clinical events in those patients. Liver cT_1 was also predictive of outcomes in patients with compensated disease only, in which a screening tool would be most useful. In a head-to-head analysis comparing the effect of failure and unreliability rates on MRI and TE, liver cT_1 > 825 ms was predictive of liver-related clinical outcomes, but not TE > 8kPa. This is important, as TE is currently used widely in clinical practice, and although costed at less than MRI with Liver*MultiScan*™ (£199[334] vs £331 (£200 for analysis and reporting + £131 for MRI scan time)), its failure rate means that these costs could be recouped by using MRI and avoiding biopsies (£497-553)[335].

Although we have followed up almost twice as many patients, over longer follow-up than previously, more data from independent cohorts are necessary to further build on this evidence base.

We believe that liver cT_1 might have particular use in early detection of liver disease. The UKBB resource would be an appropriate dataset in which to investigate this. Participants within the UKBB should be stratified into groups, e.g. those with no signs of liver fibroinflammation, those with MR-based signs of fibroinflammation but no blood-based signs, those with blood-based signs of liver fibroinflammation but no MR-based signs; and those with both MR and blood-based signs of liver fibroinflammation. The UKBB aims to release participant outcome data every 3 years, and these participants' development of clinical outcomes or first referral to liver services should be compared. This would tell us a great deal about liver cT_1's sensitivity in relation to other non-invasive biomarkers for early detection of outcomes.

9.3 Liver cT_1 for monitoring of disease progression/regression

This thesis has shown liver cT_1 can accurately monitor patients' response to therapy. We observed significant changes in fibroinflammatory disease and liver fat up to 48 weeks after end of treatment (EoT) in patients with HCV treated with DAAs. Patients with cirrhosis were observed to regress their liver cT_1 more slowly than those without.

Long-term regression in liver MR parameters should also be assessed, especially in patients with cirrhosis prior to treatment. These patients also require extended follow-up for development of hepatocellular carcinoma (HCC), a limitation of the relatively short follow-up in this study. Additionally, in these particular populations of patients treated for HCV, different cut-offs for liver cT_1 may be necessary for prediction of HCC risk.

Response to treatment could be extended to assessing response to abstinence from alcohol consumption in patient with alcohol-related liver disease (ArLD). Abstinence from drinking is the strongest predictor of survival in patients with alcohol-related liver disease, but there are still significant numbers who go on to develop poor outcomes even if they stop drinking excess alcohol[336]. In the same way as we have described the importance of assessing those at risk of HCC development even after successful treatment for HCV, if changes in liver cT_1 post abstinence from alcohol can predict those who are at continued higher risk of developing complications, it would be of significant clinical utility.

Changes in liver T_2* were observed over the course of this study. Further work should extend the assessment of changes in MR parameters from focussing on fibroinflammatory change to changes in liver iron content (LIC). The T_2* measurement of LIC underpins the iron correction algorithm and therefore its sensitivity to change would be useful to know for future longitudinal studies. This could be assessed in patients undergoing iron infusion or venesection for treatment of anaemia or haemochromatosis, in whom we might expect to observe significant changes in LIC.

9.4 Spleen T_1 and cT_1 for assessment of portal hypertension

Following on from research in our group, this thesis has further established that both spleen T_1/cT_1 and liver cT_1 has clinical utility in classification of patients with portal hypertension (PH). This work compares similarly to studies involving magnetic resonance elastography (MRE), and also sits within the range of diagnostic accuracies reported for liver and spleen stiffness by TE. Taking into consideration the failure/unreliability rate and coefficient of variation (CoV) of TE on its clinical application, along with studies showing spleen stiffness measurements by acoustic radiation force impulse (ARFI) fail to take measurements of

smaller spleens, these encourage us to generate further evidence on the utility of spleen T_1 and cT_1 as surrogate markers of portal pressure. In particular, larger patient cohorts, with more balanced numbers of patients with and without clinically significant PH (CSPH) as defined by hepatic venous pressure gradient (HVPG) are required.

We presented a reference range of 'normal' spleen T_1 and cT_1 values, utilising the United Kingdom Biobank (UKBB) resource, thereby contributing significantly to the current literature relating to normal spleen T_1 ranges.

Paediatric spleen T_1/cT_1 was shown to be significantly higher amongst patients with liver disease (autoimmune liver disease (AILD)) than in adults with CLD, and we also observed a difference between our adult 'normal' spleen T_1 values and those of paediatric healthy volunteers from another study. This shows that in addition to significant differences in spleen T_1/cT_1 observed in adult women and men, different ranges may be required for the paediatric population. These differences were large and comparable with differences in spleen T_1/cT_1 disease status, and therefore require addressing if spleen T_1/cT_1 is to become clinically useful.

Lastly, albeit in a small cohort of patients with PH who were on nonselective beta-blocker (NSBB) treatment, spleen cT_1 showed excellent prediction of future variceal bleeding, potentially able to identify those at risk of failing primary prophylaxis for variceal haemorrhage.

Further work could involve assessing patient pre- and post- transjugular intrahepatic shunt (TIPS) procedure to investigate the ability of abdominal T_1 mapping to assess effectiveness of the procedure.

9.5 Advancement in MR techniques

Although we have used magnetic spectroscopy (MRS) in this study, Liver*MultiScan*™ currently uses MRI proton density fat fraction (PDFF) techniques, which correlate excellently with MRS[337], and combined with T_1 and T_2* mapping allow a patient examination in approx. 15 mins, allowing high throughput assessment as in the UK Biobank. The Liver*MultiScan*™ protocol should be trialled in multiple centres in an intention-to-diagnose setting alongside TE to assess which methods provides best overall value for money to a healthcare system.

Inflammation increases liver cT_1, in some cases to such an extent that the cT_1 value is comparable with that of a patient with cirrhosis. MR techniques that might separate contributions from fibrosis and inflammation in the liver should be pursued (intravoxel incoherent motion diffusion weighted imaging for example) as they could be of clinical importance especially in the tracking of patients undergoing treatment for their underlying liver disease and in clinical trials for antifibrotic treatments and NASH treatments specifically. Ideally one would want multiparametric MRI to be a 'one-stop shop' for liver disease assessment. An important part of liver disease assessment is screening for hepatocellular carcinoma (HCC) which is currently performed with by ultrasound or contrast enhanced MRI. Our technique currently measures liver cT_1 through an imaging 'slice', but 3 dimensional T_1 mapping techniques could map the whole liver in a much shorter time. This could improve detection of smaller HCCs, which can be missed on US, and also help surgical planning for resection. General applicability is another important facet of a diagnostic tool. Currently the shMOLLI[197] sequence is mainly available on MRI scanners that are either active in research or in those who update their software regularly. This is often not done in purely clinical centres - across the US in particular - and therefore the current MRI acquisition technique we use is

suboptimal for the widest coverage. 3D T_1 mapping techniques that can be used on a wider variety of scanners than shMOLLI can are in development.

9.6 Practical Utility

From the data presented here showing cT_1's prognostic capability and its ability to quantify changes in liver disease over the medium term, cT_1 could be useful as a prognostic or monitoring biomarker, or potentially even as a clinical trial endpoint. Drug trials for non-alcoholic steatohepatitis (NASH) and antifibrotic therapies need better recruitment and retainment, which are affected by the FDA's requirement for multiple biopsies. Liver cT_1 could significantly impact on these statistics and the process of drug trials/approval.

In clinical care, we would envisage MRI, perhaps alongside blood biomarkers to be used as a screening tool to avoid unnecessary biopsies, HVPG procedures or endoscopies, thereby reducing healthcare costs and increasing patient comfort.

Multiparametric MRI's evidence base as non-invasive diagnostic tool for liver disease is being bolstered continuously. However, there are other liver-specific tools, namely TE and MRE that also show good diagnostic accuracy and prediction of clinical outcomes. Although we have shown advantages of cT_1, these techniques should be compared head-to-head over multiple centres, and feasibility and cost-benefit analyses performed to determine practical utility.

References

1. James, S. L. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1789–1858 (2018).
2. Mokdad, A. A. *et al.* Liver cirrhosis mortality in 187 countries between 1980 and 2010: A systematic analysis. *BMC Med.* **12**, 1–24 (2014).
3. Younossi, Z. M. *et al.* Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* **64**, 73–84 (2016).
4. World Health Organization. *Global Hepatitis Report 2017. Geneva: World Health Organization; 2017 Licence: CC BY-NC-SA 3.0 IGO. WHO* (2017). doi:10.1149/2.030203jes
5. World Health Organization. *Global status report on alcohol and health 2018. Geneva: World Health Organization; 2018. Licence: CC BY-NC-SA 3.0 IGO*. (2018).
6. Scaglione, S. *et al.* The Epidemiology of Cirrhosis in the United States. *J. Clin. Gastroenterol.* **49**, 690–696 (2015).
7. 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.* **44**, 217–231 (2006).
8. World Health Organization. *Global Health Estimates 2016: Disease burden by Cause, Age, Sex, by Country and by Region, 2000-2015*. (2018).
9. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* **68**, 394–424 (2018).
10. National Cancer Institute. SEER Cancer Statistics Review, 1975-2015. *SEER* (2019). Available at: https://seer.cancer.gov/csr/1975_2015/index.html.
11. Tengowski, M., Fuerst, T. & Sirlin, C. Magnetic resonance imaging fatty liver changes following surgical , lifestyle , or drug treatments in obese , non-alcoholic fatty liver disease , or non-alcoholic steatohepatitis subjects. *Imaging Med* **9**, 195–214 (2017).
12. Lackner, C. *et al.* Histological parameters and alcohol abstinence determine long-term prognosis in patients with alcoholic liver disease. *J. Hepatol.* **66**, 610–618 (2017).

