Non-Invasive Characterization of Liver Disease

By Multimodal Quantitative Magnetic Resonance

Markus Karlsson
Non-Invasive Characterization of Liver Disease
By Multimodal Quantitative Magnetic Resonance

Markus Karlsson
“The greatest teacher, failure is”
-Yoda
# CONTENTS

ABSTRACT .................................................................................................................. I
SVENSK SAMMANFATTNING ................................................................................... III
LIST OF PAPERS ..................................................................................................... V
ABBREVIATIONS .................................................................................................. IX
ACKNOWLEDGEMENTS ...................................................................................... XI

1 INTRODUCTION ............................................................................................. 1
   1.1 The Liver .................................................................................................. 1
   1.2 Chronic Liver Disease ......................................................................... 3
   1.3 Magnetic Resonance ........................................................................... 6

2 AIMS & THESIS OUTLINE ........................................................................... 9
   2.1 Specific Aims of Each Paper ............................................................... 9
   2.2 Outline of the Thesis .......................................................................... 9

3 FAT ..................................................................................................................... 11
   3.1 Magnetic Resonance Spectroscopy ................................................... 12
   3.2 Chemical Shift Encoded Imaging ....................................................... 14
   3.3 MRS vs CSE ........................................................................................ 16

4 IRON .................................................................................................................... 19
   4.1 R2*- Mapping ...................................................................................... 19
   4.2 R2-Mapping ........................................................................................ 24
   4.3 R2* vs R2 ........................................................................................... 26

5 FIBROSIS & INFLAMMATION ...................................................................... 29
   5.1 Magnetic Resonance Elastography .................................................... 29
   5.2 T1-Mapping ........................................................................................ 34
   5.3 MRE vs T1 .......................................................................................... 39

6 FUNCTION ........................................................................................................ 41
   6.1 Gadoxetate-Enhanced MRI ................................................................. 41
   6.2 Gadoxetate-Enhanced MRI in Liver Disease .................................... 51
   6.3 Association between Gadoxetate-Enhanced MRI and MRCP .......... 54
   6.4 Translating Liver Function Between Humans and Rats .................. 57
   6.5 Future of Liver Function and Gadoxetate-Enhanced MRI ............... 62

7 SUMMARY & FUTURE OUTLOOK ................................................................. 65

8 BIBLIOGRAPHY .............................................................................................. 67
ABSTRACT

There is a large and unmet need for diagnostic tool that can be used to characterize chronic liver diseases (CLD). In the earlier stages of CLD, much of the diagnostics involves performing biopsies, which are evaluated by a histopathologist for the presence of e.g. fat, iron, inflammation, and fibrosis. Performing biopsies, however, have two downsides: i) biopsies are invasive and carries a small but non-negligible risk for serious complications, ii) biopsies only represents a tiny portion of the liver and are thus prone to sampling error. Moreover, in the later stages of CLD, when the disease has progressed far enough, the ability of the liver to perform its basic function will be compromised. In this stage, there is a need for better methods for accurately measuring liver function. Additionally, measures of liver function can also be used when developing new drugs, as biomarkers for drug-induced liver injury (DILI), which is a serious drug-safety issue.

Magnetic resonance imaging (MRI) is a non-invasive medical imaging modality, which have shown much promise with regards to characterizing liver disease in all of the above-mentioned aspects. The aim of this PhD project was to develop and validate MR-based methods that can be used to non-invasively characterize liver disease.

Paper I investigated if R2* mapping, a MR-method for measuring liver iron content, can be confounded by liver fat. The results show fat does affect R2*. The conclusion was therefore that fat must be taken into account when measuring small amounts of liver iron, as a small increase in R2* could be due to either small amounts of iron or large amounts of fat.

Paper II examined whether T1 mapping, which is another MR-method, can be used for staging liver fibrosis. The results of previous research have been mixed; some studies have been very promising, whereas other studies have been less promising. Unfortunately, the results in Paper II belongs to the less promising studies.

Paper III focused on measuring liver function by dynamic contrast-enhanced MRI (DCE-MRI) using a liver specific contrast agent, which is taken up the hepatocytes and excreted to the bile. The purpose of the paper was to extend and validate a method for estimating uptake and efflux rates of the contrast agent. The method had previously only been applied in health volunteers. Paper II showed that the method can be applied to CLD patients and that the uptake of the contrast agent is lower in patients with advanced fibrosis.
Paper IV also used studied liver function with DCE-MRI in patients with primary sclerosing cholangitis (PSC). PSC is a CLD where the bile ducts are attacked by the immune system. When diagnosing PSC patients, it is common to use magnetic resonance cholangiopancreatography (MRCP), which is a method for imaging the bile ducts. Paper IV examined if there was any correlation between number and severity of the morphological changes, seen on MRCP, and measures of liver function derived using DCE-MRI. However, the results showed no such correlation. The conclusion was that the results indicate that MRCP should not be used to predict parenchymal function.

Paper V developed a method for translating DCE-MRI liver function parameters from rats to humans. This translation could be of value when developing new drugs, as a tool for predicting which drugs might cause drug-induced liver injury.

In summary, this thesis has shown that multimodal quantitative MR has a bright future for characterizing liver disease from a range of different aspects.
SVENSK SAMMANFATTNING


Magnetisk resonanstomografi (MRT) är medicinsk avbildningsteknik med möjlighet utvärdera levern. Möjlet med detta doktorandprojekt var att utveckla och utvärdera MRT-baserade metoder för att kunna mäta fett, järn, inflammation, fibros och funktion i levern.

Det första delarbete undersökte om R2*-mätningar, en metod för att mäta järn i levern, kan påverkas av förekomsten av fett i levern. Resultaten visade att fett kan påverka R2*. Slutsatsen var att man måste ta hänsyn till mängden fett i levern när man uppmäter små mängder järn. Detta då en liten ökning av R2* kan bero på både en liten ökning av mängden järn i levern, eller en stor ökning av mängden fett i levern.

Det andra delarbetet undersökte huruvida mätningar av T1 i levern kan användas för att upptäcka och gradera leverfibros. Tidigare studier inom fältet har uppvisat blandade resultat. Vissa studier har varit lovande, medans andra har varit mindre lovande. Tyvärr har resultaten i detta arbete mindre lovande.

Det tredje delarbetet handlade om att mäta leverfunktion med gadoxetat-förstärkt MRT. Gadoxetat är ett kontrastmedel som tas upp av cellerna i levern och hastigheten med vilken kontrastmedlet tas upp kan användas som ett mått på allmän. Syftet med studien var att vidareutveckla och validera en metod för att mäta upptagshastigheten av gadoxetat, då metoden tidigare endast hade testats på friska frivilliga. Studien visade att metoden går att applicera även på patienter med kronisk leverjukdom, samt att upptagshastigheten av gadoxetate är lägre hos patienter med avancerad fibros.
Det fjärde delarbetet använde gadoxetat-förstärkt MRT för att studera patienter med primär skleroserande kolangit (PSC), som är en sjukdom där gallgångarna blir inflamminerade. Vid diagnos av PSC används ofta magnetisk resonanskolangio-pankreatikografi (MRCP) för att avbilda gallgångarna. Syftet med studien var att undersöka huruvida morfologiska gallgångsförändringar, som är synliga på MRCP, korrelerar med leverfunktion uppmätt via gadoxetat-förstärkt MRT. Resultaten visade ingen sådan korrelation, vilket indikerar att MRCP inte ska användas för att prediktera sjukdomsprogression och minskad leverfunktion.

I det femte och sista delarbetet utvecklades en metod för att kunna översätta mätningar av leverfunktion, baserade på gadoxetat-förstärkt MRT, från råttor till människor. Denna metod ska kunna användas vid utveckling av nya läkemedel, som ett verktyg för att förutse om vilka läkemedel som kan ge leverskador hos människor.

Sammanfattningsvis har detta avhandlingsarbete visat att det finns en ljus framtid för att använda multimodal kvantitativ MRT för att karakterisera leversjukdomar.
LIST OF PAPERS

i. Liver R2* Is Affected by Both Iron and Fat: A Dual Biopsy-Validated Study of Chronic Liver Disease

ii. Early Stage Chronic Liver Disease: T1 Relaxation Times and Liver Fibrosis
    *M Karlsson*, T Romu, A Razavi, N Dahlström, O Dahlqvist Leinhard, MF Forsgren, M Ekstedt, S Kechagias, P Lundberg
    In submission.

iii. Model-inferred mechanisms of liver function from magnetic resonance imaging data: Validation and variation across a clinically relevant
    MF Forsgren*, *M Karlsson*, O Dahlqvist Leinhard, N Dahlström, B Norén, T Romu, S Ignatova, M Ekstedt, S Kechagias, P Lundberg*, G Cedersund*

    *M Karlsson*, W Bartholomä, N Dahlström, M Ekstedt, S Kechagias, P Lundberg
    In submission.

v. Translational Modelling Framework for Predicting Human Liver Function Based on Gadoxetate Enhanced Magnetic Resonance Imaging
    *M Karlsson*, N Dahlström, G Cedersund*, P Lundberg*
    In submission.

* Equal contribution
Other Contributions

1. **Nonlinear mixed-effects modelling for single cell estimation: when, why, and how to use it**
   M Karlsson, DLI Janzén, L Durrieu, A Colman-Lerner, MC Kjellsson, G Cedersund

2. **Biomarkers of Liver Fibrosis: Prospective Comparison of Multimodal Magnetic Resonance, Serum Algorithms, and Transient Elastography**
   MF Forsgren*, P Nasr*, M Karlsson, N Dahlström, B Norén, S Ignatova, R Sinkus, G Cedersund, O Dahlqvist Leinhard, M Ekstedt, S Kechagias*, and P Lundberg*
   In submission.

Peer Reviewed Conference Abstracts:

1. **Increased Bile Excretion of Gd-EOB-DTPA in Diffuse Liver Disease: Mechanistic Modeling of qDCE-MRI in Patients With Severe Fibrosis**
   M Karlsson, MF Forsgren, N Dahlström, O Dahlqvist Leinhard, B Norén, M Ekstedt, S Kechagias, G Cedersund, P Lundberg
   ESMRMB, Vienna, Austria, 2016

2. **Diffuse Liver Disease: Measurements of Liver Trace Metal Concentrations and R2* Relaxation Rates**
   M Karlsson, MF Forsgren, N Dahlström, B Norén, M Ekstedt, S Kechagias, P Lundberg
   ESMRMB, Vienna, Austria, 2016

3. **Non-Linear Mixed-Effects Modelling Can Reduce the Acquisition Time When Measuring Liver Function Using Gadoxetate Enhanced MRI**
   M Karlsson, MF Forsgren, G Cedersund, P Lundberg
   ESMRMB, Barcelona, Spain, 2017

4. **Estimating Liver Function in Chronic Liver Disease Patients Using DCE-MRI and Whole-Body Pharmacokinetic Modeling**
   M Karlsson, MF Forsgren, O Dahlqvist Leinhard, N Dahlström, B Norén, M Ekstedt, S Kechagias, G Cedersund, P Lundberg
   ISMRM, Honolulu, USA, 2017
5. **R2*-Relaxometry Can Replace Histology for Detecting Slight Iron Overload in Patients with Early Stage Chronic Liver Disease: A Comparison of R2*, Histology, and Mass-Spectrometry**  
   *M Karlsson, M Ekstedt, N Dahlström, MF Forsgren, S Ignatova, B Norén, O Dahlqvist Leinhard, S Kechagias, P Lundberg*  
   ISMRM, Paris, France, 2018

   *M Karlsson, G Cedersund, P Lundberg*  
   ISMRM, Paris, France, 2018

7. **T1 Relaxation for Measuring Hepatic Fibrosis in a Cohort of Early Stage Chronic Liver Disease**  
   *M Karlsson, T Romu, A Razavi, N Dahlström, O Dahlqvist Leinhard, MF Forsgren, B Norén, M Ekstedt, S Kechagias, P Lundberg*  
   ISMRM, Paris, France, 2018

8. **Assessment of Liver Fibrosis Stage Using Machine Learning and Feature Extraction of Gadoxetate-Enhanced MR Images**  
   *M Karlsson, YC Lu, P Lundberg*  
   ESMRMB, Rotterdam, The Netherlands, 2019

9. **Determining Liver Function: Comparison of Gadoxetate Pharmacokinetic Models Using Perfusion Imaging**  
   *M Karlsson, S Basak, D Longbotham, S Sourbron, G Cedersund, P Lundberg*  
   ISMRM, Montreal Canada, 2019
## ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALT</td>
<td>Alanine Aminotransferase</td>
</tr>
<tr>
<td>AUROC</td>
<td>Area Under Receiver Operating Characteristic Curve</td>
</tr>
<tr>
<td>CLD</td>
<td>Chronic Liver Disease</td>
</tr>
<tr>
<td>CSE</td>
<td>Chemical Shift Encoded Imaging</td>
</tr>
<tr>
<td>DCE-MRI</td>
<td>Dynamic Contrast-Enhanced MRI</td>
</tr>
<tr>
<td>DILI</td>
<td>Drug-Induced Liver Injury</td>
</tr>
<tr>
<td>FA</td>
<td>Flip Angle</td>
</tr>
<tr>
<td>HSC</td>
<td>Hepatic Stellate Cells</td>
</tr>
<tr>
<td>ICP-SFMS</td>
<td>Inductively Coupled Plasma Sector Field Mass-Spectrometry</td>
</tr>
<tr>
<td>ICG</td>
<td>Indocyanine Green</td>
</tr>
<tr>
<td>IR</td>
<td>Inversion Recovery</td>
</tr>
<tr>
<td>LIC</td>
<td>Liver Iron Content</td>
</tr>
<tr>
<td>LSC</td>
<td>Liver-To-Spleen Contrast Ratio</td>
</tr>
<tr>
<td>MELD</td>
<td>Model for End-Stage Liver Disease</td>
</tr>
<tr>
<td>MR</td>
<td>Magnetic Resonance</td>
</tr>
<tr>
<td>MRE</td>
<td>MR elastography</td>
</tr>
<tr>
<td>MRI</td>
<td>MR Imaging</td>
</tr>
<tr>
<td>MRP2</td>
<td>Multidrug Resistance-Associated Protein 2</td>
</tr>
<tr>
<td>MRP3</td>
<td>Multidrug Resistance-Associated Protein 3</td>
</tr>
<tr>
<td>MRP4</td>
<td>Multidrug Resistance-Associated Protein 4</td>
</tr>
<tr>
<td>MRS</td>
<td>Magnetic Resonance Spectroscopy</td>
</tr>
<tr>
<td>NAFLD</td>
<td>Non-Alcoholic Fatty Liver Disease</td>
</tr>
<tr>
<td>NASH</td>
<td>Non-Alcoholic Steatohepatitis</td>
</tr>
<tr>
<td>OATP1</td>
<td>Organic-Anion-Transporting Polypeptide 1</td>
</tr>
<tr>
<td>PDFF</td>
<td>Proton Density Fat Fraction</td>
</tr>
<tr>
<td>PRESS</td>
<td>Point-Resolved Spectroscopy</td>
</tr>
<tr>
<td>PSIR</td>
<td>Phase Sensitive Inversion Recovery</td>
</tr>
<tr>
<td>PSC</td>
<td>Primary Sclerosing Cholangitis</td>
</tr>
<tr>
<td>RE</td>
<td>Relative Enhancement</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver Operating Characteristic</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of Interest</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>T2DM</td>
<td>Type 2 Diabetes Mellitus</td>
</tr>
<tr>
<td>TE</td>
<td>Echo Time</td>
</tr>
<tr>
<td>TI</td>
<td>Inversion Time</td>
</tr>
<tr>
<td>TR</td>
<td>Repetition Time</td>
</tr>
<tr>
<td>VCTE</td>
<td>Vibration Controlled Transient Elastography</td>
</tr>
<tr>
<td>VOI</td>
<td>Volume of Interest</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

First of all, I want to thank my supervisor, Peter L, for giving me this opportunity to become a doctor, as well as for always supporting me and pushing me out of my comfort zone.

