



Quantification of Hepatic Iron Overload using MRI

Andrea Brigliadori¹

¹Computer Science, University College London

Abstract

This work focuses on quantifying iron overload in the liver from magnetic resonance images. Measuring the amount of iron in the body is essential for the treatment of patients suffering from thalassemias, a class of widespread genetic diseases compromising the production of red blood cells. Without frequent blood transfusions, many of these patients do not reach the age of five. In turn, transfusions may lead to toxic accumulations of iron, potentially inducing severe health complications such as cardiac failure. In order to prevent this, it is critical to remove the excess iron in the body. At present, the only effective method is to use medications able to bind iron and assist in its excretion. However, it is necessary to tailor the length and intensity of this therapy to the specific patient, hence clinicians need to know the amount of surplus iron to remove. To approach this problem, computational models have been developed and applied on magnetic resonance images to estimate liver iron overload, which has been shown to closely relate to the total body iron content. In this study, we contribute to the field with an in-depth analysis of one of these methods, obtaining promising estimation results.

Key words: Thalassemias, iron overload, liver, T2*-weighted MRI

Background

At the heart of a successful application of computational models to medical problems is an intuitive understanding of the underlying biology. This section is our attempt to convey the basic concepts to the reader, gradually introducing the techniques we can use to accurately quantify liver iron overload in order to correctly inform treatment.

Thalassemias

Haemoglobin is a protein located in red blood cells, primarily responsible for the transportation of oxygen from the lungs to the tissues [45]. This important role is made possible by its structure comprising four globin chains (proteins), each containing an ion of iron (Fe^{2+}) to which oxygen can bind. Alterations in the production of these microscopic molecules may have a giant impact on the host organism. One example of such alterations is represented by thalassemias, genetic disorders resulting from a decreased or absent synthesis of a globin chain [28]. Globin chain imbalances compromise the production of healthy red blood cells, resulting in chronic anemia. This persistent deficiency is the primary reason why patients with severe thalassemia depend on regular and lifelong blood transfusions to survive. Thalassemias pose a major public health challenge, representing one of the most common genetic diseases and affecting nearly 68,000 newborns each year [29]. Symptoms range from fatigue and weakness to life-threatening complications such as cardiac failure [29]. In the absence of transfusions, approximately 85% of patients with severe thalassemia do not survive beyond the age of five [38]. On the other hand, transfusions often cause an excessive accumulation of iron, for which the human body lacks physiological excretion mechanisms

[43]. Elevated concentrations of iron are toxic, and potential outcomes include heart failure, cirrhosis, diabetes mellitus, infertility, and neurodegenerative diseases [10, 40]. Today, the average life expectancy of patients with severe thalassemia remains around 17 years, with cardiac complications from iron overload being the leading cause of death [28]. A therapy known as iron chelation uses medications able to bind the extra iron so that it can be removed from the body, preventing damages to the organs. Iron chelation is currently the only method proven effective in extending patients' survival [38, 40].

Liver, the body's iron warehouse

The current need for treatments such as chelation therapy is rooted in evolutionary history. Around two billion years ago, Earth experienced a period known as the Great Oxygenation Event, characterized by the accumulation of free oxygen (O_2) in the atmosphere [10, 34]. This rise in oxygen led to the oxidation of iron into a less soluble and potentially toxic ion (Fe^{3+}), which remains present in biological systems and continues to drive the need for strategies to control its accumulation, such as the one proposed in this work. In response to this environmental shift, organisms evolved proteins to safely store, transport, and regulate iron. In humans, the liver is an exemplary result of this evolutionary adaptation. Containing around 20% to 30% of the body's total iron, it serves essential functions of storage and detoxification [10, 44]. While these functions are insufficient to handle the excess iron introduced by frequent transfusions, a significant portion of the surplus iron still accumulates in the liver. In particular, liver iron concentration (LIC) has been shown to correlate reliably with total body iron stores [3, 39]. As a result, this organ remains a valuable source of information for the assessment and planning of chelation therapy. Indeed, in order