13. Altamirano, J. *et al.* Alcohol abstinence in patients surviving an episode of alcoholic hepatitis: Prediction and impact on long-term survival. *Hepatology* **66**, 1842–1853 (2017).
14. Williams, R. *et al.* Addressing liver disease in the UK: a blueprint for attaining excellence in health care and reducing premature mortality from lifestyle issues of excess consumption of alcohol, obesity, and viral hepatitis. *Lancet* **384**, 1953–1997 (2014).
15. Foundation for Liver Research. Financial case for action on liver disease. (2017).
16. Public Health England. Corporate report: PHE research 2016 to 2017: annual review. (2018). Available at: https://www.gov.uk/government/publications/phe-research-annual-review/phe-research-2016-to-2017-annual-review.
17. Public Health England. The 2nd Atlas of variation in risk factors and healthcare for liver disease in England. 12–30 (2017).
18. Stepanova, M. *et al.* Direct and Indirect Economic Burden of Chronic Liver Disease in the United States. *Clin. Gastroenterol. Hepatol.* **15**, 759-766.e5 (2017).
19. Regev, A. *et al.* Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. *Am. J. Gastroenterol.* **97**, 2614–2618 (2002).
20. Gronbaek, K. *et al.* Interobserver variation in interpretation of serial liver biopsies from patients with chronic hepatitis C. *J. Viral Hepat.* **9**, 443–449 (2002).
21. Poynard, T. *et al.* Concordance in a World without a Gold Standard: A New Non-Invasive Methodology for Improving Accuracy of Fibrosis Markers. *PLoS One* **3**, e3857 (2008).
22. Thalheimer, U., Bellis, L., Puoti, C. & Burroughs, A. K. 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.* **22**, 5–7 (2011).
23. Patch, D. *et al.* Single portal pressure measurement predicts survival in cirrhotic patients with recent bleeding. *Gut* **44**, 264–269 (1999).
24. Ripoll, C. *et al.* Hepatic Venous Pressure Gradient Predicts Clinical Decompensation in Patients With Compensated Cirrhosis. *Gastroenterology* **133**, 481–488 (2007).
25. Ripoll, C. *et al.* Hepatic venous pressure gradient predicts development of hepatocellular carcinoma independently of severity of cirrhosis. *J Hepatol* **50**, 923–928 (2009).
26. Rincón, D. *et al.* Prognostic value of hepatic venous pressure gradient in patients with

compensated chronic hepatitis C-related cirrhosis. *Scand. J. Gastroenterol.* **48**, 487–95 (2013).

27. de Franchis, R. Expanding consensus in portal hypertension. *J. Hepatol.* **63**, 743–752 (2015).
28. Dufour, D. R. *et al.* Diagnosis and monitoring of hepatic injury. I. Performance characteristics of laboratory tests. *Clin. Chem.* **46**, 2027–49 (2000).
29. Newsome, P. N. *et al.* Guidelines on the management of abnormal liver blood tests. *Gut* **67**, 6–19 (2018).
30. Nunes, D. *et al.* Noninvasive Markers of Liver Fibrosis Are Highly Predictive of Liver-Related Death in a Cohort of HCV-Infected Individuals With and Without HIV Infection. *Am. J. Gastroenterol.* **105**, 1346–1353 (2010).
31. Vergniol, J. *et al.* Noninvasive Tests for Fibrosis and Liver Stiffness Predict 5-Year Outcomes of Patients With Chronic Hepatitis C. *Gastroenterology* **140**, 1970-1979.e3 (2011).
32. Hsu, W. F. *et al.* Rapid decline of noninvasive fibrosis index values in patients with hepatitis C receiving treatment with direct-acting antiviral agents. *BMC Gastroenterol.* **19**, 1–9 (2019).
33. Parkes, J. *et al.* Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. *Gut* **59**, 1245–1251 (2010).
34. Ngo, Y. *et al.* A prospective analysis of the prognostic value of biomarkers (FibroTest) in patients with chronic hepatitis C. *Clin. Chem.* **52**, 1887–1896 (2006).
35. Naveau, S. *et al.* Diagnostic and prognostic values of noninvasive biomarkers of fibrosis in patients with alcoholic liver disease. *Hepatology* **49**, 97–105 (2009).
36. Sherman, K. E. *et al.* Validation and Refinement of Noninvasive Methods to Assess Hepatic Fibrosis: Magnetic Resonance Elastography Versus Enhanced Liver Fibrosis Index. *Dig. Dis. Sci.* (2019). doi:10.1007/s10620-019-05815-z
37. Irvine, K. M. *et al.* The Enhanced liver fibrosis score is associated with clinical outcomes and disease progression in patients with chronic liver disease. *Liver Int.* **36**, 370–377 (2016).
38. National Institute for Health and Care Excellence. *Cirrhosis in over 16s: assessment and management*. (2016).
39. National Institute for Health and Care Excellence (NICE). *Hepatitis B (chronic): diagnosis*

and management. (2018).

40. Hsu, C. *et al.* Magnetic Resonance vs Transient Elastography Analysis of Patients With Nonalcoholic Fatty Liver Disease: A Systematic Review and Pooled Analysis of Individual Participants. *Clin. Gastroenterol. Hepatol.* **17**, 630-637.e8 (2019).
41. Castéra, L. *et al.* Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. *Hepatology* **51**, 828–35 (2010).
42. Ferraioli, G. Point shear wave elastography method for assessing liver stiffness. *World J. Gastroenterol.* **20**, 4787 (2014).
43. Lupsor, M. *et al.* Performance of a new elastographic method (ARFI technology) compared to unidimensional transient elastography in the noninvasive assessment of chronic hepatitis C. Preliminary results. *J. Gastrointestin. Liver Dis.* **18**, 303–10 (2009).
44. Friedrich-Rust, M. *et al.* Performance of Transient Elastography for the Staging of Liver Fibrosis: A Meta-Analysis. *Gastroenterology* **134**, 960–974 (2008).
45. Eddowes, P. J. *et al.* Accuracy of FibroScan Controlled Attenuation Parameter and Liver Stiffness Measurement in Assessing Steatosis and Fibrosis in Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology* **156**, 1717–1730 (2019).
46. Bazerbachi, F. *et al.* Range of Normal Liver Stiffness and Factors Associated With Increased Stiffness Measurements in Apparently Healthy Individuals. *Clin. Gastroenterol. Hepatol.* **17**, 54-64.e1 (2019).
47. Berzigotti, A. Review Non-invasive evaluation of portal hypertension using ultrasound elastography. *J. Hepatol.* **67**, 399–411 (2017).
48. Song, J. *et al.* Comparison of three cut-offs to diagnose clinically significant portal hypertension by liver stiffness in chronic viral liver diseases: a meta-analysis. *Eur. Radiol.* **28**, 5221–5230 (2018).
49. Colecchia, A. *et al.* Measurement of Spleen Stiffness to Evaluate Portal Hypertension and the Presence of Esophageal Varices in Patients With HCV-Related Cirrhosis. *Gastroenterology* **143**, 646–654 (2012).
50. Stefanescu, H. *et al.* Spleen stiffness measurement using Fibroscan for the noninvasive assessment of esophageal varices in liver cirrhosis patients. *J. Gastroenterol. Hepatol.* **26**, 164–70 (2011).
51. Sharma, P. *et al.* Spleen stiffness in patients with cirrhosis in predicting esophageal varices. *Am. J. Gastroenterol.* **108**, 1101–7 (2013).

52. Zykus, R., Jonaitis, L., Petrenkienė, V., Pranculis, A. & Kupčinskas, L. Liver and spleen transient elastography predicts portal hypertension in patients with chronic liver disease: a prospective cohort study. *BMC Gastroenterol.* **15**, 183 (2015).
53. Tseng, Y. *et al.* Spleen and liver stiffness for noninvasive assessment of portal hypertension in cirrhotic patients with large esophageal varices. *J. Clin. Ultrasound* **46**, 442–449 (2018).
54. Colecchia, A. *et al.* A combined model based on spleen stiffness measurement and Baveno VI criteria to rule out high-risk varices in advanced chronic liver disease. *J. Hepatol.* **69**, 308–317 (2018).
55. Bastard, C. *et al.* A Novel FibroScan Examination Dedicated to Spleen Stiffness Measurement. *Ultrasound Med. Biol.* **44**, 1616–1626 (2018).
56. Stefanescu, H. *et al.* A novel spleen-dedicated stiffness measurement by FibroScan® improves the screening of high-risk oesophageal varices. *Liver Int.* **40**, 175–185 (2020).
57. RONTGEN, W. C. ON A NEW KIND OF RAYS. *Science (80-.).* **3**, 227–231 (1896).
58. Ludwig, G. D. & Struthers, F. W. *Considerations Underlying the use of Ultrasound to Detect Gallstones and Foreign Bodies in Tissue*. *Naval Medical Research Institute* (1949).
59. Bloch, F. Nuclear Induction. *Phys. Rev.* **70**, 460–474 (1946).
60. Purcell, E. M., Torrey, H. C. & Pound, R. V. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys. Rev.* **69**, 37–38 (1946).
61. Damadian, R. Tumor Detection by Nuclear Magnetic Resonance. *Science (80-.).* **171**, 1151–1153 (1971).
62. LAUTERBUR, P. C. Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. *Nature* **242**, 190–191 (1973).
63. Mansfield, P. & Maudsley, A. A. Medical imaging by NMR. *Br. J. Radiol.* **50**, 188–194 (1977).
64. Mansfield, P., Pykett, I. L., Morris, P. G. & Coupland, R. E. Human whole body line-scan imaging by NMR. *Br. J. Radiol.* **51**, 921–922 (1978).
65. Brant-Zawadzki, M. MR imaging of the brain. *Radiology* **166**, 1–10 (1988).
66. Modic, M. T. *et al.* Nuclear magnetic resonance imaging of the spine. *Radiology* **148**, 757–762 (1983).
67. Chezmar, J. L. *et al.* Liver and abdominal screening in patients with cancer: CT versus

MR imaging. *Radiology* **168**, 43–47 (1988).