I’m also grateful for my co-supervisors. Starting with Gunnar C. I want to thank you for getting me into science and always being inspiring. Continuing with, Nils D. I have appreciated work closely with you, learning so much about radiology and anatomy. Ending with Stergios K. I want to thank you for sharing your wisdom on the body’s most important organ.

I want to acknowledge all the great colleagues who have contributed in all the projects I have worked in. I especially want to thank Mikael F for all the help in the beginning, Mattias E for giving me the opportunity to work on all the exciting NAFLD projects, and Olof D for always providing challenging inputs and inspiration. However, Thobias R, Bengt N, Wolf B, Carl-Johan C, Christian S, Fredrik T, Tino E are not forgotten.

I want to thank Sofie T for making those boring days in the office at little less boring and being a good friend. I also want to thank André A for keeping an eye us youngsters.

I have really enjoyed working at the exciting environment at CMIV, which would be nothing if it wasn’t for Anders P. Also, not many patients would be scanned without the great clinical staff: Henrik E, Christer H, Marcelo M, Emelie B, Mona C, Ninni H, Mirjana V. Additionally, CMIV would be a boring place without all other great staff Dennis C, Catrin N, Maria K, Marie W, Joel H, Patrik H. Let’s not forget all the other PhD students at CMIV, Annette K, Lillian H, Karin L, Barhi K, Milda P, Sebastian S, Natasha M, Robin K, David A, and more, who also make CMIV a great place.

I’m also pleased to have worked with and supervised a number of great students: Jasin B, Jens T, Tetyana B, Yi-Chen L, Faisal Z.

The systems biology group at IMT is also remembered as the place where I started my journey in science. I will especially remember William L, Elin N, Rikard J.

Finally, I want to thank my family for all support during the years.
1 INTRODUCTION

1.1 The Liver

1.1.1 Liver Anatomy

The liver and is the largest internal organ, weighing 1.2-1.5 kg [1]. It is situated in the upper right part of the abdomen (Figure 1.1A).

The liver has a dual blood supply, with blood being supplied by both the hepatic artery and the portal vein. The hepatic artery originates from the aorta and carries oxygenated blood to the liver, while the portal vein returns venous blood from the gastrointestinal system and the spleen. In a healthy individual, the hepatic artery supplies 20-25% of the blood, while the portal vein supplies the remaining 75-80%. Moreover, both blood supplies enter the liver at the liver hilum, together with the common hepatic bile duct. The common hepatic bile duct transport bile from the liver down to the gallbladder and the towards the duodenum.

Figure 1.1 Anatomical overview. (A) The liver (red) is situated in the upper right quadrant of the abdomen. (B) The liver can be divided into eight different segments.

(A) Reproduced from: https://lifesciencedb.jp.
(B) Case courtesy of Dr Craig Hacking, Radiopaedia.org, rID: 62589

After entering the liver, all three vessels divide into their left and right branches, supplying the different liver lobes. The two min branches then divide further into eight different branches supplying eight different
segments (Figure 1.1B). The segments not only have an independent blood supply and bile drainage, each segment also have their independent blood drainage. Blood is drained from each segment into either the left, right, or middle hepatic veins, which are subsequently drained into the inferior vena cava.

On a microscopic level, each segment can be divvied into lobules, which is the basic functional unit of the liver (Figure 1.2). The lobule has the shape of a hexagon, with each corner containing a portal tract, which consists of portal venule and a and arteriole supplying blood, as well as a bile duct draining bile. In the middle of the hexagon, the blood is drained into a central vein, which drains the blood towards the hepatic veins.

![Hepatic lobule](image)

*Figure 1.2 Hepatic lobule. The basic functional unit of the liver is the hepatic lobule.*

Reproduced and modified from: *Accessory Organs in Digestion: The Liver, Pancreas, and Gallbladder - Anatomy and Physiology - OpenStax.*

A more detailed image of the lobule structure is shown in Figure 1.3. At the portal triad, the blood from the arteriole and venule joins into the sinusoid, who’s walls are covered by endothelial, and is eventually drained into the central vein. On the other side of the endothelial cells, the hepatocytes are lined up from the portal tract to the central vein, with the space between the endothelial cells and hepatocytes called space of Disse. On the other side of the hepatocytes, bile is collected and drained towards the bile duct. Finally, the liver also contains other cell types, such as kuppfer cells and hepatic stellate cells (HSC).
1.1.2 Normal Liver Functions

The liver performs an array of different functions in the body. While listing all of them would be out of scope for this thesis, some of the most important ones will be described below.

One function is metabolism, with the liver being central to maintaining the body’s glucose homeostasis, as the organ is one of the body’s main storages of glycogen. Another function is synthesis, liver synthesizing a number of different proteins that are circulating in the blood, such as albumin and transferrin. The liver is also responsible for synthesizing molecules such as cholesterol and bile acids. An additional function of the liver is to clear and metabolize drugs and other substances, such as bilirubin, from the blood.

1.2 Chronic Liver Disease

Chronic liver disease (CLD) is a large and growing problem [2]. CLD is not just one disease; it is rather a collection of diseases affecting the liver, such as viral hepatitis B (HBV) or C (HCV) infection, autoimmune hepatitis, cholestatic disorders, such as primary biliary cirrhosis (PBS) and primary sclerosing cholangitis (PSC). CLD can also be caused by hereditary hemochromatosis, i.e. accumulation of iron in the liver, or by the accumulation of fat, due to either alcoholic liver disease (ALD) or non-alcoholic fatty liver disease (NAFLD).

Regardless of etiology, many CLDs can manifest itself in several different ways. These ways include the accumulation of fat (steatosis) or iron (hemochromatosis), the development of inflammation and fibrosis, and a reduction of liver function.
1.2.1 Steatosis
Steatosis, or excess accumulation of fat in the liver, can occur in several different CLDs. The most common of those diseases are NAFLD and ALD. Steatosis is diagnosed via histopathology of a biopsy and typically graded on a scale from 0 (healthy) to 3 corresponding to fat deposition in < 5 %, 5-33 %, 34-66 %, and > 67 % of the hepatocytes [3].

1.2.2 Iron Overload
Excess iron can be accumulated in the liver in CLDs. One of those diseases is hereditary hemochromatosis [4], where genetic defects distort the body’s regulation of iron storage, which causes iron to accumulate in the liver. Iron can also be accumulated in other CLDs such as NAFLD [5]. Finally, patients receiving blood transfusions, such as thalassemia patients, can also accumulate iron, as the transfused blood contains lots of iron [6].

Iron overload can be assessed from biopsies in two different ways. The first way is to simply measure the liver iron content (LIC) by analyzing the entire biopsy using e.g. mass spectrometry. Measurement of LIC is considered a gold standard for iron accumulation. However, since the whole biopsy is consumed, it cannot also be used for other purposes, such as staging of steatosis and fibrosis. Fortunately, iron can also be assessed by histopathology, using Perls’ Prussian blue stain with iron graded on a scale from 0 (healthy) to 4, with 0: iron granules absent or barely discernible at × 400 magnification; 1: iron granules barely discernible at × 250 magnification and easily confirmed at × 400 magnification; 2: discrete iron granules resolved at × 100 magnification; 3: discrete iron granules resolved at × 25 magnification; and 4: Masses visible at × 10 magnification, or naked eye [7].

1.2.3 Inflammation
All CLDs will produce some form of liver injury and inflammatory response. This response is a key driver of fibrosis development [8]. Inflammation is assessed using a biopsy and histopathology, although there is no universal staging for all CLD, as the histological features of inflammation can vary depending on etiology.

1.2.4 Fibrosis
Fibrosis is the excess accumulation of scar tissue and will eventually occur in all CLDs. The development of fibrosis can take many years to develop and the early stages are reversible. Although, if the underlying liver injury is not stopped, the fibrosis progression will eventually lead to cirrhosis.

HSCs, which normally sit in the space of Disse in a quiescent state, plays a central role in fibrosis development [9]. During the inflammation process, the HSCs will become activated and start promoting the
deposition of scar tissue. The HSCs will also excrete molecules that promotes proliferation and migration of more HSCs, as well as other proinflammatory and profibrotic cells, thus starting a positive feedback loop, increasing the fibrosis development.

Liver fibrosis is also diagnosed using biopsies and typically staged on a scale from 0 to 4 where the different stages are defined as F0: no fibrosis; F1: portal or perisinusoidal fibrosis; F2: portal and perisinusoidal fibrosis; F3: bridging fibrosis; and F4: cirrhosis [10].

1.2.5 Liver Biopsy
As described above, a liver biopsy is the reference standard for assessing the amount of fat, iron, inflammatory activity, and fibrosis. However, using a biopsy has several drawbacks. Performing a biopsy is an invasive procedure and carries a small but not insignificant risk for serious complications [11]. Additionally, a biopsy only represents a tiny fraction of the liver volume and fat, iron, and fibrosis is known to not be homogeneously distributed in the liver [12, 13]. Furthermore, it is difficult for a pathologist to estimate accurately how much e.g. how much fat or fibrosis is present in a biopsy specimen. Therefore, the biopsy method also suffers from intra reader inter-reader and variability [14].

1.2.6 Liver Function
Liver function is more difficult to define and therefore more challenging to quantify. The simplest way to quantify liver function is to use blood panels, such as Model for End-Stage Liver Disease (MELD) [15] and Child-Pugh [16]. Although blood panels are cheap and easy to measure, they are static measurements and only represent liver function indirectly. The blood panels measure proteins that were produced by the liver at some previous stage.

More dynamic and direct alternatives are methods that measure the ability of the liver to take up, metabolize, and dispose of certain substances. The earliest of these methods relied on $^13$C-labeled molecules, which are metabolized in the liver [17, 18]. The metabolized molecules will be broken down into $^13$C-labeled labeled CO$_2$, which can be measured in the expiration air. The downside of these methods is that they require special measurement equipment.

An alternative to the metabolism-based methods are dyes, which are cleared through the liver. A commonly used dye is indocyanine green (ICG), which after injection can be easily measured with a fingertip optical sensor [19]. The dye clearance methods do not measure the livers metabolic capacity, but rather the hepatocytes capacity to transport certain molecules from blood to the bile. A possible drawback of this is that the clearance rate can also be affected by the hepatic blood flow, which can be reduced in
patients with end-stage liver disease, e.g. due to portal hypertension. A second drawback is that the clearance is only measured on a global level, i.e. it is impossible to take into account different functional reserves in different segments of the liver.

1.2.7 Non-Alcoholic Fatty Liver Disease
Non-Alcoholic Fatty Liver Disease (NAFLD) is the hepatic manifestation of the metabolic syndrome, which also includes conditions such as obesity, diabetes, hypertension, and hyperglycemia [20]. NAFLD is characterized by steatosis, i.e., the accumulation of fat in the liver, without any other cause, such as excessive alcohol consumption.

When steatosis is accompanied by histologic features of inflammation and cellular injury, the disease is called non-alcoholic steatohepatitis (NASH). NASH is thought to be a more aggressive form of NAFLD, with a greater risk of fibrosis progression. The severity of NASH is often graded using the NAFLD activity score (NAS), with a combined measure of steatosis, hepatocellular ballooning, and lobular inflammation [3].

1.2.8 Primary Sclerosing Cholangitis
Primary sclerosing cholangitis (PSC) is a chronic immune-mediated liver disease affecting the intra- and extrahepatic bile ducts. The disease is characterized by inflammation and fibrosis around the bile ducts [21]. As PSC progresses, the bile ducts will be destroyed, and the patient will develop end-stage liver disease. Additionally, PSC is also associated with an increased risk for malignancies, hepatobiliary cancer [22] as well as extrahepatic malignancies such as colorectal cancer [23, 24]. Lacking effective treatment options, liver transplantation is often the only treatment [25].

1.2.9 Drug-Induced Liver Injury
Drug-induced liver injury (DILI) is an adverse reaction to a drug affecting the liver and can manifest itself in different ways, such as steatosis, inflammation, or necrosis [26]. DILI is often caused by a direct toxic effect of either drug itself, or one of its metabolites, as many drugs are metabolized in the liver. Although DILI can also be caused by an immunological response to the drug and its metabolites. Moreover, DILI is a potentially serious condition which can cause acute liver failure, and even require transplantation [27]. Therefore, DILI is a major concern when developing new drugs [28].

1.3 Magnetic Resonance
1.3.1 Nuclear Magnetic Resonance
Magnetic resonance (MR), or nuclear magnetic resonance, stems from the fact that the angular momentum, or spin, of a hydrogen proton (or other nuclei with half integer spins) can interact with an external magnetic field.
When a hydrogen proton is placed in a strong magnetic field, $B_0$, the spins will start to precess around the direction of $B_0$, with a frequency known as the Larmor frequency. The spins will align with $B_0$, either in the same direction or in the opposite, with more spins aligning with $B_0$ rather than against $B_0$. This will lead to a net magnetization along $B_0$.

Using radio frequency (RF) pulses, at the Larmor frequency, the net magnetization can be tipped from the direction of $B_0$ (z-axis) into the xy-plane. When the magnetization is in the xy-plane, it will start to precess around the z-axis. It is this xy-component of the net magnetization, $M_{xy}$ or transverse magnetization, that will generate the MR-signal.