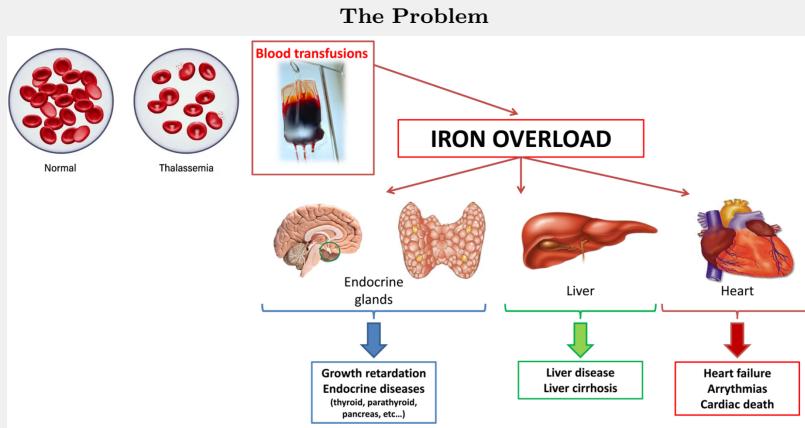


Fig. 1: Thalassemias compromise the formation of red blood cells, with potentially fatal outcomes. Consequently, often patients depend on regular blood transfusions. Nonetheless, these can lead to the accumulation of a toxic amount of iron in the body, posing new critical challenges to the patient's health. Hence, excess iron must be removed. Image adjusted from [19, 32].

to effectively remove the accumulated iron from the body, clinicians must first determine how much excess iron is present, and LIC can be considered a useful marker for this purpose. Once the amount of surplus iron is known, the intensity and duration of chelation therapy can be adjusted accordingly. This targeted approach is essential, as chelation therapy is both expensive and associated with potential side effects. The core objective of the computational methods used in this study (and several others) is precisely that of accurately quantifying liver iron overload to correctly initiate, monitor, and tailor treatment [39]. Traditionally, liver biopsy has represented the gold standard procedure for this assessment [11]. Nonetheless, biopsy is invasive, expensive, and potentially leading to complications [42]. For these reasons, methods based on magnetic resonance imaging (MRI) have emerged as an effective non-invasive alternative for estimating the iron accumulation [49].

T2*-weighted MRI

In simplistic terms, MRI techniques use a scanner that, by means of magnetic fields, can encode the value of a signal which decays in time. Depending on the particular technique, the decaying time is influenced by specific variables. For example, T2-weighted MRI is especially affected by a tissue property known as relaxation time, corrected for imperfections in the scanner. These corrections, while useful in some application (e.g. cancer detection), can suppress effects that are attributable to the presence of iron. Therefore, when considering the problem of iron overload, such corrections are generally not applied, and the resulting technique is called T2*-weighted MRI. This method allows for an accurate description of LIC and is therefore an useful resource for the early detection of iron overload. Moreover, without applying the corrections, the signal decays faster, so that the scan can be taken in a single patient breath hold. Simple and fast measurements are important not only for a correct initiation of chelation therapy, but also for a regular monitoring of its progresses. These advantages have allowed T2*-weighted MRI to receive the largest consensus as the favoured imaging technique for assessing iron overload [39]. In terms of computational modelling, T2* represents the time constant of signal decay and is generally the parameter to be estimated. Indeed, the signal decays faster for higher LIC, meaning that smaller T2* (measured in ms) indicates higher iron content. For simplicity, we will

often consider the reciprocal of T2*, denoted R2* (measured in $1/s$), since it is larger for higher iron concentration. The numerous studies that have considered T2*-MRI for the quantification of iron overload represent one more demonstration of its importance in the field.