68. Heiken, J. P. & Lee, J. K. T. MR imaging of the pelvis. *Radiology* **166**, 11–16 (1988).
69. Qayyum, A. *et al.* MRI Steatosis Grading: Development and Initial Validation of a Color Mapping System. *Am. J. Roentgenol.* **198**, 582–588 (2012).
70. McPherson, S. *et al.* Magnetic resonance imaging and spectroscopy accurately estimate the severity of steatosis provided the stage of fibrosis is considered. *J. Hepatol.* **51**, 389–397 (2009).
71. Dixon, W. T. Simple proton spectroscopic imaging. *Radiology* **153**, 189–194 (1984).
72. Hong, C. W. *et al.* MRI proton density fat fraction is robust across the biologically plausible range of triglyceride spectra in adults with nonalcoholic steatohepatitis. *J. Magn. Reson. Imaging* **47**, 995–1002 (2018).
73. Haufe, W. M. *et al.* Accuracy of PDFF estimation by magnitude-based and complex-based MRI in children with MR spectroscopy as a reference. *J. Magn. Reson. Imaging* **46**, 1641–1647 (2017).
74. Di Martino, M. *et al.* Comparison of magnetic resonance spectroscopy, proton density fat fraction and histological analysis in the quantification of liver steatosis in children and adolescents. *World J. Gastroenterol.* **22**, 8812 (2016).
75. Roldan-Valadez, E. *et al.* In vivo 3T spectroscopic quantification of liver fat content in nonalcoholic fatty liver disease: Correlation with biochemical method and morphometry. *J. Hepatol.* **53**, 732–737 (2010).
76. Raptis, D. A. *et al.* MRI: the new reference standard in quantifying hepatic steatosis? *Gut* **61**, 117–127 (2012).
77. Hallsworth, K. *et al.* Resistance exercise reduces liver fat and its mediators in non-alcoholic fatty liver disease independent of weight loss. *Gut* **60**, 1278–1283 (2011).
78. Pooler, B. D. *et al.* Monitoring Fatty Liver Disease with MRI Following Bariatric Surgery: A Prospective, Dual-Center Study. *Radiology* **290**, 682–690 (2019).
79. Wilman, H. R. *et al.* Characterisation of liver fat in the UK Biobank cohort. *PLoS One* **12**, e0172921 (2017).
80. Gandon, Y. *et al.* Non-invasive assessment of hepatic iron stores by MRI. *Lancet* **363**, 357–362 (2004).
81. Wood, J. C. MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients. *Blood* **106**,

1460–1465 (2005).

82. St. Pierre, T. G. *et al.* Noninvasive measurement and imaging of liver iron concentrations using proton magnetic resonance. *Blood* **105**, 855–861 (2005).

83. St Pierre, T. G. *et al.* Multicenter validation of spin-density projection-assisted R2-MRI for the noninvasive measurement of liver iron concentration. *Magn. Reson. Med.* **71**, 2215–2223 (2014).

84. d'Assignies, G. *et al.* Non-invasive measurement of liver iron concentration using 3-Tesla magnetic resonance imaging: validation against biopsy. *Eur. Radiol.* **28**, 2022–2030 (2018).

85. European Association for the Study of the Liver. EASL clinical practice guidelines for HFE hemochromatosis. *J. Hepatol.* **53**, 3–22 (2010).

86. McKay, A. *et al.* Measurement of liver iron by magnetic resonance imaging in the UK Biobank population. *PLoS One* **13**, e0209340 (2018).

87. Ryan, J. D. *et al.* Hepatic iron is the major determinant of serum ferritin in NAFLD patients. *Liver Int.* **38**, 164–173 (2018).

88. Schwenzer, N. F. *et al.* T2* relaxometry in liver, pancreas, and spleen in a healthy cohort of one hundred twenty-nine subjects-correlation with age, gender, and serum ferritin. *Invest. Radiol.* **43**, 854–860 (2008).

89. Lin, H. *et al.* Quantitative susceptibility mapping in combination with water-fat separation for simultaneous liver iron and fat fraction quantification. *Eur. Radiol.* **28**, 3494–3504 (2018).

90. Akpinar, S., Yilmaz, G., Bulakbasi, N. & Kocaoglu, M. The role of diffusion-weighted imaging in prediction liver iron concentration in beta-thalassaemia patients. *J. Med. Imaging Radiat. Oncol.* **62**, 169–173 (2018).

91. Yin, M. *et al.* Assessment of Hepatic Fibrosis With Magnetic Resonance Elastography. *Clin. Gastroenterol. Hepatol.* **5**, 1207–1213 (2007).

92. Asbach, P. *et al.* Viscoelasticity-based staging of hepatic fibrosis with multifrequency MR elastography. *Radiology* **257**, 80–86 (2010).

93. Imajo, K. *et al.* Magnetic Resonance Imaging More Accurately Classifies Steatosis and Fibrosis in Patients With Nonalcoholic Fatty Liver Disease Than Transient Elastography. *Gastroenterology* **150**, 626-637.e7 (2016).

94. Park, C. C. *et al.* Magnetic Resonance Elastography vs Transient Elastography in

Detection of Fibrosis and Noninvasive Measurement of Steatosis in Patients With Biopsy-Proven Nonalcoholic Fatty Liver Disease. *Gastroenterology* **152**, 598-607.e2 (2017).

95. Chen, J. *et al.* Diagnostic Performance of MR Elastography and Vibration-controlled Transient Elastography in the Detection of Hepatic Fibrosis in Patients with Severe to Morbid Obesity. *Radiology* **283**, 418–428 (2017).
96. Caussy, C. *et al.* Association Between Obesity and Discordance in Fibrosis Stage Determination by Magnetic Resonance vs Transient Elastography in Patients With Nonalcoholic Liver Disease. *Clin. Gastroenterol. Hepatol.* **16**, 1974-1982.e7 (2018).
97. Ghoz, H. M. *et al.* Hepatic iron overload identified by magnetic resonance imaging-based T2* is a predictor of non-diagnostic elastography. *Quant. Imaging Med. Surg.* **9**, 921–927 (2019).
98. Serai, S. D. & Trout, A. T. Can MR elastography be used to measure liver stiffness in patients with iron overload? *Abdom. Radiol.* **44**, 104–109 (2019).
99. Gharib, A. M. *et al.* Magnetic Resonance Elastography Shear Wave Velocity Correlates with Liver Fibrosis and Hepatic Venous Pressure Gradient in Adults with Advanced Liver Disease. *Biomed Res. Int.* **2017**, 1–8 (2017).
100. Guo, J. *et al.* In vivo abdominal magnetic resonance elastography for the assessment of portal hypertension before and after transjugular intrahepatic portosystemic shunt implantation. *Invest. Radiol.* **50**, 347–51 (2015).
101. Yoon, H. *et al.* Predicting gastroesophageal varices through spleen magnetic resonance elastography in pediatric liver fibrosis. *World J. Gastroenterol.* **25**, 367–377 (2019).
102. Wagner, M. *et al.* Noninvasive prediction of portal pressure with MR elastography and DCE-MRI of the liver and spleen: Preliminary results. *J. Magn. Reson. Imaging* **48**, 1091–1103 (2018).
103. Cheng, Y. W., Chang, Y. C., Chen, Y. L., Chen, R. C. & Chou, C. Te. Feasibility of measuring spleen stiffness with MR elastography and splenic volume to predict hepatic fibrosis stage. *PLoS One* **14**, 1–13 (2019).
104. Ronot, M. *et al.* Assessment of portal hypertension and high-risk oesophageal varices with liver and spleen three-dimensional multifrequency MR elastography in liver cirrhosis. *Eur. Radiol.* **24**, 1394–402 (2014).
105. Shin, S. U. *et al.* Prediction of esophageal varices in patients with cirrhosis: Usefulness

of three-dimensional MR elastography with echo-planar imaging technique. *Radiology* **272**, 143–153 (2014).

106. Sun, H. Y., Lee, J. M., Han, J. K. & Choi, B. I. Usefulness of MR elastography for predicting esophageal varices in cirrhotic patients. *J. Magn. Reson. Imaging* **39**, 559–566 (2014).
107. Polasek, M. *et al.* Molecular MR imaging of liver fibrosis: A feasibility study using rat and mouse models. *J. Hepatol.* **57**, 549–555 (2012).
108. Fuchs, B. C. *et al.* Molecular MRI of collagen to diagnose and stage liver fibrosis. *J. Hepatol.* **59**, 992–998 (2013).
109. Farrar, C. T. *et al.* 3D molecular MR imaging of liver fibrosis and response to rapamycin therapy in a bile duct ligation rat model. *J. Hepatol.* **63**, 689–696 (2015).
110. Erstad, D. J. *et al.* Molecular magnetic resonance imaging accurately measures the antifibrotic effect of EDP-305, a novel farnesoid X receptor agonist. *Hepatol. Commun.* **2**, 821–835 (2018).
111. Zhu, B. *et al.* Combined magnetic resonance elastography and collagen molecular magnetic resonance imaging accurately stage liver fibrosis in a rat model. *Hepatology* **65**, 1015–1025 (2017).
112. Baik, S. K. Haemodynamic evaluation by Doppler ultrasonography in patients with portal hypertension: a review. *Liver Int.* **30**, 1403–1413 (2010).
113. Gouya, H. *et al.* Chronic liver disease: Systemic and splanchnic venous flow mapping with optimized cine phase-contrast MR imaging validated in a phantom model and prospectively evaluated in patients. *Radiology* **261**, 144–155 (2011).
114. Yzet, T. *et al.* Hepatic vascular flow measurements by phase contrast MRI and doppler echography: A comparative and reproducibility study. *J. Magn. Reson. Imaging* **31**, 579–588 (2010).
115. Gouya, H. *et al.* Portal hypertension in patients with cirrhosis: indirect assessment of hepatic venous pressure gradient by measuring azygos flow with 2D-cine phase-contrast magnetic resonance imaging. *Eur. Radiol.* **26**, 1981–1990 (2016).
116. McDonald, N. *et al.* Assessment of Haemodynamic Response to Nonselective Beta-Blockers in Portal Hypertension by Phase-Contrast Magnetic Resonance Angiography. *Biomed Res. Int.* **2017**, 1–8 (2017).
117. Palaniyappan, N. *et al.* Non-invasive assessment of portal hypertension using quantitative magnetic resonance imaging. *J. Hepatol.* **65**, 1131–1139 (2016).