After the RF-pulse, the z-component of the magnetization, $M_z$ or longitudinal magnetization, will start to regrow back to the equilibrium (while $M_{xy}$ decreases) in a process called spin-lattice relaxation, or T1-relaxation. This relaxation occurs with a time constant $T_1$, which is defined as the time it takes for 63% of the relaxation to regrow.

Simultaneously with the T1-relaxation, $M_{xy}$ will also decrease in another process called spin-spin-relaxation, or T2-relaxation. T2-relaxation is caused by dephasing of the spins in the xy-plane. Directly after the RF-pulse, all the spins will be in phase. However, as the spins randomly interact with each other, they will temporarily encounter slightly different magnetic fields strengths and will therefore precess at slightly different frequencies. When the spin precess at different frequencies, they will dephase, thus causing $M_{xy}$ to decrease.

Apart from T2-relaxation, there is another process that will cause $M_{xy}$ to decrease. The process is called T2*-relaxation, or apparent T2-relaxation. T2*-relaxation is caused by the fact that the magnetic field inside an MR-scanner will not be perfectly spatially homogeneous. This means that there will be small differences in field strengths which will generate additional dephasing. This dephasing due to the static inhomogeneity can be reversed with a 180° refocusing RF-pulse, used in spin-echo sequences. Therefore, spin echo sequences depend on T2, while gradient echo sequences depend on T2*.

1.3.2 Magnetic Resonance Imaging
In magnetic resonance imaging (MRI), images are created based on the principle described above, that spins experiencing different field strengths will have different precession frequencies. This principle is used by applying magnetic field gradients, which a well-defined spatial (and temporal) variation in the magnetic field across the geometry of the MR-scanner. This means that the frequency of the spins will depend on their location. Based on this, images can be calculated by Fourier transforming the signal.
There are a number of different parameters that can be changed when acquiring MR images. Three of the most common basic are repetition time (TR), echo time (TE), and flip angle (FA).

It typically takes multiple consecutive RF-pulses, and subsequent reading of the signal, to generate an image. TR is the time between two such RF-pulses. TR is generally related to T1-weighting in the images, i.e. when different types of tissue have different signal intensity based on T1-relaxation time. Varying TRs result in different amounts of non-equilibrium $M_z$, thus a varying degree of saturation of the signal in the z-direction.

TE is the time between the RF-pulse and the detection of the signal. TE is generally related to T2- or T2*-weighting, as different T2- or T2*-relaxation times means that different amount of spins will have had time to diphase.

FA is simply the angle that the magnetization is deviates from the z-axis after the RF-pulse. A FA of 90° means that all magnetization is flipped into the xy-plane, while a FA of 180° means that the magnetization will be completely inverted, and no magnetization component will remain in the xy-plane.
2 AIMS & THESIS OUTLINE

The overarching aim of this PhD project was to develop and validate MR-based methods that can be used to characterize liver disease non-invasively. More specifically, this meant working with methods for fat and iron quantification and the staging of fibrosis and inflammation, and determining overall liver function.

2.1 Specific Aims of Each Paper

**Fat & Iron**

I. The aim of Paper I was to investigate how liver iron measurement, based on $R_2^*$ relaxation rates, is affected by the presence of fat.

**Fibrosis & Inflammation**

II. The aim of Paper II was to evaluate how T1 relaxation times can be used to stage fibrosis and how the measurements are affected by inflammation and the presence of iron.

**Function**

III. The aim of Paper III was to extend and validate a previously developed method for determining liver function, based on gadoxetate-enhanced MRI and pharmacokinetic modeling.

IV. The aim of Paper IV was to evaluate the utility of gadoxetate-enhanced MRI as a biomarker for liver function in PSC.

V. The aim of Paper V was to develop a modeling framework for translating gadoxetate-enhanced MRI-based biomarkers for DILI from preclinical animal models to humans.

2.2 Outline of the Thesis

The coming chapters will describe the most common MR methods for measuring fat (Chapter 3), iron (Chapter 4), fibrosis & inflammation (Chapter 5), and function (Chapter 6). These chapters will also present and discuss key results generated in each area during the PhD project, both results that are included in the papers, but also unpublished results. The last chapter will present the final conclusions and present a final outlook for quantitative MRI in liver disease.
Magnetic resonance is an excellent modality for quantifying liver fat, since fat molecules contain plenty of hydrogen protons. However, the liver also contains an even larger water, which also has plenty of hydrogen protons. One therefore needs to be able to separate the signal coming from the water protons from the signal coming from the fat protons. This is commonly done by exploiting the fact that protons in fat and water actually have slightly different resonance frequencies [29].

This physical effect is called a chemical shift (Figure 3.1). A chemical shift could be reported as the difference in resonance frequency between different nuclei. However, that would mean the exact chemical shift was dependent on the magnetic field strength. Chemical shifts, δ, are therefore often reported in ppm relative to the chemical shift of a chosen reference molecule (typically Tetramethylsilane):

\[ \delta = \frac{\nu - \nu_{\text{ref}}}{\nu_{\text{ref}}} \times 10^6, \]  

where \( \nu \) and \( \nu_{\text{ref}} \) are the resonance frequencies of the compound of interest and the reference molecule respectively.

Figure 3.1 Schematic illustration of chemical shift. (A) shows a molecule where the electronegative oxygen atom attracts parts of the electron cloud from the blue hydrogen atom. The electron cloud around the blue hydrogen will therefore be less dense, meaning that the blue hydrogen will experience less magnetic shielding and thus a stronger effective magnetic field. This will in turn lead to a higher resonance frequency and a larger chemical shift. (B) shows a hypothetical proton spectrum of the molecule in (A)

The water proton has a single resonance at a chemical shift of 4.7 ppm. The typical fat molecule on the other hand has multiple resonances, since the protons in fat can have different atoms surrounding it (Figure 3.2),
unlike the symmetrical water molecule. Typically, there are six different resonances in a fat molecule that need to be taken into account [30]. These resonances are shown in Figure 3.2 along with their relative contribution to the total fat signal.

![Fat Resonances](image)

Figure 3.2 Fat Resonances. Schematic representation of typical triglyceride molecule. The chain shown is linoleic acid. R indicates the other fatty acid chains in the triglyceride. Also, a schematic representation of a human liver fat spectrum. The area under each resonance is proportional to the relative contribution from the different resonances to the total fat signal.

There are two different MR-methods that can be used to measure liver fat, MR spectroscopy (MRS) and chemical shift encoded imaging (CSE).

**Table 3.1 Chemical shift and relative signal of the different fat resonances**

<table>
<thead>
<tr>
<th>Resonance</th>
<th>Chemical shift [ppm]</th>
<th>Percentage of total fat signal [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.3</td>
<td>4.7</td>
</tr>
<tr>
<td>2</td>
<td>4.2</td>
<td>3.9</td>
</tr>
<tr>
<td>3</td>
<td>2.75</td>
<td>0.6</td>
</tr>
<tr>
<td>4</td>
<td>2.1</td>
<td>12.0</td>
</tr>
<tr>
<td>5</td>
<td>1.3</td>
<td>70.0</td>
</tr>
<tr>
<td>6</td>
<td>0.9</td>
<td>8.8</td>
</tr>
</tbody>
</table>

### 3.1 Magnetic Resonance Spectroscopy

In MRS, the gradients and RF pulses of the MR scanner are used to select a localized volume of interest (VOI) in the liver, typically of the size of 10x10x10 – 30x30x30 mm³. The signal from this VOI is recorded and Fourier-transformed to obtain a spectrum where the water and fat signals can
be seen. Figure 3.3 shows an example of a VOI selected in a liver as well as a spectrum, acquired at 1.5 T. From the figure, it is clear that not all six resonances, seen in Figure 3.2, are distinguishable. There are two reasons for this. First, the ability to separate different resonances is dependent on field strength and with clinical scanners (1.5-3.0 T), it is not always possible to separate all the resonances. Second, the large water resonance covers some of the smaller fat resonances.

![Figure 3.3 MRS. The figure illustrates a VOI selected in a human subject, as well as a spectrum, acquired at 1.5 T. The green area in the spectrum represents the water resonance, plus some hidden fat resonances. The blue area represents the main fat resonances.](image)

Even though not every individual resonance can be distinguished in the spectrum, the liver fat fraction can still be quantified. One way of doing this is to first measure the signal, i.e. the area under the spectrum, from the main fat resonances, (Figure 3.3; blue area) and then measure the area under the water resonance (Figure 3.3; green area), which also contains some hidden fat hidden resonances. After this, the fat fraction, \( FF \), can be calculated:

\[
FF = \frac{\text{area under main fat resonances}}{\text{area under all resonances}},
\]

although this measure will be confounded by the hidden fat resonances. However, if the fat spectrum can be assumed to be the same in all humans, the FF can be corrected and the signal fat fraction, \( SFF \), can be calculated [31]:

\[
SFF = \frac{FF}{1.138 - 0.339 \times FF}
\]

Note that this is only called the signal fat fraction. The reason for this is that the signals from fat and water are not only affected by the amount of fat and water protons, i.e. the proton density, but also by the \( T_1 \) and \( T_2 \) relaxation times. The magnitudes of these effects will be determined by the
echo time (TE) and repetition time (TR) of the pulse sequence. Fortunately, these effects can be corrected for by

\[ S_0 = \frac{S}{(1 - e^{-\text{TR}/T_1}) \cdot e^{-\text{TR}/T_2}} \]  

(3.4)

where \( S_0 \) and \( S \) are the corrected and uncorrected signals respectively.

One way to implement the T1 and T2 correction is to use literature values for the relaxation times. While easy, this approach will be confounded if the relaxation times are changed, i.e. if water T2 is lowered in patients with iron overload [32]. Another approach is to acquire several spectra with different TE, so that T2 can be estimated from the data. This only leaves the effect of T1, which can be neglected if a sufficiently long TR is selected, typically \( \geq 3000 \text{ ms} \) [29]. Once all of the above corrections have been done, the proton density fat fraction (PDFF) can be calculated.

### 3.2 Chemical Shift Encoded Imaging

CSE, also known as Dixon imaging, acquires images (typically spoiled gradient echo images) with different TEs [33]. The basic idea is that as the fat and water protons precess at different frequencies, their signals can either add up, if they are in phase, or cancel each other out, if they are out of phase, depending on the TE.

The simplest example of this a 2-point-dixon with out-of-phase and in-phase images. The out-of-phase image is acquired when the water protons are out of phase with the main fat resonance at 1.3 ppm (TE = 2.3/1.15 ms at 1.5/3 T). This means that (magnitude) signal, \( S_{OP} \), in the image will be

\[ S_{OP} = |S_W - S_F| \]  

(3.5)

where \( S_W \) and \( S_F \) are the signals from water and fat respectively. Meanwhile, the in-phase images are acquired when those protons are in phase (TE = 4.6/2.3 ms at 1.5/3 T), meaning that (magnitude) signal, \( S_{IP} \), in the image will be

\[ S_{IP} = |S_W + S_F|. \]  

(3.6)

From these two images the fat and water signals can easily be calculated:

\[ S_W = \frac{S_{IP} + S_{OP}}{2} \]  

(3.7)

\[ S_F = \frac{S_{IP} - S_{OP}}{2}. \]  

(3.8)

Finally, these two images can be used to calculate the signal fat fraction, SFF:

\[ \text{SFF} = \frac{S_F}{S_W - S_F}. \]  

(3.9)
Unfortunately, this fat fraction has several problems. The first problem is that it only considers the main fat resonance at 1.3 ppm. The most common way to resolve this is to use a predetermined fat spectrum, such as the one in Figure 3.2/Table 3.1. This spectrum can then be fitted to the signal intensities from the images using the following equation:

$$S(TE) = S_W + S_F \sum_{p=1}^{6} r_p e^{i2\pi f_p T_E},$$

(3.10)

where $r_p$ is the relative amplitudes of the six fat resonances (Figure 3.2/Table 3.1), and $\Delta f_p$ is the frequency shift of the fat resonances relative to water. Since literature values are used for $r_p$ and $\Delta f_p$, the model has two parameters to be estimated: $S_W$ and $S_F$.

The second problem is that the signal fat fraction does not consider the R2* relaxation taking place between the echoes. This is resolved by adding a monoexponential R2*-decay to equation (3.10):

$$S(TE) = S_W + S_F \sum_{p=1}^{6} r_p e^{i2\pi f_p T_E} e^{-T_E R2*}.$$

(3.11)

As Equation 11 contains three unknowns, more than two images have to be used. Commonly, six images are acquired six images (Figure 3.4) with TEs equal to the first six times when water and the main fat resonance are in and out of phase [34]. Finally, as with MRS, the different T1 values of fat and water must also be considered. This is typically done by using a low flip angle, typically 3-10°, thus avoiding T1 bias. If all the above-mentioned corrections are performed, PDFF [34, 35] can be calculated as:

$$PDFF = \frac{S_F}{S_W + S_F}.$$

(3.12)
The method above, based on magnitude images does have one significant limitation. The range of PDFF values is limited to 0-50% [36], as the analysis requires the assumption that the signal is water-dominant. Fortunately, even patients with severe steatosis still have PDFF < 50% [37]. The problem stems from the fact that if you have a voxel with a water/fat content of 80/20% and another voxel with a water/fat content of 20/80%, the voxels proton density-weighted signals will have the same magnitudes, but different phases. To overcome this limitation, complex images must be used, so that the phase of the signal is preserved. The signal model is then changed to:

\[
S(TE) = \left( S_W + S_F \sum_{p=1}^{6} S_p e^{i2\pi\Delta f_p T_E} \right) e^{i2\pi\psi T_E} e^{-T_E\cdot R^2}, \tag{3.13}
\]

where \( \psi \) is the local field inhomogeneity. While this method is the most accurate, it also requires that potential errors in phase of the signal, e.g. due to Eddy currents, are corrected [38].

### 3.3 MRS vs CSE

In general, both MRS and CSE are accurate and reliable techniques, which have both been validated against histopathology [31, 37, 39-43]. However, the agreement with histopathology is often not perfect. This is probably due to the usual limitations of biopsies and histopathology, namely the sampling variability and the fact that it is difficult for a human reader to estimate the amount of fat in a biopsy sample [44]. While the sampling
variability is not a problem for CSE, as it covers the whole liver, it could be a problem for MRS. However, a typical voxel for fat quantification is still much larger than a biopsy. For example, in Paper I, a 30x30x30 mm³ voxel was used. This, combined with the fact that the inter-reader agreement and reproducibility for the MR methods are high [34, 45, 46], offers an argument that MR should actually be considered the gold standard for liver fat quantification, or at least suggests it is better than histopathology.