Related work

To orient ourselves within the current research landscape, we now examine its major milestones. Many studies have assessed R2*-based techniques against liver iron concentration (LIC) from biopsy samples [12, 17, 31]. In the study by Wood et al. [48] 22 patients underwent both MRI scanning and liver biopsy. The aim was defining a mathematical relationship (known as calibration) that allows estimating LIC from measured R2* values. They accomplished this by fitting a constant-offset exponential model (C-EXP), demonstrating strong correlation between R2* and LIC, though the results were limited by sample size. Hankins et al. [15] extended this with 43 patients, using a truncated exponential model [16], and found excellent agreement with Wood's calibration curve. They also reported high interobserver reproducibility (ICC = 0.98), highlighting the robustness of T2*-weighted MRI. Christoforidis et al. [8] analysed one of the largest cohorts (94 thalassemia patients), confirming R2*'s reliability over other imaging methods. Several studies obtained similar results [2, 13, 17, 18], all using single exponential models (with or without offset or truncation), typically fitted with the Levenberg–Marquardt (LM) method [24, 26], which appears to be the most widely used in the literature. Despite broad adoption of computational methods for T2*-weighted MRI, there is no consensus on the optimal image analysis pipeline. Variability exists in the choice of model, fitting algorithm, ROI selection, and whether fitting is performed pixel-wise or on average ROI signals. Addressing this, Positano et al. [37] compared SEXP, C-EXP, and BiEXP models using synthetic ground truths, but no single model was found to consistently outperform the others. Their study also highlighted the impact of ROI variability [2, 25, 33, 37, 47, 49], an aspect that can heavily influence the estimates. A global pixel-wise approach across the whole liver was proposed and shown to alleviate the problem. Ibrahim et al. [22] similarly support pixel-level analysis as reliable when using single-exponential models. In 2023, Eldaly et al. [11] extended this work by proposing a standardised method for

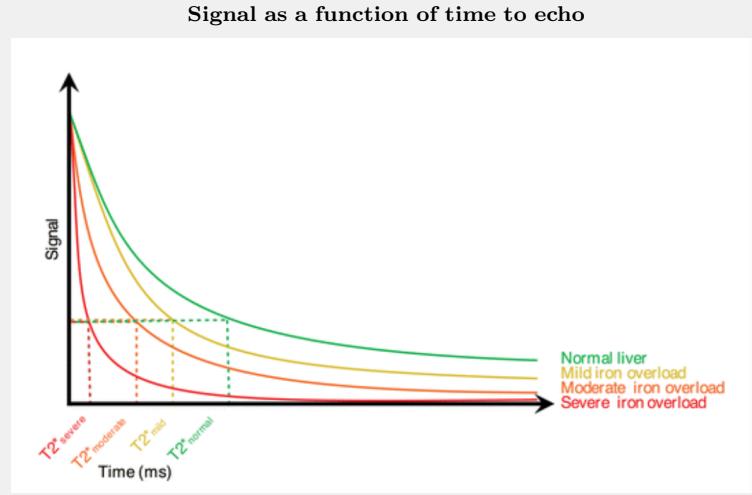


Fig. 2: This idealised picture illustrates the typical signal decay as a function of time to echo. The curve assumes a shape resembling an exponential decay. Moreover, the decay is faster for smaller $T2^*$. As faster decay means more iron, smaller $T2^*$ values suggest a more severe iron overload. Image obtained from [23].

$R2^*$ estimation. They employed exponential models fitted pixel-by-pixel and optimised via the alternating direction method of multipliers (ADMM). This strategy was also applied by Huang et al. [20] for similar purposes, and in both cases the results demonstrated superior performance compared to traditional techniques. In 2024, Huang et al. [21] proposed a further alternative for model fitting, based on a Bayesian approach. The main limitation of all these techniques remains their significant computational cost. Nonetheless, compared to traditional methods, they are typically less sensitive to noise and therefore tend to produce better estimates [21]. A particularly innovative direction was explored by Positano et al. [36], who developed a deep learning model for the unsupervised classification of LIC. Three different convolutional neural networks (CNNs) architectures were trained on augmented MRI images, and the predicted LIC values were compared against synthetic ground truth data. Their best model achieved accuracy comparable to that of expert reviewers and generalised well to an external dataset acquired on different scanners. Given its performance, this method may contribute to improve consistency and efficiency in clinical procedures, potentially supporting radiologists in their decisions.

Having established a solid overview of the field, we are now in a position to identify how to contribute to it.