118. Lee, C. *et al.* Diagnosis of Clinically Significant Portal Hypertension in Patients with Cirrhosis: Splenic Arterial Resistive Index versus Liver Stiffness Measurement. *Ultrasound Med. Biol.* **42**, 1312–1320 (2016).
119. Roldán-Alzate, A. *et al.* In vivo validation of 4D flow MRI for assessing the hemodynamics of portal hypertension. *J. Magn. Reson. Imaging* **37**, 1100–1108 (2013).
120. Bane, O. *et al.* Hemodynamic measurements with an abdominal 4D flow MRI sequence with spiral sampling and compressed sensing in patients with chronic liver disease. *J. Magn. Reson. Imaging* **49**, 994–1005 (2019).
121. Roldán-Alzate, A., Rutkowski, D. R., Fernandez, L. A. & Reeder, S. B. Longitudinal characterization of liver regeneration and portal haemodynamics in living donor liver transplant. *ISMRM* (2017).
122. Motosugi, U. *et al.* Four-dimensional Flow MRI as a marker for risk stratification of gastroesophageal varices in patients with liver cirrhosis. *Radiology* **290**, 101–107 (2019).
123. Stankovic, Z. *et al.* Effect of TIPS placement on portal and splanchnic arterial blood flow in 4-dimensional flow MRI. *Eur. Radiol.* **25**, 2634–2640 (2015).
124. Bannas, P. *et al.* Longitudinal Monitoring of Hepatic Blood Flow before and after TIPS by Using 4D-Flow MR Imaging. *Radiology* **281**, 574–582 (2016).
125. Koinuma, M., Ohashi, I., Hanafusa, K. & Shibuya, H. Apparent diffusion coefficient measurements with diffusion-weighted magnetic resonance imaging for evaluation of hepatic fibrosis. *J. Magn. Reson. Imaging* **22**, 80–85 (2005).
126. Boulanger, Y. *et al.* Diffusion-weighted MR imaging of the liver of hepatitis C patients. *NMR Biomed.* **16**, 132–136 (2003).
127. Bonekamp, S., Torbenson, M. S. & Kamel, I. R. Diffusion-weighted Magnetic Resonance Imaging for the Staging of Liver Fibrosis. *J. Clin. Gastroenterol.* **45**, 885–892 (2011).
128. Bakan, A. A., Inci, E., Bakan, S., Gokturk, S. & Cimilli, T. Utility of diffusion-weighted imaging in the evaluation of liver fibrosis. *Eur. Radiol.* **22**, 682–687 (2012).
129. Verloh, N. *et al.* DWI - histology: A possible means of determining degree of liver fibrosis? *Oncotarget* **9**, 20112–20118 (2018).
130. Bülow, R. *et al.* Diffusion-weighted magnetic resonance imaging for staging liver fibrosis is less reliable in the presence of fat and iron. *Eur. Radiol.* **23**, 1281–1287 (2013).
131. Manning, P. *et al.* Liver histology and diffusion-weighted MRI in children with

nonalcoholic fatty liver disease: A MAGNET study. *J. Magn. Reson. Imaging* **46**, 1149–1158 (2017).

132. Besheer, T. *et al.* Does steatosis affect the performance of diffusion-weighted MRI values for fibrosis evaluation in patients with chronic hepatitis C genotype 4? *Turkish J. Gastroenterol.* **28**, 283–288 (2017).
133. Dijkstra, H., Handayani, A., Kappert, P., Oudkerk, M. & Sijens, P. E. Clinical Implications of Non-Steatotic Hepatic Fat Fractions on Quantitative Diffusion-Weighted Imaging of the Liver. *PLoS One* **9**, e87926 (2014).
134. Chandarana, H. *et al.* The effect of liver iron deposition on hepatic apparent diffusion coefficient values in cirrhosis. *Am. J. Roentgenol.* **199**, 803–808 (2012).
135. Zou, L. Q., Chen, J., Pan, L., Jiang, J. Z. & Xing, W. Comparison of magnetic resonance elastography and diffusion-weighted imaging for staging hepatic fibrosis. *Chin. Med. J. (Engl).* **128**, 620–625 (2015).
136. Hennedige, T. P. *et al.* Magnetic resonance elastography and diffusion weighted imaging in the evaluation of hepatic fibrosis in chronic hepatitis B. *Gut Liver* **11**, 401–408 (2017).
137. Dyvorne, H. A. *et al.* Prospective comparison of magnetic resonance imaging to transient elastography and serum markers for liver fibrosis detection. *Liver Int.* **36**, 659–666 (2016).
138. Le Bihan, D. *et al.* Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* **168**, 497–505 (1988).
139. Hu, F. *et al.* Liver fibrosis: in vivo evaluation using intravoxel incoherent motion-derived histogram metrics with histopathologic findings at 3.0 T. *Abdom. Radiol.* **42**, 2855–2863 (2017).
140. Sandrasegaran, K. *et al.* Does intravoxel incoherent motion reliably stage hepatic fibrosis, steatosis, and inflammation? *Abdom. Radiol.* **43**, 600–606 (2018).
141. Ichikawa, S. *et al.* MRI-based staging of hepatic fibrosis: Comparison of intravoxel incoherent motion diffusion-weighted imaging with magnetic resonance elastography. *J. Magn. Reson. Imaging* **42**, 204–210 (2015).
142. Zhang, J. *et al.* MRI-based estimation of liver function by intravoxel incoherent motion diffusion-weighted imaging. *Magn Reson Imaging* **34**, 1220–1225 (2016).
143. Li, Y. T. *et al.* Liver intravoxel incoherent motion (IVIM) magnetic resonance imaging: a

comprehensive review of published data on normal values and applications for fibrosis and tumor evaluation. *Quant. Imaging Med. Surg.* **7**, 59–78 (2017).

144. Fu, F. *et al.* Non-invasive assessment of hepatic fibrosis: comparison of MR elastography to transient elastography and intravoxel incoherent motion diffusion-weighted MRI. *Abdom. Radiol.* **45**, 73–82 (2020).
145. Ramasawmy, R. *et al.* Hepatic arterial spin labelling MRI: An initial evaluation in mice. *NMR Biomed.* **28**, 272–280 (2015).
146. Byk, K. *et al.* MRI-based assessment of liver perfusion and hepatocyte injury in the murine model of acute hepatitis. *Magn. Reson. Mater. Physics, Biol. Med.* **29**, 789–798 (2016).
147. Pan, X. *et al.* Quantification of liver perfusion using multidelay pseudocontinuous arterial spin labeling. *J. Magn. Reson. Imaging* **43**, 1046–1054 (2016).
148. Martirosian, P. *et al.* Spatial-temporal perfusion patterns of the human liver assessed by pseudo-continuous arterial spin labeling MRI. *Z. Med. Phys.* **29**, 173–183 (2019).
149. Schalkx, H. J. *et al.* Arterial and portal venous liver perfusion using selective spin labelling MRI. *Eur. Radiol.* **25**, 1529–1540 (2015).
150. Cox, E. F., Palaniyappan, N., Aithal, G. P., Guha, I. N. & Francis, S. T. Using MRI to study the alterations in liver blood flow, perfusion, and oxygenation in response to physiological stress challenges: Meal, hyperoxia, and hypercapnia. *J. Magn. Reson. Imaging* **49**, 1577–1586 (2019).
151. Look, D. C. & Locker, D. R. Time Saving in Measurement of NMR and EPR Relaxation Times. *Rev. Sci. Instrum.* **41**, 250–251 (1970).
152. Smith, F. W., Mallard, J. R., Reid, A. & Hutchison, J. M. S. NUCLEAR MAGNETIC RESONANCE TOMOGRAPHIC IMAGING IN LIVER DISEASE. *Lancet* **317**, 963–966 (1981).
153. Doyle, F. *et al.* Nuclear magnetic resonance imaging of the liver: initial experience. *Am. J. Roentgenol.* **138**, 193–200 (1982).
154. Keevil, S. F. *et al.* Non-invasive assessment of diffuse liver disease by in vivo measurement of proton nuclear magnetic resonance relaxation times at 0.08 T. *Br. J. Radiol.* **67**, 1083–1087 (1994).
155. Thomsen, C., Christoffersen, P., Henriksen, O. & Juhl, E. Prolonged T1 in patients with liver cirrhosis: An in vivo MRI study. *Magn. Reson. Imaging* **8**, 599–604 (1990).
156. The Clinical NMR Group. Magnetic resonance imaging of parenchymal liver disease: a

comparison with ultrasound, radionuclide scintigraphy and X-ray computed tomography. *Clin. Radiol.* **38**, 495–502 (1987).