As a rule of thumb, MRS is a little more accurate than CSE, particularly at low fat levels, while CSE is easier to use. In fact, many studies of CSE methods actually use MRS as a reference standard [34, 38, 45-48]. In terms of practical use, the post-processing of CSE images is typically fully automated and built into modern scanners, while the processing of spectra often requires manual work. Furthermore, it is also common that many MR scanners outside of research centers do not have access to spectroscopy, while commercial implementations of CSE are becoming more and more widespread. All in all, it would therefore probably be better to implement CSE in a clinical workflow, while MRS could be used in clinical studies where the extra accuracy might be needed, and extra post-processing resources are available.

Since PDFF is such an accurate non-invasive measurement of liver fat, it is ideal for use in clinical studies. An example of this is shown in Figure 3.5. In short, the data come from a study (still unpublished) on diabetes mellitus type 2 (T2DM), where 46 T2DM patients were included, along with 46 controls matched with respect to age, sex, and smoking status. The participants underwent an extensive MR examination at 1.5 T, including MRS to measure PDFF. Other measured parameters included liver iron and fibrosis, heart function, and body composition. MRS was performed using a point-resolved spectroscopy (PRESS) sequence (TR = 1500 ms, TE = 35 ms). The spectra were acquired using an NSA = 8, and two dummy acquisitions for the magnetization to reach steady state. PDFF was calculated using the method described in equations 3.3-3.5.
Figure 3.5 Example of MRS PDFF applied in clinical research. 92 subjects (46 T2DM patients and 46 controls) had PDFF measured using MRS. The results, which are part of a larger, still unpublished, MR study, show that PDFF is high in T2DM patients. The study also included further measurements on the liver, as well as measurements of heart function and body composition.
4 IRON

Liver iron content (LIC) is another parameter that is well suited to being measured using MRI, due to iron being strongly ferromagnetic. In the cells of the liver, iron ions are bound to ferritin proteins, which are aggregated into larger hemosiderin complexes. All these iron-containing complexes act like tiny magnets, affecting the local magnetic field. While doing this, they create tiny local inhomogeneities in the magnetic field. These static inhomogeneities cause an increased signal dephasing in gradient echo images, i.e. a shortening of $T_2^*$ (or an increase in $R_2^*$ as $R_2^* = 1/T_2^*$). When using a spin echo sequence on the other hand, such static signal dephasing is rephased by the refocusing pulse. This means that that the static inhomogeneities do not affect $T_2$. However, the water protons are also subject to random diffusion. During the diffusion, the proton spins randomly pass through inhomogeneities creating signal dephasing. Unlike static inhomogeneities, these random effects are not rephased by the spin echo refocusing pulse, thus causing a shortening of $T_2$ (or an increase in $R_2$ as $R_2 = 1/T_2$).

4.1 $R_2^*$- Mapping

The most common method to quantify $R_2^*$ is to acquire $T_2^*$-weighted gradient echo images with different TE [49-51] (Figure 4.1). Ideally, the signal from these images should decrease monoexponentially as a function of TE:

$$S(TE) = S_0 * e^{-TE* R_2^*}, \quad (4.1)$$

where $S$ is the signal intensity and $S_0$ is the signal intensity before any transverse relaxation.
In Paper I, the ability of R2* to measure LIC was investigated. In short, 81 patients undergoing liver biopsy due to suspected CLD underwent an MR examination (1.5 T Philips), including R2* measurement. Directly after the MR examination, two liver biopsies were obtained. The first biopsy was sent for standard histopathological examination, including semi-quantitative staging of iron using Perls’ Prussian blue stain [7]. The second biopsy was used to quantify LIC using inductively coupled plasma sector field mass-spectrometry (ICP-SFMS) to obtain a reference measure of LIC.

R2* of the water protons were quantified from images acquired using an axial 3D multi-echo turbo filed echo sequence with spectral pre-saturation with inversion recovery fat saturation (Figure 4.1). The imaging used the following parameters: repetition time (TR) = 26.0 ms, echo time (TE) = 4.6/9.2/13.8/18.4/23.0 ms, flip angle (FA) = 20°, spatial resolution = 1.25x1.25x5.00 mm³, and field of view (FOV) = 320x290x80 mm³. A single region of interest (ROI) was placed in the liver. R2* was calculated by fitting equation 4.1 to the mean signal intensities from the ROIs at the different echo times.

R2* was correlated with LIC and compared to histopathology. The results are shown in Figure 4.2A, with the different markers indicating the score given by the histopathologist. The figure shows two things. First, there is a linear relationship between R2* and LIC. Second, R2* is better than the histology in terms of identifying patients with elevated iron levels (LIC > 1.2 mg/g) [52].
Figure 4.2 Association between LIC and R2*. (A) Correlation between LIC and R2*. The different markers indicate the Perls’ score given by the histopathologist. The black line is a linear regression. (B) Our calibration curve for R2* and LIC compared to other calibration curves found in the literature [53-55]. Note that the curve by Hankins is almost completely covered by the curve by Karlsson. Issues with R2*


This correlation between R2* and LIC can also be compared with similar studies. Figure 4.2B shows the data and calibration curve from Paper I together with calibration curves from other studies [53-55]. It can be noted that these studies used different patient cohorts. In general, the patients in Paper I had very little iron, and only eight patients had elevated levels of LIC. The other studies used patients receiving blood transfusions. Such patients typically have much higher levels of LIC, and many of them are receiving chelation therapy. Almost all of the patients in those studies had an LIC between 5 and 40 mg/g, i.e., they would therefore be outside the range of the graphs in Figure 4.2. Despite this, the figure shows that there is a good agreement between the different studies.

4.1.1 Limitations of R2*

There are two major limitations of the approach described above, the effects of noise bias and the effect of fat.

4.1.1.1 Noise Bias

The problem of noise bias stems from the fact that noise in magnitude MR images is not normally distributed, but rather Rician distributed [56]. An effect of this distribution is that the signal will not decay toward zero, but rather to an offset slightly above zero. This can in turn lead to an underestimation of R2*, at least when TEs are too long and the signal actually approaches zero. Alternatively, one can say that the problem occurs when R2*
is too high, i.e. in patients with the most iron. In general, this problem is also larger at 3 T than at 1.5 T, as R2* increases roughly by a factor of two when going from 1.5 T to 3 T [57, 58].

There are different approaches that can be used to address the noise bias issue, either alone or in combination. First, use a 1.5 T scanner for the R2*-mapping. Second, use TEs that are as short as possible. Third, truncate the signal, i.e. remove the last time points, where the signal has approached zero, from the fitting. However, this complicates the analysis process, especially if the analysis is to be automated. Fourth, change the signal model to include a constant offset, C:

\[
S(TE) = S_0 \times e^{-TE \times R2^*} + C
\]

which can account for the noise floor effect [53, 59]. Last, use complex images instead of the magnitude images [50]. When the complex images are used, the noise becomes normally distributed in the complex space and the noise floor issues disappear (see section 3.2).

### 4.1.1.2 Fat

As described in the previous chapter, the presence of fat in the liver will affect the signal from gradient echo images, as the signal from the fat protons goes in and out of phase with the signal from the water protons. There are three main methods for handling this issue. The first method is to select TEs so that the main fat resonance at 1.3 ppm (containing about 70 % of the fat signal) is always in phase with the water signal. This method is easy, but it does have the drawbacks of not addressing the signal from the full fat spectrum (Figure 3.2), as well as forcing the use of longer TEs. Many authors recommend using a minimum TE around 1 ms [53, 60, 61], while the shortest in-phase TE is 4.6 ms at 1.5 T. The second method is to acquire images with fat saturation [49]. Again, this is also a simple solution, but it can have a problem if the fat suppression is imperfect. Also, some implementation of fat suppression may only target the main fat resonance. The third method is to acquire non-fat suppressed images and model the presence of fat using the same methods used for CSE imaging and fat quantification (see section 3.2) [50].

Apart from the signal from fat protons confounding the water signal, there is also the question of whether the presence of fat in the liver can affect the R2* of the water protons themselves. On a cellular level, fat is stored in lipid droplets distributed inside the hepatocytes (Figure 4.3). If these droplets have a susceptibility that is different from the cytosol around them [62], they will create small static inhomogeneities in the magnetic field, just like iron.
Figure 4.3 Histological example of liver fat. The image shows a part of a liver biopsy. The pink background is formed of cells stained with hematoxylin and eosin, while the "white bubbles" are fat droplets.

Image courtesy of Mattias Ekstedt.

This question of whether fat could affect the water R2* has been investigated in a few studies. First, Mamidipalli et al. found a correlation between R2* and PDFF in a pediatric NAFLD cohort using CSE [63]. However, the study was limited by the fact that no biopsies were obtained from the patients, meaning that the researchers could not measure the actual LIC. Since they could not measure LIC, they could not exclude the possibility that their results were just an effect of a correlation between PDFF and LIC. In a second study, Bashir et al. examined the same question, this time including the histopathological grading of biopsies, in a cohort of adult NAFLD patients [64]. They found that the R2* of water in the liver was more strongly associated with PDFF than the histological iron score was. Again, further suggesting that fat affects the R2* of water. However, the study only used the semi-quantitative histopathological iron grade, and not quantitative LIC. As is shown in Figure 4.2A, histopathological iron grade is not necessarily a very good measure in patients with lower iron levels.

The effect of fat on water R2* was also studied in Paper I. Since this study included an actual measure of LIC, the patients with elevated iron levels could be excluded. Liver PDFF was measured using MRS (same procedure as described in section 3.3.

The measurements of R2*, LIC, and PDFF resulted in a moderate correlation between PDFF and LIC (Figure 4.4A) with the following linear regression:

\[ R2^* = 28.6 + 1.04 \cdot PDFF, \]  
(4.3)

indicating that for every percentage point of fat a patient has in the liver, R2* increases by close to one s⁻¹. Moreover, a linear regression analysis with R2* as the response variable and LIC and PDFF as predictors gave the following model:

\[ R2^* = 16.1 + 25.4 \cdot LIC + 0.96 \cdot PDFF. \]  
(4.4)
From the equation, two conclusions can be drawn. One, the equation affirms the notion that $R_2^*$ increases by close to one s$^{-1}$ for every percentage point of PDFF. Two, the effect of iron on $R_2^*$ is much larger than the effect of fat. In general, patients with severe steatosis can have a PDFF of up to 30 % [37]; $R_2^*$ may thus be confounded by up to 30 s$^{-1}$. Patients with a clinically significant iron overload can have LIC levels of tens of mg/g and $R_2^*$ of several hundreds of s$^{-1}$. The confounding effect of fat is therefore largest when the absolute levels of LIC are lower, as in NAFLD patients [65].

Finally, equation 4.5 was used to create a PDFF-corrected $R_2^*$ (PDFF-$R_2^*$), by subtracting 0.96 times PDFF for each patient:

$$PDFF-R_2^* = R_2^* - 0.96 \cdot PDFF.$$  \hspace{1cm} (4.5)

When comparing the PDFF-$R_2^*$ to LIC, the correlation increased from $R = 0.82$ to $R = 0.87$ (Figure 4.4B).

#### 4.2 R2-Mapping

To quantify $R_2$, the approach is similar to $R_2^*$ quantification, with the major difference that spin echo images are acquired instead of gradient echo images. $R_2$ can also be estimated by fitting a monoexponential decay to the signal intensities from the different TEs:

$$S(TE) = S_0 \cdot e^{-TE \cdot R_2}. \hspace{1cm} (4.6)$$
Figure 4.5 Example of how to calculate R2. The figure shows an experiment where 16 T2-weighted (fat-suppressed) images acquired at different TEs. Note that only five of the images are shown. A region of interest (red) is drawn in the liver and the signal intensity from the ROI is plotted against the TE. To this, a mono-exponential decay-function (blue) can be fitted in order to estimate R2.

In the same cohort as in Paper I, R2 of water was also measured using a single-slice axial turbo gradient spin echo sequence with spectral pre-saturation with inversion recovery fat saturation. The following imaging parameters were used: TR = 190 ms, TE = 6xn ms, with n = 1,..., 16, FA = 90°, spatial resolution = 0.98x0.98x10.00 mm³, FOV = 375x294x10 mm³. As with R2*, an ROI was drawn and the signal was fitted to a mono-exponential decay.

Figure 4.6A shows the correlation between R2 and LIC. The figure shows that the correlation seems to be linear, but is lower compared to R2*. Furthermore, Figure 4.6B shows a correlation between R2 and PDFF, after patients with elevated LIC have been removed. Unlike R2*, there seems to be little to no correlation between R2 and PDFF, due to the lack of lipid sphere induced inhomogenous broadening of R2.
This R2 calibration curve can also be compared to other studies in the literature that have found that the relationship between R2 and LIC follows a curvilinear relationship, rather than a linear one [32, 53]. A likely explanation for this is the small range of LIC in Paper I. The curvilinear behavior in [32, 53] was shown over a LIC range of up to 50 mg/g, which is much larger than the 4 mg/g in Paper I.

4.3 R2* vs R2

An upside of the R2 method [32] is that it exists as a commercial service, where images are acquired and sent to an external company (FerriScan®, Resonance Health, Australia) for analysis. This procedure has the advantage of being easy to use and implement, as not much extra work is required. This is particularly useful for a hospital with only a limited number of cases. The downsides, on the other hand, are that the external analysis is expensive, and that the image acquisition is long [32]. Furthermore, if an R2 method is to be set up in-house, the curvilinear nature of the calibration curve could complicate things as it would require more data to be accurately determined, compared to a simpler linear curve.

In the literature, the R2* method is more popular than R2. There are several reasons for this. First, calculating R2* is quick, as an R2*-map with full liver coverage can easily be acquired in a single breath hold. Second, the linear relationship between R2* and LIC makes the calibration curve easier to establish. Third, R2* mapping can easily be combined with measures of liver fat, as most imaging-based methods for liver fat quantification estimate R2* in order to correct for signal decay (see section 3.2). This means that both fat and iron content can be estimated in a single breath hold, which is good for implementation in a clinical workflow.
A limitation of the combined fat and iron measurement is that the TEs of such sequences are typically optimized for fat quantification, which often uses longer TEs than is optimal for iron quantification. This can lead to issues with noise. To what extent that this is a problem can depend on the application. The problem would be greater if the purpose of the iron measurement is quantifying LIC in a patient group known to have very high levels of LIC. This could be the case for example when monitoring the response to therapeutic phlebotomy or iron chelating therapy in patients receiving blood transfusions [6, 66]. On the other hand, a longer TE may not be a problem if the patient group being studied is not expected to have severe iron overload. This could be the case, for example, when assessing if NAFLD patients, or other patients with different chronic liver diseases, have concomitant iron overload [49, 65, 67]. In such cases the longer TEs could be enough.