Research gap

Fitting methods that include regularisation terms, such as ADMM, typically are more computational intensive than simpler techniques. Nonetheless, they have shown promising results, demonstrating good robustness against noise. Yet, the number of studies examining these methods for applications to LIC estimation from $T2^*$ -MRI images remains reduced. Also because of this, the field currently lacks general consensus on how to approach this problem. Several steps of the necessary pipeline remain subject to arbitrary choices, and a standardisation could guarantee a major ease in the use, comparison, and interpretability of the results, both for researchers and for medical practitioners. With the aim of filling this research gap, we continue on the path indicated by Eldaly et al. [11], exploring in depth the possibility of performing LIC estimation by means of ADMM fitting.

Contributions

In this study, we approach the problem of quantifying iron overload in thalassemia patients using $T2^*$ -weighted MRI. The contributions of this paper are four-fold:

- We focus on the estimation of $R2^*$, which can be used as a proxy for the amount of iron in the liver;
- We formulate the estimation task as an optimisation problem and we carefully outline the steps prescribed by the ADMM, the fitting method to be employed;
- We run experiments on synthetic and real images, offering a needed analysis of the ADMM-based approach;
- We repeat the experiments using LM, presenting a useful comparison between alternative methods.

Methodology

In this section, we propose a systematic method to obtain $T2^*$ estimations by performing curve fitting on the signal of magnetic resonance images. Obtaining an estimate of $T2^*$ allows us to have a proxy measure for LIC, which in turn is a reliable indicator of total body iron content. However, in the presence of noise and vasculature, this estimation becomes difficult. In order to approach this task, the very first step involves the choice of an appropriate relation between the measured signal and the $T2^*$. This decision corresponds to the definition of the model we want to use to obtain a signal value given $T2^*$ as input parameter. Then, we can define a loss function to compare the reconstructed signal with the actual measurement. Our estimated $T2^*$ will be the one leading to the minimal loss value.

Model

Figure 2 illustrates curves which are akin to the evolution of signal values as TE increases. In particular, the signal appears to show an exponential decay, which becomes faster for lower $T2^*$ values. More intuitively, the decay rate increases as $R2^*$ grows. In order to define this relation with an explicit model, we now present the needed notation. At each of the M echo times used, an MRI scan provides an image

of generic shape K by L . Consequently, the final result is an image tensor of shape K by L by M , with $N = K \times L \times M$ voxels containing M measurement each. We denote this signal tensor by $\mathbf{S} \in \mathbb{R}^{K \times L \times M}$ and we write $\mathbf{s}_n = \{s_1, s_2, \dots, s_M\}$ for the signal at each voxel n . Moreover, r_n will represent the corresponding scalar R2^* value. Finally, for simplicity, we denote by $\mathbf{t} \in \mathbb{R}^M$ the vector of TEs. With these ingredients at hand, an intuitive model definition could be:

$$\mathbf{s}_n = a_n e^{-r_n \mathbf{t}}. \quad (1)$$

Here, a_n is an additional scalar parameter (the "intercept") useful to avoid imposing that $\mathbf{s}_n = 1$ at $\text{TE} = 0$ ms. This intuitive model is the one considered for conducting the experiments described in this paper.

Optimisation problem

Chosen the model, the next step is defining a loss function that, given a model output and the true signal, quantifies the distance between the two. This is a useful proxy for the distance between the estimated T2^* and its ground truth, which is our true objective but is assumed to be unknown. Moreover, while T2^* is our parameter of interest, we underscore that we will also need to estimate the intercept. It is customary and convenient to employ, as a loss function, the sum of squared differences between the true signal and the model's output. In addition, to introduce global information and increase robustness to noise, we include regularisation terms for both parameters. Specifically, we use Total Variation, often considered the regularisation of favour in signal processing for its ability to discard excessive details caused by noise. More explicitly, we consider the following unconstrained minimisation problem:

$$\min_{\mathbf{a}, \mathbf{r}} \frac{1}{2} \sum_{i=1}^N \|\mathbf{s}_i - a_i e^{-r_i \mathbf{t}}\|_2^2 + \lambda_1 \|\mathbf{a}\|_{TV} + \lambda_2 \|\mathbf{r}\|_{TV}. \quad (2)$$