157. Kim, K. A. *et al.* Quantitative evaluation of liver cirrhosis using T1 relaxation time with 3 tesla MRI before and after oxygen inhalation. *J. Magn. Reson. Imaging* **36**, 405–410 (2012).
158. Heye, T. *et al.* MR relaxometry of the liver: significant elevation of T1 relaxation time in patients with liver cirrhosis. *Eur. Radiol.* **22**, 1224–1232 (2012).
159. Goldberg, H. I. *et al.* Hepatic cirrhosis: magnetic resonance imaging. *Radiology* **153**, 737–739 (1984).
160. Aisen, A. M., Doi, K. & Swanson, S. D. Liver Fibrosis with Magnetic Cross-Relaxation. *Magn. Reson. Med.* **31**, 551–556 (1994).
161. Chamuleau, R. A. F. M., De Nie, J. H. N. C. I., Moerland, M. A., Van der Lende, O. R. & Smidt, J. Is the magnetic resonance imaging proton spin-lattice relaxation time a reliable noninvasive parameter of developing liver fibrosis? *Hepatology* **8**, 217–221 (1988).
162. Cassinotto, C. *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.* **84**, 1459–1465 (2015).
163. Luetkens, J. A. *et al.* Quantification of liver fibrosis at T1 and T2 mapping with extracellular volume fraction MRI: Preclinical results. *Radiology* **288**, 748–754 (2018).
164. Hoad, C. L. *et al.* A study of T_1 relaxation time as a measure of liver fibrosis and the influence of confounding histological factors. *NMR Biomed.* **28**, 706–14 (2015).
165. Obmann, V. C. *et al.* Liver MR relaxometry at 3T – segmental normal T1 and T2* values in patients without focal or diffuse liver disease and in patients with increased liver fat and elevated liver stiffness. *Sci. Rep.* **9**, 1–11 (2019).
166. Allkemper, T. *et al.* Evaluation of Fibrotic Liver Disease with Whole-Liver T1ρ MR Imaging: A Feasibility Study at 1.5 T. *Radiology* **271**, 408–415 (2014).
167. Chen, W. *et al.* Quantitative assessment of liver function with whole-liver T1rho mapping at 3.0 T. *Magn. Reson. Imaging* **46**, 75–80 (2018).
168. Singh, A. *et al.* T1ρ MRI of healthy and fibrotic human livers at 1.5 T. *J. Transl. Med.* **13**, 292 (2015).
169. Xie, S. *et al.* Impact of liver fibrosis and fatty liver on T1rho measurements: A

prospective study. *Korean J. Radiol.* **18**, 898–905 (2017).

170. Zhao, F. *et al.* MR T1ρ as an imaging biomarker for monitoring liver injury progression and regression: an experimental study in rats with carbon tetrachloride intoxication. *Eur. Radiol.* **22**, 1709–1716 (2012).
171. Zhang, H. *et al.* Comparison of T2, T1rho, and diffusion metrics in assessment of liver fibrosis in rats. *J. Magn. Reson. Imaging* **45**, 741–750 (2017).
172. Juluru, K. *et al.* Diagnostic accuracy of intracellular uptake rates calculated using dynamic Gd-EOB-DTPA-enhanced MRI for hepatic fibrosis stage. *J. Magn. Reson. Imaging* **45**, 1177–1185 (2017).
173. Norén, B. *et al.* Separation of advanced from mild hepatic fibrosis by quantification of the hepatobiliary uptake of Gd-EOB-DTPA. *Eur. Radiol.* **23**, 174–181 (2013).
174. Hako, R. *et al.* Noninvasive assessment of liver fibrosis in patients with chronic hepatitis b or c by contrast-enhanced magnetic resonance imaging. *Can. J. Gastroenterol. Hepatol.* **2019**, 1–9 (2019).
175. Ning, J. *et al.* Hepatic function imaging using dynamic Gd-EOB-DTPA enhanced MRI and pharmacokinetic modeling. *Magn. Reson. Med.* **78**, 1488–1495 (2017).
176. Haimerl, M. *et al.* Volume-assisted estimation of liver function based on Gd-EOB-DTPA–enhanced MR relaxometry. *Eur. Radiol.* **26**, 1125–1133 (2016).
177. Karageorgis, A. *et al.* A multi-center preclinical study of gadoxetate DCE-MRI in rats as a biomarker of drug induced inhibition of liver transporter function. *PLoS One* **13**, 1–18 (2018).
178. Huh, J. *et al.* Gadoxetate-enhanced dynamic contrast-enhanced MRI for evaluation of liver function and liver fibrosis in preclinical trials. *BMC Med. Imaging* **19**, 89 (2019).
179. Zhou, L. *et al.* Liver dynamic contrast-enhanced MRI for staging liver fibrosis in a piglet model. *J. Magn. Reson. Imaging* **39**, 872–878 (2014).
180. Keller, S. *et al.* Prospective comparison of diffusion-weighted MRI and dynamic Gd-EOB-DTPA-enhanced MRI for detection and staging of hepatic fibrosis in primary sclerosing cholangitis. *Eur. Radiol.* **29**, 818–828 (2019).
181. Cannella, R., Borhani, A. A., Tublin, M., Behari, J. & Furlan, A. Diagnostic value of MR-based texture analysis for the assessment of hepatic fibrosis in patients with nonalcoholic fatty liver disease (NAFLD). *Abdom. Radiol.* **44**, 1816–1824 (2019).
182. Chen, Y. *et al.* MR Fingerprinting for Rapid Quantitative Abdominal Imaging. *Radiology*

279, 278–286 (2016).

183. Feier, D. *et al.* 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. *Eur. Radiol.* **26**, 539–546 (2016).
184. Bradley, C. R. *et al.* Multi-organ assessment of compensated cirrhosis patients using quantitative magnetic resonance imaging. *J. Hepatol.* **69**, 1015–1024 (2018).
185. Bradley, C. *et al.* Short-term changes observed in multiparametric liver MRI following therapy with direct-acting antivirals in chronic hepatitis C virus patients. *Eur. Radiol.* **29**, 3100–3107 (2019).
186. Tunnicliffe, E. M., Banerjee, R., Pavlides, M., Neubauer, S. & Robson, M. D. A model for hepatic fibrosis: the competing effects of cell loss and iron on shortened modified Look-Locker inversion recovery T1(shMOLLI-T1) in the liver. *J. Magn. Reson. Imaging* **45**, 450–462 (2017).
187. Pavlides, M. *et al.* Multiparametric magnetic resonance imaging predicts clinical outcomes in patients with chronic liver disease. *J. Hepatol.* **64**, 308–15 (2016).
188. Bachtiar, V. *et al.* Repeatability and reproducibility of multiparametric magnetic resonance imaging of the liver. *PLoS One* **14**, e0214921 (2019).
189. Mojtahed, A. *et al.* Reference range of liver corrected T1 values in a population at low risk for fatty liver disease—a UK Biobank sub-study, with an appendix of interesting cases. *Abdom. Radiol.* **44**, 72–84 (2018).
190. Harrison, S. A. *et al.* NGM282 Improves Liver Fibrosis and Histology in 12 Weeks in Patients With Nonalcoholic Steatohepatitis. *Hepatology* **0**, 1–15 (2019).
191. Pavlides, M. *et al.* Multiparametric magnetic resonance imaging for the assessment of non-alcoholic fatty liver disease severity. *Liver Int.* **37**, 1065–1073 (2017).
192. Mole, D. J. *et al.* Study protocol: HepaT1ca – an observational clinical cohort study to quantify liver health in surgical candidates for liver malignancies. *BMC Cancer* **18**, 890 (2018).
193. Levick, C. *et al.* Non-invasive assessment of portal hypertension by multi-parametric magnetic resonance imaging of the spleen: A proof of concept study. *PLoS One* **14**, e0221066 (2019).
194. Sharp, K. & Matschinsky, F. Translation of Ludwig Boltzmann's Paper "On the Relationship between the Second Fundamental Theorem of the Mechanical Theory of

Heat and Probability Calculations Regarding the Conditions for Thermal Equilibrium" Sitzungberichte der Kaiserlichen Akademie d. *Entropy* **17**, 1971–2009 (2015).

195. Roujol, S. *et al.* Accuracy, precision, and reproducibility of four T1 mapping sequences: a head-to-head comparison of MOLLI, ShMOLLI, SASHA, and SAPPHIRE. *Radiology* **272**, 683–9 (2014).
196. Kellman, P. & Hansen, M. S. T1-mapping in the heart: accuracy and precision. *J. Cardiovasc. Magn. Reson.* **16**, 2 (2014).
197. Piechnik, S. K. *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.* **12**, 69 (2010).
198. Bernstein, M. A., King, K. F. & Zhou, X. J. *Handbook of MRI Pulse Sequences*. *Handbook of MRI Pulse Sequences* (2004). doi:10.1016/B978-0-12-092861-3.X5000-6
199. Mozes, F. E. Magnetic resonance imaging of the liver. T1 dynamics: confounders and modelling. (2018).
200. Messroghli, D. R. *et al.* Modified Look-Locker inversion recovery (MOLLI) for high-resolutionT1 mapping of the heart. *Magn. Reson. Med.* **52**, 141–146 (2004).
201. Reeder, S. B. *et al.* Iterative decomposition of water and fat with echo asymmetry and least-squares estimation (IDEAL): Application with fast spin-echo imaging. *Magn. Reson. Med.* **54**, 636–644 (2005).
202. Robson, M. D., Piechnik, S. K., Tunnicliffe, E. M. & Neubauer, S. T 1 measurements in the human myocardium: The effects of magnetization transfer on the SASHA and MOLLI sequences. *Magn. Reson. Med.* **70**, 664–670 (2013).
203. Banerjee, R. *et al.* Multiparametric magnetic resonance for the non-invasive diagnosis of liver disease. *J. Hepatol.* **60**, 69–77 (2014).
204. Frahm, J., Merboldt, K. ., Hänicke, W. & Haase, A. Stimulated echo imaging. *J. Magn. Reson.* **64**, 81–93 (1985).
205. Purvis, L. A. B. *et al.* OXSA: An open-source magnetic resonance spectroscopy analysis toolbox in MATLAB. *PLoS One* **12**, (2017).
206. Vanhamme, L., van den Boogaart, A. & Van Huffel, S. Improved Method for Accurate and Efficient Quantification of MRS Data with Use of Prior Knowledge. *J. Magn. Reson.* **129**, 35–43 (1997).
207. Rial, B., Robson, M. D., Neubauer, S. & Schneider, J. E. Rapid quantification of

myocardial lipid content in humans using single breath-hold 1H MRS at 3 Tesla. *Magn. Reson. Med.* **66**, 619–24 (2011).