Can MR replace biopsies for LIC measurements? The results in Paper I indicate that if the question is to rule out iron overload due to, e.g. elevated ferritin levels, R2*, is better than a biopsy used for histopathology. As with grading of fat content it can be difficult for a human reader to make a quantitative grading. Additionally, since iron needs to be stained [7], there is also the possibility that the level of staining could affect the human grading. Alternatively, if the biopsies are used for quantitative LIC measurement, e.g. by ICP-SFMS, MR measurements may not be better, since the quantitative biopsy analysis can be considered close to a true gold standard, although, there is still an element of sampling variability. However, the results in this thesis, as well as the extensive literature on MR-based iron-quantification, show that both R2* and R2 can be accurate and reliable biomarkers. However, as calibration curves have not been identical between studies, there is a need for consistency in terms of acquisitions parameters and post-processing at each site.
5 FIBROSIS & INFLAMMATION

There has been extensive research into developing MR-based methods for staging fibrosis. The most common methods are magnetic resonance elastography (MRE) and T1 mapping. In the last few years, there has also been an increasing interest in staging inflammation in the liver. In part this is driven by the fact that both MRE and T1 mapping can be confounded by inflammatory activity. This creates a need to be able to separate the two effects. Moreover, staging inflammation is also interesting in itself, since inflammation is thought to precede fibrosis development.

5.1 Magnetic Resonance Elastography

MRE measures mechanical properties (e.g. stiffness, elasticity, viscosity) of tissue. This has turned out to be useful in CLD as the process of fibrosis development, with liver parenchymal cells being replaced by collagen containing scar tissue, increases liver stiffness.

5.1.1 MRE-Techniques

Performing an MRE-experiment consists of three principal steps (Figure 5.1): i) introduce motion into the liver; ii) image the motion; iii) calculate the mechanical parameters.

![Figure 5.1 Principle of an MRE experiment. First, shear waves are introduced in the liver, with the scanner acquiring special phase-contrast images simultaneously. Second, the images are used to visualize traveling through the liver. Third, the wave information is used to calculate parametric maps of the mechanical properties.](image)

5.1.1.1 Introducing Motion into the Liver

In order to induce shear waves in the liver, a form of vibrating device, called a transducer, is strapped to the patient’s ribcage. When performing liver MRE in human subjects, a frequency of 60 Hz is typically used, at least for fibrosis staging [68], while a slightly larger range of frequencies have been used in exploratory investigation inflammation staging [69, 70].

There are a range of different methods and transducers, based on different physical principles. The most common, and commercially available, method is to use an acoustic transducer (Figure 5.2). The acoustic
transducer has two parts, an active driver, placed outside of the scanner room, and a passive driver, which is strapped to the patient inside the scanner. The active driver is a loudspeaker, producing airwaves at a specific frequency. Those airwaves are transferred through a plastic tube into the scanner. The tube connects to the passive driver, which is a plastic membrane that bulges out due the airwaves, thus vibrating the ribcage.

Figure 5.2 Schematic description of the set-up of MRE with an acoustic transducer. Airwaves are generated by an active transducer, placed outside the scanner room. The airwaves are transferred through a plastic tube to the passive driver, which is placed over the liver, on the ribcage.


An alternative method is to use an electromagnetic transducer (Figure 5.3) [71]. When using the electromagnetic transducer, an active driver is strapped to the right side of the patient, inside the scanner. In this case, the active transducer is essentially an electric motor that drives a piston, which vibrates against the ribcage. The motor is driven by an alternating electrical current with the desired MRE frequency, generated by a signal generator outside the scanner room.
Figure 5.3 Electromagnetic transducer. (A) The electromechanical transducer is strapped to the right side of the patient’s ribcage. The transducer is essentially an electric motor that drives a piston (white circle). (B) The transducer is driven by a signal generator (top), together with an amplifier (bottom), which are placed in a cabinet outside the scanner room. The signal generator is synchronized with the MRI data acquisition. The cabinet also contains an oscillator (middle) for displaying the electric signal.

A third method is to use a gravitational transducer [72] (Figure 5.4). The gravitational transducer consists of a plastic box, which is strapped to the patient. The plastic box contains a rotating off-centric mass. As the off-centric mass rotates, the entire transducer will start vibrating. The rotation is driven by a flexible axis. The axis itself is rotated by a stepper motor, which is located outside the scanner room.
5.1.1.2 Imaging Motion in the Liver

The motion of the liver is imaged using special phase-contrast sequences [73-76]. These sequences are created through modification of basic sequences by adding motion-encoding gradients, which have to be synchronized to the vibration of the transducer. These gradients encode the motion of the moving liver protons into the phase of the MR signal. These phase images can then be used to visualize the shear waves (Figure 5.1 Middle).

For the image readout, both spin echo and gradient echo sequences can be used [77]. Gradient echo sequences are the most common, although, they can be more sensitive to iron overload [78], as the iron overload leads to a decreased T2*, which creates issues with low SNR.
5.1.1.3 Calculating Elastograms
From the images of the shear waves, special algorithms are used to calculate elastograms, which are parametric maps of the mechanical parameters (Figure 5.1 Right) [79]. The early algorithms only performed the calculations in 2D. This assumes that the shear waves only travel along the plane of the imaged slice and not between slices. Moreover, the 2D method only provides shear stiffness. If a 3D method is used on the other hand, the complex shear modulus, $G^*$, can be estimated:

$$G^* = G' + iG''$$

where $G'$ is the storage modulus, corresponding to elastic effects, and $G''$ is the loss modulus, corresponding to viscous effects. Stiffness, is the absolute value of the shear modulus, $|G^*|$. In other words, when a 3D sequence is used, the effects of elasticity and viscosity can be separated.

5.1.2 MRE and Fibrosis
The ability of MRE to detect liver fibrosis has been studied extensively for over ten years [80]. In general, most studies have shown that measuring stiffness at a frequency in the 50-60 Hz range can identify patients with advanced fibrosis (F3-F4). Typically, studies where MRE has been compared to biopsies have shown an area under the receiver operating characteristic curve (AUROC) around or above 0.9 [81]. For instance, a meta-analysis of studies on patients with HBV and HCV gave an AUROC of 0.97 and 0.94 respectively [82] for detection of advanced fibrosis. Another meta-analysis focusing on NALFD found a AUROC of 0.9 [83].

5.1.3 MRE and Inflammation
It is not only fibrosis that can affect the mechanical properties of the liver, so can inflammation. For instance, Shi et al. and Ichikawa et al. both performed MRE on patients with viral hepatitis and found that patients who had the same fibrosis stage, but a higher inflammation stage, tended to have a higher stiffness [84, 85].

Following these studies, a range of different MRE-based approaches have been suggested for staging inflammation. These approaches have involved both the use of different frequencies as well as calculating different mechanical parameters. In a first example, Yin et al. used several different animal models of liver disease to show that $G''$ and the damping ratio, $\xi = G''/(2G')$, increase in the liver during the early stages of liver injury, before fibrosis could be detected by histopathology [86]. In another example, Allen et al. continued the previous work by studying NAFLD in bariatric surgery patients [70]. Their results suggested that PDFF, stiffness at 60 Hz, and damping ratio at 40 Hz, combined in a statistical model, can predict NASH. They defined NASH based on a pathologist’s impression and reported an AUROC of 0.85. In a third example, Yin et al. tried to predict NAS...
in a mouse model [87]. They also used a statistical model, this time consisting of PDFF, stiffness and loss modulus.

5.2 T1-Mapping

5.2.1 T1-Mapping Techniques
After an RF-pulse, the magnetization along the z-axis (Mz) will start to recover toward the equilibrium (M0) according to:

\[
Mz(t) = M0 - (M0 - Mz(0))e^{-t/T1}
\]  
(5.2)

where Mz(0) is the magnetization directly after the RF pulse and T1 is the time it takes for 63 % of the magnetization to recover (Figure 5.5).

![Figure 5.5 Example of T1-relaxation time. The figure shows how the magnetization (T1 = 400 ms) along the z-axis regrows toward M0 after a 90° RF-pulse (Mz(0) = 0). T1-relaxation time is the time it takes for 63 % of the magnetization to regrow.](image)

The gold standard method for quantifying T1 is an inversion recovery (IR) experiment. In an IR experiment, Mz is inverted by an 180° RF pulse and several images are at different inversion times (TI), i.e. time points after the inversion. The signal intensity from the different images can be fitted to equation 5.2 in order to estimate T1. However, a drawback with the IR method is that it is time consuming as the magnetization has to be inverted before each individual k-space line is acquired (as Mz will be affected by the readout of the images). Additionally, TR also needs to be sufficiently long to allow Mz to relax back to M0 before each inversion.

There are alternative and faster methods for quantifying T1. One such method was used in Paper II. In the paper, T1 was estimated using a phase sensitive IR (PSIR) sequence [88] using two TIs (Figure 5.6). Additionally, this 2D method only performs the inversion once per slice and the entire slice is acquired using a single turbo field echo readout with a short TR, at each TI. This means that equation 5.2 cannot be used since the Mz does not relax by T1 during the acquisition of the
slices, because the repeated RF pulses experienced by the spins. Instead, Mz will relax by an apparent T1, called T1*, toward an apparent M0, M0*:

\[
\frac{M0^*}{M0} = \frac{T1^*}{T1} = \frac{TR}{TR - T1^* \ln(\cos(FA))}. \tag{5.3}
\]

After the inversion, the recovery will start from M_A:

\[
M_A = M_0 \times \cos(180 \times \delta), \tag{5.4}
\]

where \(\delta\) is the inversion efficiency. From the inversion to the first TI at 500 ms, the magnetization undergoes T1 relaxation from M_A to M_B:

\[
M_B = M_0 - (M_0 - M_A) \times e^{-500/T1}, \tag{5.5}
\]

During the acquisition, for a time of T_acq = 317 ms, the magnetization undergoes T1* relaxation:

\[
M_C = M_0^* - (M_0^* - M_B) \times e^{-T_acq/T1^*}, \tag{5.6}
\]

Between the two acquisitions, the magnetization again undergoes T1 relaxation for 1000 ms:

\[
M_D = M_0 - (M_0 - M_C) \times e^{-(1000-T_acq)/T1}, \tag{5.7}
\]

---

Even though measurements of T1 in theory constitutes objective quantitative measurements, there are a number of different factors that can affect the accuracy and reproducibility of such measurement. One important such factor is the effective inversion angle, which may deviate from 180° [89, 90]. An imperfect inversion can be caused by a number of reasons, such as B1 uniformity [90], non-perfect 2D slice profiles [91], and T2 relaxation effects [92].
The effect of an imperfect inversion was assessed in Paper II by performing a phantom experiment. In short, a phantom with gel filled tubes with different T1 values was imaged using the PSIR sequence, as well as an IR sequence for reference. The IR sequence had eight TIs, meaning that any imperfect inversion in the IR sequence could be corrected for by including the inversion angle as a parameter in the fitting of the signal model. This, however, cannot be done using the PSIR method, since the use of only two TIs means that only two parameters can be fitted (M0 and T1). During post-processing, the inversion angle used in the PSIR method was manually changed in order to maximize the concordance between the measured T1 from the PSIR and IR methods. This analysis showed that the optimal inversion angle was 77% of 180°, with the correlation between the two methods shown in Figure 5.7.

![Figure 5.7](image)

*Figure 5.7 Results from a phantom showing correlation between T1 measured using our PSIR method and an inversion recovery method. The dashed line is the unity line.*

Reproduced from: Karlsson et al. (2019). Early Stage Chronic Liver Disease: T1 Relaxation Times and Liver Fibrosis. In manuscript.

Apart from the sequence used in Paper II, there are a range of other sequences that can be used to measure T1 in the liver. One common sequence is the MOLLI sequence [93]. Like the method in Paper II, MOLLI is also a method for accelerating an IR experiment, and a 2D slice of the liver can be acquired in a single breath hold [94, 95]. Moreover, there are also IR methods that quantify liver T1 in multiple slices during free breathing acquisition [96]. Finally, T1 can also be measured without inverting the magnetization, e.g. by using a variable flip angle approach, i.e. acquiring spoiled gradient echo images with different flip angles [97, 98].

### 5.2.2 T1-Mapping and Fibrosis and Inflammation

The ability to use liver T1 as a biomarker for fibrosis and/or inflammation was studied in Paper II, using the same cohort as in Paper I. Figure 5.8A-B
shows examples of T1 maps for one non-fibrotic patient with a lower T1 and one cirrhotic patient with a higher T1. On a group level, there was a weak trend between liver T1 and increasing fibrosis stage, as is shown in Figure 5.8C. A one-way ANOVA found that the difference between the mean liver T1 measurements of the different groups was borderline insignificant (p = 0.052). Figure 5.8D shows liver T1 for the different fibrosis stages grouped by inflammation. In the figure, no significant effect of inflammation on liver T1 is apparent. A two-way ANOVA did not find any significant effect of either fibrosis (p = 0.172) or inflammation (p = 0.430). Finally, a receiver operating characteristic (ROC) curve was calculated to determine the diagnostic accuracy of liver T1, with respect to diagnosing advanced fibrosis (F3-F4). The area under the curve was 0.70.

As stated in the previous chapter, iron is stored in the liver and the amount of iron can be affected by diseases. Iron in the liver does not only affect T2 and T2*, but also T1 [99]. Therefore, the LIC measurements from Paper I were included. The LIC levels of all patients are shown in Figure 5.9A. The figure shows that only two patients had a clinically significant iron overload (LIC > 1.8 mg/g) [100]. Nonetheless, even LIC levels below

![Figure 5.8 Example of a T1 map for one patient with low liver T1. (B) Example of a T1 map for one patient with high liver T1. (C) Liver T1 grouped by fibrosis. (D) Liver T1 grouped by fibrosis and inflammation. The error bars show mean and standard deviation.

Reproduced from: Karlsson et al. (2019), Early Stage Chronic Liver Disease: T1 Relaxation Times and Liver Fibrosis. In manuscript.](image-url)
what can be considered clinically significant can induce changes in liver T1 [99]. Patients were therefore stratified on LIC > 1 mg/g. This is shown in Figure 5.9B, where the liver T1 for all patients is grouped by fibrosis and LIC. The figure shows that liver T1 tends to be lower in patients with LIC > 1 mg/g. Last, a two-way ANOVA with fibrosis and > 1.0 mg/g as independent factors gave both fibrosis (P = 0.032) and LIC (P = 0.001) as significant factors.

![Figure 5.9](image)

Figure 5.9 Effects of LIC on liver T1. (A) Liver iron concentration (LIC) in all patients in the cohort. Dashed line indicates LIC > 1.0 mg/g dw. (B) Liver T1 grouped by fibrosis and LIC greater or lower than 1.0 mg/g dw. Liver T1 grouped by fibrosis and inflammation for patients with LIC > 1.0 mg/g dw. The error bars show mean and standard deviation.