Despite its convexity, this optimisation problem cannot be solved with gradient-based methods. In particular, a complication is represented by the dependence of both the sum of squared errors and the Total Variation terms on the parameters. On the other hand, this problem fits naturally into a class of objectives commonly addressed by the alternating direction method of multipliers (ADMM). The first step prescribed by ADMM is known as variable splitting. This involves introducing new variables (the "splitting parameters", denoted \mathbf{g} and \mathbf{f}) to replace the parameters appearing in the regularisation terms. Then, equality constraints are imposed between these variables and the parameters they replace, so that the obtained problem remains equivalent to the original one. After this process, we can rewrite the objective as follows:

$$\begin{aligned} \min_{\mathbf{a}, \mathbf{r}, \mathbf{g}, \mathbf{f}} & \frac{1}{2} \sum_{i=1}^N \|\mathbf{s}_i - a_i e^{-r_i \mathbf{t}}\|_2^2 + \lambda_1 \|\mathbf{g}\|_{TV} + \lambda_2 \|\mathbf{f}\|_{TV} \\ \text{subject to} & \end{aligned} \quad (3)$$

$$\mathbf{g} = \mathbf{a},$$

$$\mathbf{f} = \mathbf{r}.$$

Subsequently, we can consider the corresponding augmented Lagrangian:

$$\begin{aligned} \mathcal{L}(\mathbf{a}, \mathbf{r}, \mathbf{g}, \mathbf{f}, \mathbf{d}_1, \mathbf{d}_2) = & \frac{1}{2} \sum_{i=1}^N \|\mathbf{s}_i - a_i e^{-r_i \mathbf{t}}\|_2^2 \\ & + \lambda_1 \|\mathbf{g}\|_{TV} + \lambda_2 \|\mathbf{f}\|_{TV} \\ & + \frac{\mu}{2} \|\mathbf{a} - \mathbf{g} - \mathbf{d}_1\|_2^2 + \frac{\mu}{2} \|\mathbf{r} - \mathbf{f} - \mathbf{d}_2\|_2^2, \end{aligned} \quad (4)$$

where \mathbf{d}_1 and \mathbf{d}_2 are the Lagrange multipliers and $\mu > 0$ is a constant. Given one such objective, the ADMM searches for the solution value of each parameter individually and iteratively.

Solving for \mathbf{a}

In order to solve for any a_n , we can simply focus on $\mathcal{L}_a(\mathbf{a}, \mathbf{r}, \mathbf{g}, \mathbf{d}_1)$, the portion of the Lagrangian that only includes terms depending on the parameter of interest:

$$\mathcal{L}_a(\mathbf{a}, \mathbf{r}, \mathbf{g}, \mathbf{d}_1) := \frac{1}{2} \sum_{i=1}^N \|\mathbf{s}_i - a_i e^{-r_i \mathbf{t}}\|_2^2 + \frac{\mu}{2} \|\mathbf{a} - \mathbf{g} - \mathbf{d}_1\|_2^2 \quad (5)$$

The solution can then be found computing the partial derivative with respect to a_n and setting it to zero:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial a_n} &= 0, \\ &- (e^{-r_n \mathbf{t}})^T (\mathbf{s}_n - a_n e^{-r_n \mathbf{t}}) + \mu(a_n - g_n - d_{1n}) = 0, \\ &- (e^{-r_n \mathbf{t}})^T \mathbf{s}_n + a_n (e^{-r_n \mathbf{t}})^T e^{-r_n \mathbf{t}} + \mu a_n = \mu(g_n + d_{1n}), \\ a_n \left[(e^{-r_n \mathbf{t}})^T e^{-r_n \mathbf{t}} + \mu \right] &= (e^{-r_n \mathbf{t}})^T \mathbf{s}_n + \mu(g_n + d_{1n}), \\ a_n &= \frac{(e^{-r_n \mathbf{t}})^T \mathbf{s}_n + \mu(g_n + d_{1n})}{(e^{-r_n \mathbf{t}})^T e^{-r_n \mathbf{t}} + \mu}. \end{aligned} \quad (6)$$