208. Reeder, S. B., Cruite, I., Hamilton, G. & Sirlin, C. B. Quantitative assessment of liver fat with magnetic resonance imaging and spectroscopy. *J. Magn. Reson. Imaging* **34**, 729–749 (2011).
209. Sharma, P. *et al.* Quantitative analysis of T2-correction in single-voxel magnetic resonance spectroscopy of hepatic lipid fraction. *J. Magn. Reson. Imaging* **29**, 629–635 (2009).
210. Wood, J. C. *et al.* Cardiac Iron Determines Cardiac T2*, T2, and T1 in the Gerbil Model of Iron Cardiomyopathy. *Circulation* **112**, 535–543 (2005).
211. Ghugre, N. R. *et al.* Multi-field behavior of Relaxivity in an Iron-rich environment. *Proc. Intl. Soc. Mag. Reson. Med.* **16**, (2008).
212. Mozes, F. E., Valkovič, L., Pavlides, M., Robson, M. D. & Tunnicliffe, E. M. Liver glycogen concentration and body hydration status are predictors of liver shMOLLI T1 measurements. *ISMRM* (2019).
213. Sandrin, L. *et al.* Transient elastography: A new noninvasive method for assessment of hepatic fibrosis. *Ultrasound Med. Biol.* **29**, 1705–1713 (2003).
214. Klibansky, D. A. *et al.* Transient elastography for predicting clinical outcomes in patients with chronic liver disease. *J. Viral Hepat.* **19**, e184–e193 (2012).
215. Pang, J. X. Q. *et al.* Liver Stiffness by Transient Elastography Predicts Liver-Related Complications and Mortality in Patients with Chronic Liver Disease. *PLoS One* **9**, e95776 (2014).
216. Arena, U. *et al.* Acute viral hepatitis increases liver stiffness values measured by transient elastography. *Hepatology* **47**, 380–384 (2007).
217. Millonig, G. *et al.* Extrahepatic cholestasis increases liver stiffness (FibroScan) irrespective of fibrosis. *Hepatology* **48**, 1718–1723 (2008).
218. Berzigotti, A. *et al.* Effect of Meal Ingestion on Liver Stiffness in Patients with Cirrhosis and Portal Hypertension. *PLoS One* (2013). doi:10.1371/journal.pone.0058742
219. Mederacke, I. *et al.* Food intake increases liver stiffness in patients with chronic or resolved hepatitis C virus infection. *Liver Int.* (2009). doi:10.1111/j.1478-3231.2009.02100.x
220. Kettaneh, A. *et al.* Features associated with success rate and performance of fibroscan

measurements for the diagnosis of cirrhosis in HCV patients: A prospective study of 935 patients. *J. Hepatol.* **46**, 628–634 (2007).

221. R Core Team. R: A language and environment for statistical computing. (2018).
222. Angulo, P. *et al.* Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology* **149**, 389-397.e10 (2015).
223. Ostovaneh, M. R. *et al.* Association of Liver Fibrosis with Cardiovascular Diseases in the General Population: The Multi-Ethnic Study of Atherosclerosis (MESA). *Circ. Cardiovasc. Imaging* **11**, 1–10 (2018).
224. Nuttall, K. L., Palaty, J. & Lockitch, G. Reference limits for copper and iron in liver biopsies. *Ann. Clin. Lab. Sci.* **33**, 443–50 (2003).
225. Szczepaniak, L. S. *et al.* Magnetic resonance spectroscopy to measure hepatic triglyceride content: prevalence of hepatic steatosis in the general population. *Am. J. Physiol. Metab.* **288**, E462–E468 (2005).
226. Fraquelli, M. *et al.* Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. *Gut* **56**, 968–973 (2007).
227. Schuetz, G. M., Schlattmann, P. & Dewey, M. Use of 3×2 tables with an intention to diagnose approach to assess clinical performance of diagnostic tests: Meta-analytical evaluation of coronary CT angiography studies. *BMJ* **345**, 1–10 (2012).
228. Westin, J., Nordlinder, H., Lagging, M., Norkrans, G. & Wejstål, R. Steatosis accelerates fibrosis development over time in hepatitis C virus genotype 3 infected patients. *J. Hepatol.* **37**, 837–842 (2002).
229. McDonald, N. *et al.* Multiparametric magnetic resonance imaging for quantitation of liver disease: a two-centre cross-sectional observational study. *Sci. Rep.* **8**, 9189 (2018).
230. Harrison, S. A. *et al.* Utility and variability of three non-invasive liver fibrosis imaging modalities to evaluate efficacy of GR-MD-02 in subjects with NASH and bridging fibrosis during a phase-2 randomized clinical trial. *PLoS One* **13**, e0203054 (2018).
231. Eddowes, P. J. *et al.* Utility and cost evaluation of multiparametric magnetic resonance imaging for the assessment of non-alcoholic fatty liver disease. *Aliment. Pharmacol. Ther.* **47**, 631–644 (2018).
232. Blake, L., Duarte, R. V & Cummins, C. Decision analytic model of the diagnostic pathways for patients with suspected non-alcoholic fatty liver disease using non-

invasive transient elastography and multiparametric magnetic resonance imaging. *BMJ Open* **6**, e010507 (2016).

233. Boursier, J. *et al.* Diagnostic accuracy and prognostic significance of blood fibrosis tests and liver stiffness measurement by FibroScan in non-alcoholic fatty liver disease. *J. Hepatol.* **65**, 570–578 (2016).
234. Sebastiani, G. *et al.* Prognostic Value of Non-Invasive Fibrosis and Steatosis Tools, Hepatic Venous Pressure Gradient (HVPG) and Histology in Nonalcoholic Steatohepatitis. *PLoS One* **10**, e0128774 (2015).
235. Boursier, J. *et al.* Combination of blood tests for significant fibrosis and cirrhosis improves the assessment of liver-prognosis in chronic hepatitis C. *Aliment. Pharmacol. Ther.* **40**, 178–188 (2014).
236. Ferenci, P. *et al.* Real-world safety and effectiveness of ombitasvir/paritaprevir/ritonavir ± dasabuvir ± ribavirin in hepatitis C virus genotype 1- and 4-infected patients with diverse comorbidities and comedications: A pooled analysis of post-marketing observational studi. *J. Viral Hepat.* **26**, 685–696 (2019).
237. Zeuzem, S. *et al.* JNJ-4178 (AL-335, Odalasvir, and Simeprevir) for 6 or 8 Weeks in Hepatitis C Virus-Infected Patients Without Cirrhosis: OMEGA-1. *Hepatology* **69**, 2349–2363 (2019).
238. Foster, G. R. *et al.* Impact of direct acting antiviral therapy in patients with chronic hepatitis C and decompensated cirrhosis. *J. Hepatol.* **64**, 1224–1231 (2016).
239. Reddy, K. R. *et al.* Ledipasvir and sofosbuvir in patients with genotype 1 hepatitis C virus infection and compensated cirrhosis: An integrated safety and efficacy analysis. *Hepatology* **62**, 79–86 (2015).
240. Lawitz, E. *et al.* Efficacy and safety of 12 weeks versus 18 weeks of treatment with grazoprevir (MK-5172) and elbasvir (MK-8742) with or without ribavirin for hepatitis C virus genotype 1 infection in previously untreated patients with cirrhosis and patients with previous . *Lancet* **385**, 1075–1086 (2015).
241. Fontana, R. J. *et al.* Efficacy and Safety of 8 Weeks of Glecaprevir/Pibrentasvir in Treatment-Naïve, HCV-Infected Patients with APRI ≤ 1 in a Single-Arm, Open-Label, Multicenter Study. *Adv. Ther.* **36**, 3458–3470 (2019).
242. De Lédinghen, V. *et al.* Retreatment with Sofosbuvir Plus Grazoprevir/Elbasvir Plus Ribavirin of Patients with Hepatitis C Virus Genotype 1 or 4 Who Previously Failed an

NS5A-or NS3-Containing Regimen: The ANRS HC34 REVENGE Study. *Clin. Infect. Dis.* **66**, 1013–1018 (2018).

243. Llaneras, J. *et al.* Effectiveness and safety of sofosbuvir/velpatasvir/voxilaprevir in patients with chronic hepatitis C previously treated with DAAs. *J. Hepatol.* **71**, 666–672 (2019).

244. Esteban, R. *et al.* Efficacy of Sofosbuvir and Velpatasvir, With and Without Ribavirin, in Patients With Hepatitis C Virus Genotype 3 Infection and Cirrhosis. *Gastroenterology* **155**, 1120-1127.e4 (2018).

245. Brown, R. S. *et al.* Glecaprevir/pibrentasvir for 8 weeks in treatment-naïve patients with chronic HCV genotypes 1–6 and compensated cirrhosis: The EXPEDITION-8 trial. *J. Hepatol.* (2019). doi:10.1016/j.jhep.2019.10.020

246. Puoti, M. *et al.* High SVR12 with 8-week and 12-week glecaprevir/pibrentasvir therapy: An integrated analysis of HCV genotype 1–6 patients without cirrhosis. *J. Hepatol.* **69**, 293–300 (2018).