Reproduced from: Karlsson et al. (2019), Early Stage Chronic Liver Disease: T1 Relaxation Times and Liver Fibrosis. In manuscript.

In a final analysis, the ten patients with LIC > 1 mg/g were excluded. Figure 5.9C shows the remaining patients grouped by fibrosis and inflammation. A two-way ANOVA did not detect any significant effect of either fibrosis (p = 0.252) or inflammation (p = 0.580). The area under the ROC for diagnosing advanced fibrosis was 0.68.

Apart from the results in Paper II, the relationship between liver T1 and fibrosis and inflammation have been examined in a number of other studies. For instance, Hoad et al. reported that liver T1 increases in patients
with advanced fibrosis [96], in a cohort of diverse CLD patients. They also found that T1 was higher in patients with higher inflammation. Finally, the study also investigated the effect of iron on T1 and found that T1 was not a reliable biomarker for fibrosis if T2* < 22.6 ms. Another study by Banerjee and colleagues presented a method for correcting T1 for the presence of iron (cT1) [94]. They showed promising results for fibrosis staging, also in a cohort of diverse CLD patients. However, there are also results suggesting that T1 has limited ability to stage fibrosis, similar to the results in Paper II. Both Pavlides et al. and Eddows et al. has reported this with respect to cT1, while at the same time showing a stronger correlation to histologic features of inflammation, such as ballooning [101, 102].

5.3 MRE vs T1

Clearly, MRE is the most validated of the MR-techniques available for determining fibrosis. MRE has been in use for over 20 years and the results of studies are fairly consistent: they show an excellent ability to detect cirrhosis in addition to a very good ability to identify patients with advanced fibrosis. These results have also been shown with different patient cohorts; from mixed cohorts with different CLDs [103, 104], to unified cohorts with only, NALFD [105] or HCB [106].

The results regarding liver T1 have been less consistent. Some studies have been really promising [94, 96], while others have been much less convincing [101, 102] (+ Paper II). It is difficult to say exactly why this is the case, although one reason could be that there were differences in the patient cohorts recruited in the different studies. It could also be that it is harder to reproduce the T1 measurement, compared to the stiffness measurements performed using MRE. It would therefore be interesting to perform multi-center studies with T1 mapping. There is also a variation in the different T1 mapping methods that have been employed. Hence, it would also be useful to directly compare different methods in order to understand how the choice of method influences the results. Nevertheless, one major advantage with T1 mapping compared to MRE is that T1 mapping does not require any additional expensive hardware that needs to be purchased and maintained by the radiology departments.

Distinguishing between lower stages of fibrosis is still challenging, even with MRE. There are probably two reasons for this. First, the amount of fibrosis in the early stages may not create very big changes in stiffness. Second, the biopsy is inaccurate, which is a well-known problem. This begs the question of whether it will ever be possible to replace biopsies for fibrosis staging. On the other hand, since the biopsy is such an imperfect “gold standard”, it is possible that the quantitative MR methods could actually be better. However, to be able to show that MRE, or T1 for that matter, is better than biopsy would require longer studies where patients can be followed
over time in order to see if MRE can predict disease progression better than biopsies. However, such studies are limited by the fact that fibrosis can take decades to progress [107]. This means that either patients need to be followed for a very long time, or that a very large number of patients are needed.

On the matter of inflammation, results with both MRE and T1 mapping have been somewhat promising. However, the research remains exploratory, and much work still needs to be done. In NAFLD, which probably the most studied CLD now, there are follow up studies, based on histology, that shows that the presence of inflammation on histology dose not predict actual clinical outcomes [107, 108]. Instead, fibrosis is the only histological factor that predict outcome. This begs the question of whether it is worth to spend lots of resources on develop non-invasive methods that are supposed to correlate with histological measures that does not predict actual clinical outcomes.
6 FUNCTION

Broadly speaking, the liver has two functions: i) to metabolize nutrients and other substances; ii) to clear substances from the blood. Therefore, any measure of liver function should be able to characterize, and preferably quantify, the liver's capacity to perform one or both tasks. There are a variety of different methods for quantifying liver function. The method used in conjunction with MRI uses gadoxetate, a liver specific contrast agent [109].

6.1 Gadoxetate-Enhanced MRI

Contrast agents are molecules that are used to increase the contrast in specific areas of medical images and are commonly used in MRI. Most MRI contrast agents consist of a gadolinium ion (Gd³⁺), which is bound to a chelating molecule. Gadolinium ions are strongly paramagnetic, and it will cause a shortening of T1 relaxation time of the water in the vicinity. This will lead to an increase in signal intensity on T1-weighted images, which in turn will cause the areas in the image where the contrast agent is situated to light up.

Gadoxetate is a liver-specific contrast agent, unlike most other extracellular contrast agents. After being injected into the blood, extracellular contrast agents will be distributed into the extracellular extravascular space (EES), but the contrast agent will not enter any cells. Eventually, all contrast agent will be cleared from the body via the kidneys. Gadoxetate on the other hand will be taken up by the hepatocytes and is mainly transported via the organic-anion-transporting polypeptide 1 (OATP1) B1/B3 (Figure 6.1). After being taken up, the gadoxetate molecules will be transported out of the hepatocytes, into the bile, a transport that is governed by the multidrug resistance-associated protein 2 (MRP2), and is an active transport, i.e. requiring ATP. To some extent, gadoxetate can also be transported back to the blood via MRP3 and MRP4. Furthermore, the efflux from the hepatocytes is slower than the uptake. This will lead to an accumulation of gadoxetate in the liver, which means that the liver will light up clearly on T1-weighted images (Figure 6.2).
Figure 6.1 Gadoxetate, a gadolinium (Gd)-containing MRI contrast agent, is taken up by the hepatocytes via the OATP1B1/B3 transport proteins. After uptake, gadoxetate is actively excreted into the bile via the MRP2 transport proteins. Gadoxetate can also to some extent be transported back to the blood via MRP3/4 transporters.

Figure 6.2 Examples of gadoxetate-enhanced images T1-weighted images. The Pre Contrast image is acquired before gadoxetate injection. During the Arterial Phase (~30 seconds), the aorta and the highly vascularized spleen lights up. The contrast agent also starts to wash into the liver via the hepatic artery. In the Portal Venous Phase (~60 seconds), gadoxetate starts to wash into the liver via the portal vein. At ~3 minutes, the gadoxetate is evenly distributed, and the signal from liver, spleen, and blood is similar. After 10-20 minutes, most of the gadoxetate has either been taken up by the hepatocytes or cleared from the blood via the kidneys. This means that the liver is brighter than the spleen and blood vessels. Gadoxetate will also be excreted to the bile ducts, making them very bright.
The principle behind using gadoxetate-enhanced MRI for liver function quantification is that the transport proteins are fundamental to the basic liver function. Therefore, if liver function is decreased, the transport capacity of OATP1 etc. will also be decreased. If the OATP1 function is decreased, less gadoxetate will be taken up by the hepatocytes. If less gadoxetate is accumulated in the hepatocytes, there will be less signal increase in the MR images.

### 6.1.1 Signal Intensity Ratios

The easiest way to analyze the gadoxetate-enhanced images is to measure signal intensity ratios. Signal intensity ratios can be calculated using images acquired at any time point, although 20 minutes after contrast injection is most common.

Several different ratios have been proposed. The most frequently used is relative enhancement (RE):

\[
RE = \frac{SI_{\text{Liver},20} - SI_{\text{Liver},0}}{SI_{\text{Liver},0}},
\]  

(6.1)

where \( SI_{\text{Liver},0} \) and \( SI_{\text{Liver},20} \) are the signal intensities in the liver before and 20 minutes after contrast injection. Another alternative is to calculate the liver-to-spleen contrast ratio (LSC) [110, 111]:

\[
LSC = \frac{SI_{\text{Liver},20}}{SI_{\text{Spleen},20}}.
\]  

(6.2)

The LSC can also be normalized (LSC_N) using the pre-contrast images [110], possibly minimizing the effects of spatially varying signal intensities:

\[
LSC_N = \frac{SI_{\text{Liver},20} / SI_{\text{Liver},0}}{SI_{\text{Spleen},20} / SI_{\text{Spleen},0}}.
\]  

(6.3)

A third alternative is to calculate a liver-to-muscle ratio instead [112]. The muscle signal intensity is often measured in one of the erector spinae muscles. Additionally, the signal intensity ratios can also be combined with the volume of the liver [113].

All the above methods consist of measuring signal intensities in T1-weighted images. As is well known in MRI, signal intensities in images are not absolute values and can be confounded by a range of factors. This could be a possible limitation to creating reliable biomarkers. An alternative strategy is to measure the actual T1 relaxation times, instead of arbitrary signal intensities [97, 114, 115]. Such a strategy could have the advantage of being less sensitive to differences in scanners, acquisition parameters, and different native T1 levels.
6.1.2 DCE-MRI

While the signal intensity ratios are simple to implement, they have the drawback that they cannot directly quantify, or separate, the gadoxetate uptake, $k_i$, and efflux, $k_{ef}$, rates. An alternative approach, which can estimate the actual gadoxetate transport rates, is to perform pharmacokinetic modeling (Figure 6.3). Unlike the signal ratio methods, such modeling typically requires a time series of gadoxetate concentrations, and thus multiple images need to be acquired. Several different modeling approaches have been introduced. These approaches vary with respect to both model complexity and structure as well as the type of input data required.

![Figure 6.3 Modelling workflow. After acquiring MR images before and after gadoxetate injection, the images are processed to get timeseries of gadoxetate concentrations. Finally, the timeseries are combined with pharmacokinetic models in order to estimate the gadoxetate transport rates, which are used as biomarkers for liver function.](image)

Reproduced and modified from: Karlsson et al. (2019), Translational Modelling Framework for Predicting Human Liver Function Based on Gadoxetate Enhanced Magnetic Resonance Imaging. In manuscript.

6.1.2.1 Patlak Model

The first approach, using what is called the Patlak model (Figure 6.4; Patlak model), uses a simple two-compartment model and estimates the gadoxetate uptake and efflux rates. The Patlak model uses an input function sampled in the blood [116, 117] or spleen [118, 119]. The model then assumes that the blood extracellular concentration in the liver is the same as the concentration in the blood, which is the same thing as assuming that the mean transit time of the contrast agent in the extracellular space is considered negligible. The Patlak model is therefore typically used when there is not enough temporal resolution in the data to properly characterize the blood flow [116, 117], e.g. when using images from standard clinical gadoxetate-enhanced MRI-protocols. Such images are typically acquired during breath-hold with enough spatial resolution to be used for radiological reading.
As implemented in Paper V, the Patlak model is described by a single differential equation:

\[
\frac{dC_{\text{hep}}}{dt} = k_i C_{\text{es}}(t) - k_{\text{ef}} C_{\text{hep}}(t),
\]

where \( C_{\text{hep}} \) and \( C_{\text{es}} \) are the gadoxetate concentrations in the hepatocytes in the extracellular space respectively, \( k_i \) is the hepatic uptake rate and \( k_{\text{ef}} \) is the efflux rate to the bile. By assuming that both the liver and spleen are highly perfused organs \[118\] and have the same extracellular Gadoxetate concentration, the extracellular gadoxetate concentration in the liver can be estimated as

\[
C_{\text{es}} = \frac{C_s}{v_{e,s}}
\]

where \( C_s \) is the total Gadoxetate concentration in the spleen and \( v_{e,s} \) is the fraction of extracellular space in the spleen, assumed to be 0.30 for humans \[120, 121\] and 0.43 for rats \[118\]. The total liver Gadoxetate concentration, \( C_l \), is calculated as
Non-Invasive Characterization of Liver Disease

\[ C_t = v_e C_{es} + v_i C_{hep}, \]  
(6.6)

where \( v_e \) and \( v_i \) are the volume fractions of the extracellular and intracellular compartments respectively. The parameters \( v_e \) and \( v_i \) are fixed at 0.32 and 0.68 [122] for humans, and 0.23 and 0.77 for rats [118]. The model then has two remaining free parameters to be optimized: \( k_i \) and \( k_{ef} \).

### 6.1.2.2 2-2CUM

The second approach (Figure 6.4; 2-2CUM) is based on perfusion MRI, which typically requires measurements of blood flow, and modeling of DCE-MRI data [123]. For DCE-MRI with gadoxetate, such models can be modified to also include hepatic uptake, and then a dual-input two-compartment uptake model (2-2CUM) is created. Thus, a model can simultaneously quantify both liver perfusion and liver function [124]. The 2-2CUM uses an impulse response, which is convolved with input data from the blood, sampled in both the aorta and portal vein (hence the dual-input name), to reflect the liver’s dual blood supply. The perfusion model has also been extended to include gadoxetate excretion [125] and transport back to the blood [126]. Furthermore, perfusion approaches typically require a high temporal resolution of the DCE-MRI data in order to be able to characterize the blood flow [123]. The 2-2CUM is essentially generalization of the Patlak model, without the assumption of the gadoxetate extracellular transit time being negligible.

In Paper V the 2-2CUM, assumes that the liver tissue consists of two compartments, an extracellular and an intracellular, with volume fractions \( v_e \) and \( v_i=1-v_e \) respectively. Gadoxetate is transported to the liver via both the liver artery and portal vein, with a total blood plasma flow, \( F_p \), and an arterial flow fraction, \( f_a \). Moreover, gadoxetate is transported from the extracellular compartment to the intracellular compartment via the uptake rate, \( k_i \), and transported out of the intracellular compartment via the efflux rate, \( k_{ef} \). The total gadoxetate concentration in the liver, \( C_i \) is defined as:

\[ C_i = F_p \left( \frac{E_i}{1 - T_i/T_e} e^{-t/T_i} + \left( 1 - \frac{E_i}{1 - T_i/T_e} \right) e^{-t/T_e} \right) \otimes c_p, \]  
(6.7)

\[ c_p(t) = (f_a c_a + (1 - f_a) c_v)/(1 - Hct), \]  
(6.8)

where the extracellular mean transit time, \( T_e=v_e/(F_p+k_i) \), the intracellular transit time, \( T_i=v_i/k_{ef} \), the hepatic uptake fraction, \( E_i=k_i/(F_p+k_i) \). The input function is the Gadoxetate concentration in the plasma, \( c_p \), consisting of the concentration in the hepatic artery, \( c_a \), and the portal vein, \( c_v \). \( Hct \) is the hematocrit level and is assumed to be 0.43, and \( \otimes \) equals the convolution operator. All in all, the model has five free parameters that are optimized: \( F_p, f_a, v_e, k_i, \) and \( k_{ef} \).
6.1.2.3 Whole-Body Model

A third approach (Figure 6.4; Whole-Body model) is to use a model that is not based on convolution with an input function [127]. This third model instead simulates coupled ordinary differential equations, describing the distribution of gadoxetate in selected compartments of the whole body. The measured outputs used to quantify the whole-body model are simulated gadoxetate concentrations in the liver and spleen [128]. As the spleen is perfused profusely, in combination with a lack of gadoxetate accumulation, it can be used as a surrogate for measurements of blood. In this way, measurements from the spleen can be used input as inputs to the Whole-Body model.