Solving for \mathbf{r}

Proceeding as before, we can define the objective to solve for any r_n as follows:

$$\mathcal{L}_r(\mathbf{a}, \mathbf{r}, \mathbf{f}, \mathbf{d}_2) := \frac{1}{2} \sum_{i=1}^N \|\mathbf{s}_i - a_i e^{-r_i \mathbf{t}}\|_2^2 + \frac{\mu}{2} \|\mathbf{r} - \mathbf{f} - \mathbf{d}_2\|_2^2 \quad (7)$$

In this case, we need to approach the minimisation of the expression using a number K (arbitrarily chosen) of gradient descent steps. Therefore, we need to compute the partial derivative of \mathcal{L}_r with respect to r_n :

$$\frac{\partial \mathcal{L}_r}{\partial r_n} = (a_n \mathbf{t} \odot e^{-r_n \mathbf{t}})^T (s_n - a_n e^{-r_n \mathbf{t}}) + \mu(r_n - f_n - d_{2n}), \quad (8)$$

where \odot stands for the component-wise (Hadamard) product. Having found the partial derivative, at each step k of gradient descent we implement the update rule:

$$r_n^{(k+1)} = r_n^{(k)} - \eta \frac{\partial \mathcal{L}_r}{\partial r_n}, \quad (9)$$

where η is known as learning rate and controls the size of the updates.

Solving for \mathbf{g}

On the same lines of the previous steps, we here consider the terms of the Lagrangian that only depend on \mathbf{g} :

$$\mathcal{L}_g(\mathbf{a}, \mathbf{g}, \mathbf{d}_1) := \lambda_1 \|\mathbf{g}\|_{\text{TV}} + \frac{\mu}{2} \|\mathbf{a} - \mathbf{g} - \mathbf{d}_1\|_2^2. \quad (10)$$

The solution to this problem is calculated using the Chambolle algorithm [7]:

$$\mathbf{g} = \text{Chambolle} \left(\mathbf{a} - \mathbf{d}_1, \frac{\lambda_1}{\mu} \right). \quad (11)$$

Solving for \mathbf{f}

Finally, we can concentrate on \mathbf{f} and define:

$$\mathcal{L}_f(\mathbf{r}, \mathbf{f}, \mathbf{d}_2) := \lambda_2 \|\mathbf{f}\|_{\text{TV}} + \frac{\mu}{2} \|\mathbf{r} - \mathbf{f} - \mathbf{d}_2\|_2^2, \quad (12)$$

an expression that can be treated using again the Chambolle algorithm:

$$\mathbf{f} = \text{Chambolle} \left(\mathbf{r} - \mathbf{d}_2, \frac{\lambda_2}{\mu} \right). \quad (13)$$

The following is a summary of the algorithmic procedure:

Algorithm 1 Iron overload estimation using ADMM.

```

1: set  $k = 0$ , choose  $\mathbf{a}^{(0)}, \mathbf{r}^{(0)}, \mathbf{g}^{(0)}, \mathbf{f}^{(0)}, \mathbf{d}_1^{(0)}, \mathbf{d}_2^{(0)}$ , and  $\mu > 0$ 
2: repeat
3:    $k \leftarrow k + 1$ 
4:    $\mathbf{a}^{(k+1)} \leftarrow \arg \min_{\mathbf{a}} \mathcal{L}(\mathbf{a}, \mathbf{r}^{(k)}, \mathbf{g}^{(k)}, \mathbf{f}^{(k)}, \mathbf{d}_1^{(k)}, \mathbf{d}_2^{(k)})$ 
5:    $\mathbf{r}^{(k+1)} \leftarrow \arg \min_{\mathbf{r}} \mathcal{L}(\mathbf{a}^{(k+1)}, \mathbf{r}, \mathbf{g}^{(k)}, \mathbf{f}^{(k)}, \mathbf{d}_1^{(k)}, \mathbf{d}_2^{(k)})$ 
6:    $\mathbf{g}^{(k+1)} \leftarrow \arg \min_{\mathbf{g}} \mathcal{L}(\mathbf{a}^{(k+1)}, \mathbf{r}^{(k+1)}, \mathbf{g}, \mathbf{f}^{(k)}, \mathbf{d}_1^{(k)}, \mathbf{d}_2^{(k)})$ 
7:    $\mathbf{f}^{(k+1)} \leftarrow \arg \min_{\mathbf{f}} \mathcal{L}(\mathbf{a}^{(k+1)}, \mathbf{r}^{(k+1)}, \mathbf{g}^{(k+1)}, \mathbf{f}, \mathbf{d}_1^{(k)}, \mathbf{d}_2^{(k)})$ 
8:    $\mathbf{d}_1^{(k+1)} \leftarrow \mathbf{d}_1^{(k)} - (\mathbf{g}^{(k+1)} - \mathbf{a}^{(k+1)})$ 
9:    $\mathbf{d}_2^{(k+1)} \leftarrow \mathbf{d}_2^{(k)} - (\mathbf{f}^{(k+1)} - \mathbf{r}^{(k+1)})$ 
10: until some stopping criterion is satisfied