247. Toyoda, H. *et al.* Real-world experience of 12-week DAA regimen of glecaprevir and pibrentasvir in patients with chronic HCV infection. *J. Gastroenterol. Hepatol.* **24**, (2019).

248. Toyoda, H. *et al.* Trends and Efficacy of Interferon-Free Anti–hepatitis C Virus Therapy in the Region of High Prevalence of Elderly Patients, Cirrhosis, and Hepatocellular Carcinoma: A Real-World, Nationwide, Multicenter Study of 10 688 Patients in Japan. *Open Forum Infect. Dis.* **6**, (2019).

249. Piñero, F. *et al.* Treatment with direct-acting antivirals for HCV decreases but does not eliminate the risk of hepatocellular carcinoma. *Liver Int.* **39**, 1033–1043 (2019).

250. Waziry, R. *et al.* Hepatocellular carcinoma risk following direct-acting antiviral HCV therapy: A systematic review, meta-analyses, and meta-regression. *J. Hepatol.* **67**, 1204–1212 (2017).

251. Nahon, P. *et al.* Eradication of Hepatitis C Virus Infection in Patients With Cirrhosis Reduces Risk of Liver and Non-Liver Complications. *Gastroenterology* **152**, 142-156.e2 (2017).

252. Ioannou, G. N. *et al.* Increased Risk for Hepatocellular Carcinoma Persists Up to 10 Years After HCV Eradication in Patients With Baseline Cirrhosis or High FIB-4 Scores. *Gastroenterology* **157**, 1264-1278.e4 (2019).

253. Ioannou, G. N. *et al.* Development of models estimating the risk of hepatocellular carcinoma after antiviral treatment for hepatitis C. *J. Hepatol.* **69**, 1088–1098 (2018).
254. Bachofner, J. A. *et al.* Direct antiviral agent treatment of chronic hepatitis C results in rapid regression of transient elastography and fibrosis markers fibrosis-4 score and aspartate aminotransferase-platelet ratio index. *Liver Int.* **37**, 369–376 (2017).
255. Dolmazashvili, E. *et al.* Regression of liver fibrosis over a 24-week period after completing direct-acting antiviral therapy in patients with chronic hepatitis C receiving care within the national hepatitis C elimination program in Georgia: results of hepatology clinic HEPA exper. *Eur. J. Gastroenterol. Hepatol.* **29**, 1223–1230 (2017).
256. Knop, V. *et al.* Regression of fibrosis and portal hypertension in HCV-associated cirrhosis and sustained virologic response after interferon-free antiviral therapy. *J. Viral Hepat.* **23**, 994–1002 (2016).
257. Chan, J. *et al.* Direct-Acting Antiviral Therapy for Chronic HCV Infection Results in Liver Stiffness Regression Over 12 Months Post-treatment. *Dig. Dis. Sci.* **63**, 486–492 (2018).
258. Elsharkawy, A. *et al.* Changes in liver stiffness measurements and fibrosis scores following sofosbuvir based treatment regimens without interferon. *J. Gastroenterol. Hepatol.* **32**, 1624–1630 (2017).
259. Laursen, T. L. *et al.* Time-dependent improvement of liver inflammation, fibrosis and metabolic liver function after successful direct-acting antiviral therapy of chronic hepatitis C. *J. Viral Hepat.* **27**, 28–35 (2020).
260. Stasi, C. *et al.* Longitudinal evaluation of liver stiffness and outcomes in patients with chronic hepatitis C before and after short- and long-term IFN-free antiviral treatment. *Curr. Med. Res. Opin.* **0**, 1–1 (2019).
261. Pan, J. J. *et al.* Morphometry Confirms Fibrosis Regression From Sustained Virologic Response to Direct-Acting Antivirals for Hepatitis C. *Hepatol. Commun.* **2**, 1320–1330 (2018).
262. Hamdane, N. *et al.* HCV-Induced Epigenetic Changes Associated With Liver Cancer Risk Persist After Sustained Virologic Response. *Gastroenterology* **156**, 2313-2329.e7 (2019).
263. Perez, S. *et al.* Hepatitis C virus leaves an epigenetic signature post cure of infection by direct-acting antivirals. *PLOS Genet.* **15**, e1008181 (2019).
264. Marrero, J. A. *et al.* α-Fetoprotein, Des-γ Carboxyprothrombin, and Lectin-Bound α-

Fetoprotein in Early Hepatocellular Carcinoma. *Gastroenterology* **137**, 110–118 (2009).

265. Lok, A. S. *et al.* Des-γ-Carboxy Prothrombin and α-Fetoprotein as Biomarkers for the Early Detection of Hepatocellular Carcinoma. *Gastroenterology* **138**, 493–502 (2010).
266. Sterling, R. K. *et al.* Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* **43**, 1317–1325 (2006).
267. Wai, C. T. *et al.* A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* **38**, 518–526 (2003).
268. Poynard, T. *et al.* Slow regression of liver fibrosis presumed by repeated biomarkers after virological cure in patients with chronic hepatitis C. *J. Hepatol.* **59**, 675–683 (2013).
269. Mozes, F. E., Tunnicliffe, E. M., Pavlides, M. & Robson, M. D. Influence of fat on liver T1 measurements using modified Look-Locker inversion recovery (MOLLI) methods at 3T. *J. Magn. Reson. Imaging* **44**, 105–11 (2016).
270. Eaton, J. E. *et al.* Changes in Liver Stiffness, Measured by Magnetic Resonance Elastography, Associated With Hepatic Decompensation in Patients With Primary Sclerosing Cholangitis. *Clin. Gastroenterol. Hepatol.* (2019). doi:10.1016/j.cgh.2019.10.041
271. Ziol, M. *et al.* Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with chronic hepatitis C. *Hepatology* **41**, 48–54 (2005).
272. Castéra, L. *et al.* Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. *Gastroenterology* **128**, 343–350 (2005).
273. D'Ambrosio, R. *et al.* The diagnostic accuracy of Fibroscan® for cirrhosis is influenced by liver morphometry in HCV patients with a sustained virological response. *J. Hepatol.* **59**, 251–256 (2013).
274. Knop, V. *et al.* Regression of fibrosis and portal hypertension in HCV-associated cirrhosis and sustained virologic response after interferon-free antiviral therapy. *J. Viral Hepat.* **23**, 994–1002 (2016).
275. Guo, J. *et al.* Treatment response of hepatic and renal stiffness in patients with chronic HCV infection monitored by multifrequency MR elastography. *Proc. Intl. Soc. Mag. Reson. Med.* **25**, (2017).
276. Jayakumar, S. *et al.* Longitudinal correlations between MRE, MRI-PDFF, and liver

histology in patients with non-alcoholic steatohepatitis: Analysis of data from a phase II trial of selonsertib. *J. Hepatol.* **70**, 133–141 (2019).

277. Poynard, T. *et al.* Impact of pegylated interferon alfa-2b and ribavirin on liver fibrosis in patients with chronic hepatitis C. *Gastroenterology* **122**, 1303–13 (2002).
278. Pekow, J. R., Bhan, A. K., Zheng, H. & Chung, R. T. Hepatic steatosis is associated with increased frequency of hepatocellular carcinoma in patients with hepatitis C-related cirrhosis. *Cancer* **109**, 2490–2496 (2007).
279. Ganne-Carrié, N. *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* **46**, 277–282 (2000).
280. Ioannou, G. N., Dominitz, J. A., Weiss, N. S., Heagerty, P. J. & Kowdley, K. V. The effect of alcohol consumption on the prevalence of iron overload, iron deficiency, and iron deficiency anemia. *Gastroenterology* **126**, 1293–1301 (2004).
281. Gronbaek, K. *et al.* Interobserver variation in interpretation of serial liver biopsies from patients with chronic hepatitis C. *J. Viral Hepat.* **9**, 443–449 (2002).
282. Robert, M. *et al.* A Comparison of Hepatopathologists' and Community Pathologists' Review of Liver Biopsy Specimens From Patients With Hepatitis C. *Clin. Gastroenterol. Hepatol.* **7**, 335–338 (2009).
283. Sun, Y. *et al.* New classification of liver biopsy assessment for fibrosis in chronic hepatitis B patients before and after treatment. *Hepatology* **65**, 1438–1450 (2017).
284. Kovalak, M. *et al.* Endoscopic screening for varices in cirrhotic patients: data from a national endoscopic database. *Gastrointest. Endosc.* **65**, 82–88 (2007).
285. Amitrano, L. *et al.* The Effectiveness of Current Acute Variceal Bleed Treatments in Unselected Cirrhotic Patients: Refining Short-Term Prognosis and Risk Factors. *Am. J. Gastroenterol.* **107**, 1872–1878 (2012).
286. Perello, A. *et al.* Wedged hepatic venous pressure adequately reflects portal pressure in hepatitis C virus-related cirrhosis. *Hepatology* **30**, 1393–1397 (1999).
287. Garcia-Tsao, G. *et al.* Portal pressure, presence of gastroesophageal varices and variceal bleeding. *Hepatology* **5**, 419–24 (1985).
288. Abraldes, J. G. *et al.* Hepatic venous pressure gradient and prognosis in patients with acute variceal bleeding treated with pharmacologic and endoscopic therapy. *J. Hepatol.* **48**, 229–236 (2008).

289. Moitinho, E. *et al.* Prognostic value of early measurements of portal pressure in acute variceal bleeding. *Gastroenterology* **117**, 626–631 (1999).

290. Rincon, D. *et al.* Prognostic value of hepatic venous pressure gradient for in-hospital mortality of patients with severe acute alcoholic hepatitis. *Aliment. Pharmacol. Ther.* **25**, 841–848 (2007).