As implemented in Paper V, the model consists of three main compartments: the blood plasma, the hepatocytes and (EES), with differential equations describing the concentrations in the three compartments defined as:

\[
\frac{dC_{\text{hep}}}{dt} = k_i C_p Alb - k_{\text{back}} C_{\text{hep}} - k_{\text{ef}} C_{\text{hep}},
\]

\[
\frac{dC_p}{dt} = \left( k_{\text{back}} C_{\text{hep}} - k_i C_p Alb \right) V_l \gamma_h - C_l R C_p Alb \frac{V_p}{V_p}
\]

\[
+ \left( k_{\text{diff}} C_{\text{ees}} - k_{\text{diff}} C_p Alb \right) V_{\text{ees}} + u \frac{V_p}{V_p}
\]

\[
\frac{dC_{\text{ees}}}{dt} = k_{\text{diff}} C_p Alb - k_{\text{diff}} C_{\text{ees}},
\]

where \( C_p \) is the Gadoxetate concentration in the blood plasma, \( Alb \) is the fraction of Gadoxetate not bound to serum albumin, assumed to be 0.9 [127], \( k_{\text{back}} \) is the Gadoxetate transport rate from the hepatocytes back to the blood plasma, \( V_l \) is the whole liver volume assumed to be 1.43 L for humans [127] and 0.0119 L for rats [129], \( \gamma_h \) is the volume fraction of hepatocytes in the liver assumed to be 0.68 for humans [127] and 0.77 for rats [118], \( C_l R \) is the renal clearance rate, assumed to be 118 mL/min for humans [127] and 1.7 mL/min for rats [130], \( k_{\text{diff}} \) is the Gadoxetate transport rate between the blood plasma and the EES, \( C_{\text{ees}} \), is the EOB concentration in the EES and \( u \) is the injection of Gadoxetate.

The model simulates two different outputs. The first output is the Gadoxetate concentration in the liver, \( C_l \), consists of contributions from all of the three model states

\[
C_l = \xi (v_h C_{\text{hep}} + v_{p,l} C_p + v_{\text{ees,l}} C_{\text{ees}}),
\]

where \( v_h \), \( v_{p,l} \), and \( v_{\text{ees,l}} \) are the volume fractions of the hepatocytes, the blood plasma, and EES in the liver. In humans, they are assumed to be 0.68, 0.12, and 0.20 respectively [127]. In rats, the ratio between \( v_{\text{ees,l}} \) and
$v_{p,l}$ are assumed to be the same, while the total extracellular volume fraction is assumed to be 0.23 [118], giving volume fractions of 0.77, 0.09, and 0.14 respectively. $\xi$ is an arbitrary constant.

On the other hand, the second output, the EOB concentration in the spleen, $C_s$, only consist of contributions from the blood plasma and EES

$$C_s = \xi (v_{p,s} C_p + v_{e,s} C_{e,s}), \quad (6.13)$$

where $v_{p,s}$, and $v_{e,s}$ are the volume fractions of the blood plasma and EES in the spleen assumed to be 0.35 and 0.20 for humans [127]. Like for the rat liver, the ratio between $v_{e,s}$ and $v_{p,s}$ are assumed to be the same for rats, while the total extracellular volume fraction is assumed to be 0.43 [118], giving volume fractions of 0.27 and 0.16 respectively. The model has five free parameters to be optimized: $k_i$, $k_{ef}$, $k_{back}$, $k_{diff}$, $\xi$.

### 6.1.2.4 Signal Equations

All of the above models require data that is a measure of the contrast agent concentration. Extracting the exact contrast agent concentration from the MR images is not straightforward, mainly for two reasons: i) as stated before, the signal intensities are arbitrary units; ii) the effect of the contrast agent concentration on the signal is not linear [110].

There are several ways to tackle this issue. The easiest way is to only calculate RE and assume that is an acceptable approximation. This has been done with apparently satisfactory results [124]. Others have also used the logarithm of RE [131]. An alternative way is to use the signal equations for a gradient echo image in order to estimate the $R_1$ relaxation rate at each time point [110]:

$$R_1(t) = -\frac{1}{TR} \left\{ \ln \left[ S(t) \ast \frac{1 - E(0)}{1 - \cos(\frac{\pi}{2}) \ast E(0)} - 1 \right] 
\right. 
\left. - \ln \left[ \cos(\frac{\pi}{2}) \ast S(t) \ast \frac{1 - E(0)}{1 - \cos(\frac{\pi}{2}) \ast E(0)} - 1 \right] \right\} \quad (6.14)$$

$$E(t) = e^{-TR \ast R_1(t)} \quad (6.15)$$

$$S(t) = \frac{SI(t)}{SI(0)} \quad (6.16)$$

From this, the contrast agent concentration, $C$, can be estimated from:

$$C = \frac{\Delta R_1}{r_1}, \quad (6.17)$$

where $\Delta R_1$ is the change in $R_1$, after contrast agent injection, and $r_1$ is the relaxivity of the contrast agent. This method requires the knowledge of $T_1$ relaxation time of the tissue before injection, as well as the relaxivity of the contrast agent. $T_1$ can either be measured by acquiring a $T_1$ map [118, 125],
or T1 can be approximated from literature values [110]. Values for relaxivity is taken from the literature and can be different in different compartments of the body, such as the blood or inside the hepatocytes [127].

6.1.2.5 **Comparing Models with Human Data**

In Paper III, all three pharmacokinetic models (Figure 6.5) were fitted to human data from 35 patients with suspected chronic liver disease [128]. The liver concentration of gadoxetate of one patient is shown in Figure 6.5A, together with liver concentrations that were simulated using the three models. The 2-2CUM was fitted using either the blood or spleen signals as inputs. The figure shows that all three models were able to fit to the data.

![Figure 6.5](image)

*Figure 6.5 All models can describe human data and gives the same uptake rates with the spleen as input. (A) Example of how well the models fitted to the data from the liver from a typical patient. (B-F) Correlations between the gadoxetate uptake rates, kᵢ, determined using different pairs of models. (B) the 2-2CUM vs. Patlak-model, with input from blood; (C) the 2-2CUM - vs. the Whole-body-model, with input from blood; (D) the Patlak vs. Whole-Body models, with input from blood; (E) the 2-2CUM vs. the Patlak model, with input from the spleen, (F) the 2-2CUM*
vs. the Whole-Body model, with input from the spleen. Black lines in (B-F) show the unity line.

Reproduced and modified from: Karlsson et al. (2019), Translational Modelling Framework for Predicting Human Liver Function Based on Gadoxetate Enhanced Magnetic Resonance Imaging. In manuscript.

Figure 6.5B-C shows the uptake rate estimated using the 2-2CUM (using blood as input) correlated with the uptake rates estimated using the Patlak and Whole-Body models (both using spleen as input). The two figures show a low correlation, (R ≤ 0.5). The correlation was higher when comparing the whole-body and Patlak models (Figure 6.5D; R = 0.81). To investigate if the discrepancies in Figure 6.5B-C were due to differences in model structure or due to differences in input signal, Figure 6.5E-F shows correlation using the 2-2CUM with the spleen as input. In this case, the correlation was larger and also similar to the correlation between the Patlak and Whole-Body models.

6.1.2.6 Comparing Models with Human Data

The models were also fitted to data from rats treated with different doses of a hepatotoxic drug [118]. Figure 6.6A shows the gadoxetate concentration in the liver from one of the rats together with a fitted simulation from the three models. Note that a larger and more complete section of the uptake curve was observed in rats compared to humans. However, only input data from the spleen (not blood) was available from the rats. The figure shows that all three models were able to fit to the data. Figure 6.6B-D shows the correlations between the uptake rates estimated using the different models.

The three figures show that the correlations were similarly large when comparing all three models, although the whole-body model did underestimate the uptake rate (Figure 6.6B). The correlation between the 2-2CUM and Patlak model was good, although the 2-2CUM overestimated the highest uptake rates (Figure 6.6C).
6.2 Gadoxetate-Enhanced MRI in Liver Disease

6.2.1 Correlation with Other Liver Function Measures
The ability of gadoxetate-enhanced MRI to measure liver function has been studied extensively. Firstly, work has been done showing that the gadoxetate uptake is lower in patient groups known to have impaired liver function [110, 132], and also showing correlation to other clinical markers of liver functions, such as Child Pugh class [97, 133], MELD score [134], and ICG retention rate [135].

Reproduced and modified from: Karlsson et al. (2019), Translational Modelling Framework for Predicting Human Liver Function Based on Gadoxetate Enhanced Magnetic Resonance Imaging. In manuscript.
6.2.2 Liver Surgery
Another section of the literature has focused on using gadoxetate-enhanced MRI to predict which patients are at risk of suffering from liver failure after resective surgery. Today, there are few very accurate liver function measures available to liver surgeons. The most common methods include measurement of the future remnant liver volume, i.e. the volume of the liver left after surgery, although, this does not consider the actual functional capacity of liver.

Asenbaum et al. recently published a study showing that a measure derived using RE was significantly better than only using liver volumes [136]. There are also a number of other similar studies. However, they have had difficulty proving the usefulness of gadoxetate-enhanced MRI since the number of patients suffering liver failure is small, thus leaving most studies underpowered [137-139].

6.2.3 Liver Fibrosis
A significant body of research exists comparing gadoxetate-enhanced MRI to liver fibrosis. This was partly done in Paper III, where the Whole-Body model was applied to a cohort of 91 patients undergoing liver biopsy due to suspected CLD. This was essentially the same cohort as in Papers I and II.

Figure 6.7A shows that ΔR1 (proportional to gadoxetate concentration) is lower in patients with increased fibrosis stage. Meanwhile, Figure 6.7B shows that ΔR1 appears to be unaffected in the spleen. Furthermore, the uptake rate of gadoxetate (called \( k_{ph} \) in Paper III), differentiated between fibrosis stages, is shown in Figure 6.7C. Figure 6.7D also shows the hepatocyte efflux rate (called \( k_{hb} \) in Paper III) differentiated between fibrosis stages.
A few observations can be made from Figure 6.7. To begin with, liver function is lowered in patients with advanced fibrosis (F3 and F4). This is visible both when looking directly at the $\Delta R_1$ time series, as well as the uptake rate. These results are similar to what has been found in studies estimating the uptake rate via modeling [116, 140] as well as studies using signal intensity ratios [141-143]. Additionally, there may be an increased efflux rate in cirrhotic patients (F4). However, this is less clear in the literature as there are studies looking at DCE-MRI as well as gene expression of MRP2 that both support [144, 145] and contradict these results [126, 146, 147]. Finally, the gadoxetate-enhanced MRI derived measures are not very useful biomarkers for fibrosis. In an ROC analysis, made with the uptake rate from Paper III with respect to detecting advanced fibrosis, the AUROC is only 0.76. That is not very impressive when compared with MRE, which in many studies shows an AUROC close to or above 0.9 [104, 105, 148].

6.2.4 Preclinical Studies

Gadoxetate-enhanced MRI can also be used to study liver function in different preclinical animal models. For example, there examples are studies using animal models to investigate the same questions as in diagnostic human studies, such as animal models of liver fibrosis [131, 149]. Other examples are studies using animal models and gadoxetate-enhanced MRI to investigate different biological mechanisms effecting OAPT1 and MRP2 function, such as a recent study showing how OATP1 function decreases in a mouse model of diabetes [150].

Several different studies have investigated modeling of gadoxetate-enhanced MRI as a biomarker for DILI. Initially, Ulloa et al. used the Patlak model to study the effects of a hepatotoxic drug, [118] and this approach was later applied and verified in a multi-center study [119]. A similar approach was also tested by Georgiou et al. [151].

6.3 Association between Gadoxetate-Enhanced MRI and MRCP

A possible use for gadoxetate-enhanced MRI would be to use the functional biomarkers to follow up disease progression. Ideally, these biomarkers should detect early changes in liver function and be able to predict which patients are at risk of serious complications. A study indicating such utility was recently published by Schulze et al. [152]. They prospectively recruited over 100 PSC patients, who were given a gadoxetate-enhanced MR examination, measuring RE. Using an RE cut-off of 0.65, they were able to predict clinical endpoints, such as cholangiocarcinoma, liver transplantation, and death, with an AUROC of 0.91.

Paper IV continued examining the utility of gadoxetate-enhanced MRI in PSC. In the paper, 26 patients who are part of a five-year follow-up study were included. The follow-up study included annual MR examinations. The MR examination itself included gadoxetate enhanced-images, as well as MR cholangiopancreatography (MRCP). MRCP is commonly used for detecting macroscopic changes in the biliary tree, such as stenoses of tumors.

The first part of Paper IV examined how gadoxetate-MRI based liver function measures vary longitudinally in patients. This is interesting since it is important to understand the accuracy of a biomarker. Ideally, this would be studied by simply repeating the measurement several times. However, this is problematic with gadoxetate, due to recent concerns over possible gadolinium deposition in tissue from gadolinium-based contrast agents [153, 154]. Even though there have been no reported clinical conditions as a result of this, it is still ethically challenging to give patients or volunteers repeated doses of contrast agent without a sufficiently strong clinical indication. For this reason, the cohort in Paper IV provided an opportunity to study the biomarker variation, as it was a group of well-
characterized patients undergoing repeated standardized gadoxetate-enhanced examinations. Moreover, to ensure that the variation in measurements was only due to the measurements or normal physiological variations, patients who were considered to have clinically progressed were excluded from this analysis. Also, patients who only had a single gadoxetate-enhanced examination were excluded. In the end, 16 patients remained with between two and five examinations.

Liver function measures were calculated for each examination. First, RE at 20 minutes was calculated. Second, the gadoxetate uptake and efflux rates were estimated using the Whole-Body model. The Whole-Body model was chosen as the MR protocol only contained a limited number of images (pre-contrast, arterial and venous phases, 3, 5, 10, 20, 30, and 40 minutes). Additionally, several patients had several missing data points.