```

Now that we have defined a model and a precise method to obtain parameter estimates, it remains to assess the quality of our approach. This can be done by first running experiments on synthetic data with known ground truth and then testing on real patient images.

Experiments

Phantom images

Data

When running experiments, the primary aim is to understand whether the recovered parameter estimates are reasonable. Ideally, we would like to compare our results with a ground truth. This motivates our decision to begin our experiments by generating phantom images, since the parameter values we arbitrarily select to produce such images can be used as ground truth. Precisely, we decide to create 4 image tensors of size $32 \times 32 \times 12$. Each of the 12 channels corresponds to one entry of the TE vector $\mathbf{t} = [2.58, 4.81, 7.59, 9.89, 12.12, 14.35, 16.58, 18.88]^T$. Moreover, we impose that for each image the parameter values are constant across voxels. Consequently, in each image the signal varies only for different time to echos, hence across different channels. With these settings, we need to choose only four values for \mathbf{a} ($[155, 255, 355, 455]$) and four for T2^* ($[5, 10, 15, 20]$). Then, given one of the four parameter pairs as input, the model will output a signal vector of the same length as \mathbf{t} , hence with one value per channel. As

signal values are constant for pixels in the same channel, this vector is enough to reconstruct the entire image (Figure 3).

Table 1. True values and mean estimates of the parameters as obtained by ADMM on phantom images. For every parameter, the colour intensity of a cell is proportional to the amount of the total absolute error that was caused by the corresponding estimate.

	Image 1	Image 2	Image 3	Image 4
Parameter \mathbf{a}				
True value	155	255	355	455
Estimated value	154.8006	254.9949	355.0071	455.0027
Parameter T2^*				
True value	5	10	15	20
Estimated value	5.0113	10.0004	14.9993	19.9997

Estimations

Obtained the image tensors, we can run our algorithm as explained before in order to recover the parameter estimates. Then, comparison with ground truth values can be performed visually as depicted in Figures 4 and 6. Moreover, it is possible to quantitatively compare the estimates with the ground truth. Throughout the experiments, the distance of the estimates from the values of reference will be measured in terms of absolute error. Table 1 confirms a high similarity between ground truth values and mean estimates for both parameters. Subsequently, we repeated this procedure adding Gaussian noise (mean 0, variance 10) to observe how it can affect the results. The estimates appear indeed less precise, but taking the average seems to be enough to mitigate the effect of noise and recover values that are close to the ground truths (Figures 5, 7, and in the Appendix Figures 18, 19).

As mentioned before, in our optimisation we used Total Variation terms to neglect unwanted details of the images. This effect can be verified by comparing the figures we obtain if we increase λ_1 and λ_2 , the coefficients weighting the regularisation terms (Figure 20 in the Appendix)

These analyses were useful to assess the suitability of our method with a simple strategy, before progressing on real images.