291. Groszmann, R. J. *et al.* Hemodynamic events in a prospective randomized trial of propranolol versus placebo in the prevention of a first variceal hemorrhage. *Gastroenterology* **99**, 1401–1407 (1990).

292. Feu, F. *et al.* Relation between portal pressure response to pharmacotherapy and risk of recurrent variceal haemorrhage in patients with cirrhosis. *Lancet* **346**, 1056–1059 (1995).

293. Turnes, J. *et al.* Pharmacological Reduction of Portal Pressure and Long-Term Risk of First Variceal Bleeding in Patients with Cirrhosis. *Am. J. Gastroenterol.* **101**, 506–512 (2006).

294. Raju, S. A., Marks, L., Ching, H.-L. & McAlindon, M. E. PTU-015 Patient perspective on the tolerance of gastroscopy. *Gut* **66**, A58 LP-A58 (2017).

295. Szakács, Z. *et al.* Baveno criteria safely identify patients with compensated advanced chronic liver disease who can avoid variceal screening endoscopy: A diagnostic test accuracy meta-analysis. *Front. Physiol.* **10**, 1–14 (2019).

296. Ying, L., Lin, X., Xie, Z.-L., Hu, Y.-P. & Shi, K.-Q. Performance of platelet count/spleen diameter ratio for diagnosis of esophageal varices in cirrhosis: a meta-analysis. *Dig. Dis. Sci.* **57**, 1672–81 (2012).

297. Berzigotti, A. *et al.* Elastography, spleen size, and platelet count identify portal hypertension in patients with compensated cirrhosis. *Gastroenterology* **144**, 102-111.e1 (2013).

298. Buechter, M. *et al.* Spleen stiffness is positively correlated with HVPG and decreases significantly after TIPS implantation. *Dig. Liver Dis.* **50**, 54–60 (2018).

299. Gharib, A. M. *et al.* Magnetic Resonance Elastography Shear Wave Velocity Correlates with Liver Fibrosis and Hepatic Venous Pressure Gradient in Adults with Advanced Liver Disease. *Biomed Res. Int.* **2017**, 1–8 (2017).

300. Vizzutti, F. *et al.* Liver stiffness measurement predicts severe portal hypertension in patients with HCV-related cirrhosis. *Hepatology* **45**, 1290–1297 (2007).

301. Levick, C. *et al.* PS-020-Spleen T1 and spleen diameter criteria can identify and exclude oesophageal varices accurately. *J. Hepatol.* **70**, e15 (2019).
302. Giuffrè, M. *et al.* Evaluation of spleen stiffness in healthy volunteers using point shear wave elastography. *Ann. Hepatol.* **18**, 736–741 (2019).
303. Balakrishnan, M. *et al.* Liver and Spleen Stiffness Measurements by Point Shear Wave Elastography via Acoustic Radiation Force Impulse. *J. Ultrasound Med.* **35**, 2373–2380 (2016).
304. Yasar, T. K. *et al.* Interplatform reproducibility of liver and spleen stiffness measured with MR elastography. *J. Magn. Reson. Imaging* **43**, 1064–1072 (2016).
305. Afdhal, N. *et al.* Thrombocytopenia associated with chronic liver disease. *J. Hepatol.* **48**, 1000–1007 (2008).
306. Gana, J. C. *et al.* Platelet count, spleen length, and platelet count/spleen length ratio for the diagnosis of oesophageal varices in patients with chronic liver disease or portal vein thrombosis. in *Cochrane Database of Systematic Reviews* (ed. Gana, J. C.) 2 (John Wiley & Sons, Ltd, 2010). doi:10.1002/14651858.CD008759
307. Zhu, J., Leng, X., Dong, N., Qi, G. & Du, R. Measurement of liver volume and its clinical significance in cirrhotic patients. *Zhonghua Wai Ke Za Zhi* **37**, 110–112 (1999).
308. Tey, T. *et al.* Application of a standardised protocol for hepatic venous pressure gradient measurement improves quality of readings and facilitates reduction of variceal bleeding in cirrhotics. *Singapore Med. J.* **57**, 132–137 (2016).
309. Zipprich, A., Winkler, M., Seufferlein, T. & Dollinger, M. M. Comparison of balloon vs. straight catheter for the measurement of portal hypertension. *Aliment. Pharmacol. Ther.* **32**, 1351–1356 (2010).
310. Da, B. L. *et al.* Portal Pressure in Non-Cirrhotic Portal Hypertension: To Measure or Not to Measure. *Hepatology* **0**, 1–3 (2019).
311. Pelizzo, G. *et al.* Spleen size evaluation in children: Time to define splenomegaly for pediatric surgeons and pediatricians. *PLoS One* **13**, 1–13 (2018).
312. Dillman, J. R. *et al.* Diagnostic performance of quantitative magnetic resonance imaging biomarkers for predicting portal hypertension in children and young adults with autoimmune liver disease. *Pediatr. Radiol.* **49**, 332–341 (2019).
313. Megremis, S. D., Vlachonikolis, I. G. & Tsilimigaki, A. M. Spleen Length in Childhood with US: Normal Values Based on Age, Sex, and Somatometric Parameters. *Radiology* **231**,

129–134 (2004).

314. Sintusek, P., Siriporn, N., Punpanich, D., Chongsrisawat, V. & Poovorawan, Y. Spleen and Liver Stiffness to Detect Esophageal Varices in Children with Biliary Atresia. *J. Pediatr. Gastroenterol. Nutr.* **69**, 411–415 (2019).
315. Sutton, H. *et al.* Transient Elastography Measurements of Spleen Stiffness as a Predictor of Clinically Significant Varices in Children. *J. Pediatr. Gastroenterol. Nutr.* **67**, 446–451 (2018).
316. Burak Özkan, M., Bilgici, M. C., Eren, E. & Caltepe, G. Diagnostic accuracy of point shear wave elastography in the detection of portal hypertension in pediatric patients. *Diagn. Interv. Imaging* **99**, 151–156 (2018).
317. Donnelly, L. F., Emery, K. H., Bove, K. E. & Bisset, G. S. Normal changes in the MR appearance of the spleen during early childhood. *Am. J. Roentgenol.* **166**, 635–639 (1996).
318. Rosenberg, H. K. *et al.* Normal splenic size in infants and children: Sonographic measurements. *Am. J. Roentgenol.* **157**, 119–121 (1991).
319. Caglar, V., Alkoc, O. A., Uygur, R., Serdaroglu, O. & Ozen, O. A. Determination of normal splenic volume in relation to age, gender and body habitus: A stereological study on computed tomography. *Folia Morphol.* **73**, 331–338 (2014).
320. Kim, D. W. *et al.* Technical performance of two-dimensional shear wave elastography for measuring liver stiffness: A systematic review and meta-analysis. *Korean J. Radiol.* **20**, 880–893 (2019).
321. Garcia-Tsao, G. *et al.* Short-term effects of propranolol on portal venous pressure. *Hepatology* **6**, 101–106 (1986).
322. Jansen, C. *et al.* Increase in liver stiffness after transjugular intrahepatic portosystemic shunt is associated with inflammation and predicts mortality. *Hepatology* **67**, 1472–1484 (2018).
323. Liu, S. *et al.* Quantification of liver iron concentration using the apparent susceptibility of hepatic vessels. *Quant. Imaging Med. Surg.* **8**, 123–134 (2018).
324. Liu, A. *et al.* Splenic T1-mapping: a novel quantitative method for assessing adenosine stress adequacy for cardiovascular magnetic resonance. *J. Cardiovasc. Magn. Reson.* **19**, 1 (2017).
325. de Bazelaire, C. M. J., Duhamel, G. D., Rofsky, N. M. & Alsop, D. C. MR Imaging

Relaxation Times of Abdominal and Pelvic Tissues Measured in Vivo at 3.0 T: Preliminary Results. *Radiology* **230**, 652–659 (2004).

326. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Med.* **12**, e1001779 (2015).
327. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with Those of the General Population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
328. UK Biobank. Protocol for a large-scale prospective epidemiological resource. (2006). Available at: https://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf.
329. Ofojekwu, M.-J. N. *et al.* Hemoglobin and Serum Iron Concentrations in Menstruating Nulliparous Women in Jos, Nigeria. *Lab. Med.* **44**, 121–124 (2013).
330. Chow, K. U., Luxembourg, B., Seifried, E. & Bonig, H. Spleen Size Is Significantly Influenced by Body Height and Sex: Establishment of Normal Values for Spleen Size at US with a Cohort of 1200 Healthy Individuals. *Radiology* **279**, 306–313 (2016).
331. Gilligan, L. A. *et al.* Magnetic resonance imaging T1 relaxation times for the liver, pancreas and spleen in healthy children at 1.5 and 3 tesla. *Pediatr. Radiol.* **49**, 1018–1024 (2019).
332. Mejias, M. *et al.* Relevance of the mTOR signaling pathway in the pathophysiology of splenomegaly in rats with chronic portal hypertension. *J. Hepatol.* **52**, 529–539 (2010).
333. O'Malley, D. P. *Atlas of Spleen Pathology*. *Atlas of Spleen Pathology* (Springer New York, 2013). doi:10.1007/978-1-4614-4672-9
334. NHS & Department of Health. *National schedule of reference costs*. (2018).
335. National Institute for Health and Care Excellence (NICE). *LiverMultiScan for liver disease: Medtech innovation briefing*. (2019).
336. Pessione, F. *et al.* Five-year survival predictive factors in patients with excessive alcohol intake and cirrhosis. Effect of alcoholic hepatitis, smoking and abstinence. *Liver Int.* **23**, 45–53 (2003).
337. Idilman, I. S. *et al.* A comparison of liver fat content as determined by magnetic resonance imaging-proton density fat fraction and MRS versus liver histology in non-alcoholic fatty liver disease. *Acta radiol.* **57**, 271–278 (2016).