The results are shown in Figure 6.8. Figure 6.8A-C shows how the three different liver function measures vary between the examinations, while Figure 6.8D-F shows the different measures grouped by patient. From the figure, two observations can be made. First, the variation is smaller for RE than for the uptake rate, and the variation for uptake rate is smaller than for the efflux rate. Second, the span of the individual patients’ longitudinal values is smaller than the mean of the inter-patient spread. This means that a decrease in patient liver function could be detected, assuming that differences between patients reflected actual differences in liver function.

Paper IV also investigated whether there were any correlations between the liver function measures and macroscopic changes in the biliary tree (stenoses), as detected using MRCP. Two radiologists evaluated MRCPs from first examination of all 26 patients. Each of the liver segments was given a score, depending on the number and severity of the stenoses affecting it. These scores were then correlated with RE, and with the uptake and efflux rates.
Figure 6.8 Longitudinal variation of liver function measures. (A-C) show how RE, gadoxetate uptake rate, and gadoxetate efflux rate varies over time for each patient. (D-F) show RE, gadoxetate uptake rate, and gadoxetate efflux rate grouped by patient. Note that the patients in (E-F) are in the same order as in (D).


Figure 6.9A-C shows the average MRCP score, averaged over all liver segments, correlated with RE, and with the uptake and efflux rates. The figure shows that there are no correlations. Moreover, to investigate whether there was a stronger correlation at a regional level, e.g. if stenoses in the left lobe would cause a regional decrease in function, Figure 6.9D-E shows RE from the right and left liver lobes correlated with the scores from those lobes. However, there were still no correlations.

In summary, two main conclusions can be drawn. First, gadoxetate-enhanced MRI is certainly promising for evaluating disease progression in PSC. Schulze et al. elegantly showed how RE could be used to predict solid clinical outcomes [152], and Paper IV showed that RE is indeed a reliable measurement. Second, the macroscopic findings on MRCP do not seem to correlate well with liver function. This could be interpreted as indicating that MRCP should not be used to assess parenchymal disease progression, despite being useful for diagnosing PSC and screening for tumors [155]. It could be that the parenchymal progression is mainly driven by microscopic changes, e.g. in the smaller bile duct, that are not visible on MRCP.
Figure 6.9 Correlation between liver function measures and biliary tree changes. (A-C) show correlations between the scoring from the MRCP images and RE, uptake rate, and excretion rate respectively. The score is the average of the total score calculated for each liver segment. (D-E) show RE correlated with average intrahepatic score from the left and right liver lobes.


6.4 Translating Liver Function Between Humans and Rats

The gadoxetate uptake and efflux have been studied and modeled in both humans and rats (Figure 6.10; yellow). The ability of these pharmacokinetic parameters to serve as biomarkers of DILI has also been studied (Figure 6.10; blue). However, there are no methods that allow for the translation of pharmacological effects observed in rats to realistic clinical situations and personalized treatments in individual human subjects (Figure 6.10; green).
Figure 6.10 Previous research has focused on how liver function and gadoxetate pharmacokinetics can be described in humans as well as in rats (yellow arrow). Other research has also studied how liver function can be affected by drugs (blue arrow). However, the clinically interesting question of how a given drug affects liver function in humans is much more difficult to answer. We therefore here introduce a translational modelling framework in order to predict how a given drug can affect liver function in humans (green arrow).


6.4.1 Estimating the Drug Effect
The question of how to translate the drug effects on gadoxetate uptake from rats to humans was approached in Paper V, with data from Ulloa et al. consisting of rats undergoing DCE-MRI after being given a hepatotoxic drug [118]. To begin with, the effects of the drug on the gadoxetate influx and efflux rates were quantified using the Patlak model in combination with nonlinear mixed-effects modeling [156], where a dose-response function was determined for both uptake and efflux rates. The dose-response function describes how pharmacokinetic parameters depend on the dose of the drug. Figure 6.11 shows the estimated parameters (ki and kef) for each rat at the given drug doses, as well as the estimated dose-response function.
6.4.2 Translation

As Patlak model has been extended to include the dose-dependent effect of a drug (Figure 6.12A), it can be used to predict how the liver function (gadoxetate uptake) would be affected in humans (assuming that the dose-response of the drug is the same in both humans and rats). Fig 8C,E shows examples of dose-response curves for three patients (P1-P3), taken from the cohort in Paper III, each having different gadoxetate uptake rates without the drug. As is shown, P1 has a much higher uptake rate (4.1 ms\(^{-1}\)) than P2 (2.5 ms\(^{-1}\)) and P3 (1.4 ms\(^{-1}\)) (Table 6.1 Column 2). Using the dose-response curve already obtained (Fig 6), it is possible to predict how these uptake rates would decrease with increasing drug doses (Figure 6.12C,E).

The drug effects can also be compared with the normal ranges of uptake rates. These normal ranges can be obtained from Paper III (Figure 6.12D), where the gadoxetate uptake rate was estimated in patients with and without liver fibrosis. The dose where the uptake rate falls below the normal limit, seen in patients without fibrosis, can be considered as the maximal allowed dose (Figure 6.12C,E; Table 6.1 Column 3). In other words, that dose is the highest that can be given before liver function is compromised beyond normal levels in an individual patient. This maximal allowed dose of the drug should then be compared to the minimal required dose, i.e., the lowest dose that must be given for the drug to have its required efficacy. Figure 6.12F shows the shape of a hypothetical dose-efficacy curve for our studied drug (Table 6.1 Column 4). A drug should only be given if the
minimal required dose is smaller than the maximal allowed dose, and only in that dosing range. For our studied drug, and the three patients, P1 and P2 satisfied these conditions, whereas P3 needs to be treated using a more effective drug (Table 6.1 Column 5). Last, Figure 6.12B shows how the uptake of gadoxetate would change if one of the human patients (P1) were given the same drug as the rats at a range of different doses.
gadoxetate would change if one of the human patients were given the same drug as the rats, assuming that the translational dose-response of the drug was the same in both humans and rats (which may or may not be correct). The different colors of the lines correspond to different doses of the drug. (C, E) Examples of a dose-response of the uptake rate for three different individuals, with different baseline uptake rates. The panels show that there are individual doses at which the uptake rates would fall below the range of normal liver function, resulting in a different maximal allowed dose for each patient. (D) The range of normal liver function is taken from [128] where normal function is defined as the uptake rate in patients without fibrosis. (F) A schematic illustration of how the efficacy of a drug is related to the translational dose. Typically, there is also a minimal required efficacy of the drug that is required for therapeutic effect; consequently there is also a minimal required dose. For the drug to be therapeutically useful, this minimal required dose needs to be lower than the maximal allowed dose (with respect to the liver function).

Reproduced and modified from: Karlsson et al. (2019), Translational Modelling Framework for Predicting Human Liver Function Based on Gadoxetate Enhanced Magnetic Resonance Imaging. In manuscript.

### Table 6.1 Illustration of suggested translational drug dose in humans.

<table>
<thead>
<tr>
<th></th>
<th>Current Uptake Rate</th>
<th>Maximal Allowed Dose</th>
<th>Minimal Required Dose</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>4.1 ms(^{-1})</td>
<td>44 mg/g</td>
<td>28 mg/kg</td>
<td>OK to use</td>
</tr>
<tr>
<td>P2</td>
<td>2.5 ms(^{-1})</td>
<td>35 mg/kg</td>
<td>28 mg/kg</td>
<td>OK to use</td>
</tr>
<tr>
<td>P3</td>
<td>1.4 ms(^{-1})</td>
<td>21 mg/kg</td>
<td>28 mg/kg</td>
<td>Need another drug</td>
</tr>
</tbody>
</table>


### 6.4.3 Limitations

The procedure described above should not be seen as complete guide on how to examine the safety of a drug. It should, instead, be seen as proof of concept on how to model the effects of a drug on liver function. Additionally, several aspects of the translation were beyond the scope of Paper IV, aspects which nonetheless needs to be considered if these methods are to be implemented in practice. The first aspect is whether the drug interacts with human hepatocytes in the same way as the rat hepatocytes, which was assumed in Paper V. This may not be entirely accurate and should therefore be controlled for by the use of in-vitro studies. The second aspect is whether the pharmacokinetics of the drug are the same in both rats and humans. Again, this this was assumed, but may not be accurate. Nevertheless, today such pharmacokinetic translations of a drug are a standard part of drug development pipelines.
6.5 Future of Liver Function and Gadoxetate-Enhanced MRI

As stated above, the role of quantitative measures from gadoxetate-enhanced MRI as biomarkers for liver fibrosis is limited. This is because fibrosis and function are two different aspects of liver disease and they only overlap in the most severe stages of fibrosis. Instead of focusing on fibrosis in CLD patients, the focus should rather be on investigating if and how the gadoxetate-enhanced MRI relates to clinical outcomes, such as the progression to decompensated cirrhosis, or even death. However, investigating this would require a shift away from cross-sectional studies, which show correlations with other liver function parameters, to longer follow-up studies. Follow-up studies are rarer, but an good example was given by Schulze et al., who showed how RE could predict clinical outcomes in PSC [152]. This, along with the results from Paper IV, showing that RE is a stable measurement, should be encouraging for further such studies in more patient groups.

Another promising application of the liver function measurement is to assess the risk of liver failure after resective surgery. This would be an elegant use of the method, as the liver function measurement could easily be included in a clinical workflow as part of a preoperative MRI-protocol. As described above, there are promising results, although, several studies have been underpowered, since only a few patients actually suffer from liver failure. The notion that few patients suffer liver failure does not necessarily mean that potential liver failure is not a serious problem. It should rather be seen as a consequence of surgeons being conservative, only operating on patients they are certain will not suffer from liver failure. If better methods for liver function evaluation were available, surgeons might be willing to operate less conservatively and on more challenging patients.

Gadoxetate-enhanced-MRI may also have an interesting future for use in drug development. First it can be used in animal models to screen possible new drugs to determine if they might induce DILI in the experimental situation. Second, the translational modeling framework presented in Paper V might also be used for understanding how the drug effects seen in rats will translate to humans. Third, gadoxetate-enhanced MRI could also be used clinically to evaluate patients if there is a suspicion of DILI, or there is a question of whether a patient has limited liver function, as well as in early clinical trials.

There are a range of different proposed methods for analyzing the gadoxetate-enhanced images. The methods range from simple signal intensity ratios on standard T1-weighted images, to T1 mapping, to pharmacokinetic modeling. However, few studies have at present compared the different methods, meaning that there is little data on whether the more advanced methods provide more clinically useful information. This is an
important factor to consider if these methods are to be introduced into clinical practice.

Using advanced pharmacokinetic modeling offers the possibility to extract more information than simple signal intensity ratios. However, modeling may be more difficult to implement, particularly at sites without dedicated specialist competence. Typically, modeling methods also require more images to be acquired, which can take more time. Taking more time is in itself always disadvantageous, especially as there is general trend to implement shorter abbreviated clinical MRI protocols. Such protocols can sometimes only contain a single hepatobiliary phase gadoxetate-enhanced image, with the contrast agent being injected prior to examination [157, 158]. With such imaging protocols, only the signal intensity ratios can be used. On the other hand, such abbreviated protocols, where no early phases of gadoxetate-enhancement are acquired, may leave time for a perfusion style image acquisition to be performed [124, 159]. If early phases are required on the other hand, the high temporal resolution of the perfusion imaging may limit the spatial temporal resolution of the images, making them unsuitable for clinical morphologic assessment. However, a solution for this might be found in recently introduced imaging methods based on compressed sensing, such as GRASP [160, 161]. Such methods continuously acquire k-space data during free breathing and images can be reconstructed retrospectively, with either high temporal resolution or high spatial resolution.

The signal intensity ratios can more easily be included in shorter clinical MRI protocols. Even more important, if a measure that does not require a pre-contrast image, such as LSC, is used, the liver function measure can also be included in the abbreviated protocols, without any modification. Furthermore, it is also not yet clear if using T1 mapping has any practical advantage over only using T1-weighted images. Testing that should be easy, since most studies of gadoxetate and T1 mapping probably already include standard T1-weighted images as part of the clinical protocol.
7 SUMMARY & FUTURE OUTLOOK

Will multimodal quantitative MR replace liver biopsies? Maybe in some situations.

First of all, there are more reasons for performing a liver biopsy, other than quantifying fat, iron, inflammation and fibrosis. For instance, a biopsy could be required for making a clinical diagnosis, in case the patient presents with atypical features. With respect to fibrosis, there is also an additional and important role for performing biopsies, as the correlation between histology and the MR-based fibrosis biomarkers is still not perfect for a number of reasons. However, as discussed in chapter 5, it is possible that this imperfect correlation occurs because histology is such a bad reference ‘gold’ standard.

Additionally, there are other competing non-invasive methods for staging fibrosis. Perhaps the most widely used method is vibration controlled transient elastography (VCTE), which is an ultrasound based elastography method [162]. The advantage of VCTE is that it is cheaper and simpler than MR and can be performed by hepatologists themselves. Even though VCTE does not perform as well as MRE [163, 164], VCTE still has a very high negative prognostic value for advanced fibrosis [165, 166]. In clinical practice, being able to exclude advanced fibrosis, at a low cost, may be sufficient. This would be especially true for diseases such as NAFLD, where the progression from mild fibrosis to severe liver disease can take decades [107].

Quantification of iron is an area in which MRI is a better alternative to biopsies. Paper I showed that measuring R2* was even better than histology for quantifying LIC. If a patient is presenting elevated serum ferritin, which is a common blood test, that may suggest elevated LIC, it is probably better to perform an MRI, rather than performing a biopsy.

Quantification of fat is another area where MR can be able to replace biopsies, particularly since MRS could be considered to as close to a gold standard there is for measuring liver fat in-vivo. However, the question is whether there is a clinical need for only quantifying liver fat. On the other hand, MR would be an excellent modality for quantifying liver fat in research projects.

In general, clinical research projects is an application where multimodal quantitative MR can really be useful. A quantitative MR-examination can easily provide excellent measures of fat and iron content from a single breath hold scan. If MRE is also included in the examination, a good measure of fibrosis and possibly also inflammation can be acquired in a few
breath holds. Additionally, if liver specific contrast agents are used in the patients, there is also the possibility to estimate liver function, both in clinical studies, as well as in drug-development.
8 BIBLIOGRAPHY


82. Singh S et al: Diagnostic performance of magnetic resonance elastography in staging liver fibrosis: a systematic review and meta-analysis of


97. Kim JE et al: T1 mapping for liver function evaluation in gadoxetic acid-enhanced MR imaging: comparison of look-locker inversion recovery and


108. Ekstedt M *et al*.: Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology* 2015, 61(5):1547-1554.


142. Feier D et al: The diagnostic efficacy of quantitative liver MR imaging with diffusion-weighted, SWI, and hepato-specific contrast-enhanced


Papers

The papers associated with this thesis have been removed for copyright reasons. For more details about these see:

http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-162653