Real images

Data

Real images do not equip us with ground truth parameter values. On the other hand, patient images provide an important opportunity to conduct more realistic experiments. In-vivo abdominal MRI scans were performed on four human subjects with sickle cell disease ¹ (mean age: 38 ± 12 years) using a 3T Siemens Skyra scanner (Siemens Healthcare, Erlangen, Germany). Imaging was conducted with a body surface coil and an 8-echo gradient echo (GRE) sequence, with echo times evenly spaced between 2.58 ms and 18.88 ms. Imaging parameters included: repetition time = 5200 ms, matrix size = 192×256 , bandwidth = 1776 Hz/pixel, flip angle = 20° , slice thickness = 10 mm, and field of view (FOV) adjusted to patient size. A single axial slice through the

¹ While the biological background presented focuses on thalassemia due to its clinical relevance in transfusion-induced iron overload, the dataset used in this work consists of T2*-weighted MRI scans from patients with sickle cell disease, who are also subject to similar iron-related complications. The imaging methodology remains applicable to both conditions.

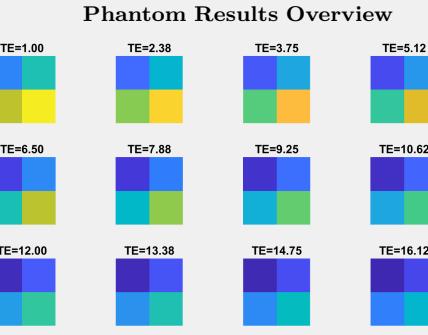


Fig. 3: Each panel shows the signal at one TE across four phantom images. For example, the top-left square of the first panel displays the first channel of the first image. As the time to echo grows, panels become darker due to signal decay (decay curves are reported in Figure 17, Appendix). The bottom-right square in each panel corresponds to the image with the highest $T2^*$, becoming darker at the slowest rate.

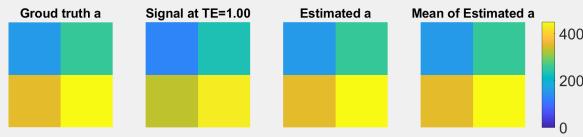


Fig. 4: Results for **a** (no noise). From left to right: ground truth, signal at first TE, estimated **a** (closely resembling the ground truth), and mean of estimated **a**.

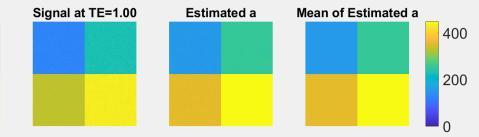


Fig. 5: Results for **a** with noise. Averaging reduces noise impact, yielding estimates close to ground truth. This effect is even more evident for more intense noise (Figure 18).

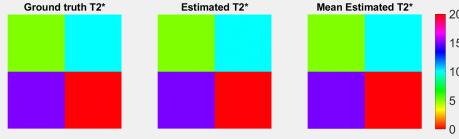


Fig. 6: Results for $T2^*$ (no noise). Similar layout to **a**. Also in this case, we can observe that the estimates are close to the ground truth.

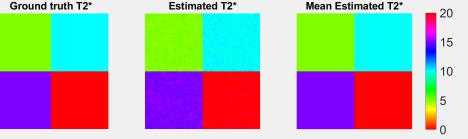


Fig. 7: $T2^*$ estimates with added noise. Averaging mitigates the noise effect. This effect is even more evident for more intense noise (Figure 19).

mid-liver was acquired during end-expiration breath hold.

Figure 8 displays all the $T2^*$ -weighted images obtained from the study participants, highlighting the signal decay in the liver with increasing TE. Due to variability in hepatic iron distribution, susceptibility artifacts, and vascular inclusion, $T2^*$ values can vary across different liver regions. As part of this study, liver biopsies were not obtained, as they were not part of standard clinical care. $T2^*$ values were instead computed using the scanner's built-in $T2^*$ mapping software, which employs a single exponential fitting model. Circular regions of interest (ROIs) approximately 4 cm in diameter were placed in the center of the vials' cross-sections and within the right hepatic lobe, avoiding vascular structures.

Estimations

Running the algorithm on these images yields the parameter estimates visualised in Figures 9 and 10. Interestingly, the $T2^*$ values for the second patient are particularly low, resulting in a dark image and suggesting high iron concentration. Combining these observations with the intuitions we gained from Figure 2, we expect that the signal decay over time to echos is especially fast for this patient. To quantitatively verify this, we focus on an ROI in the lower part of the livers' right lobe, and we compute the mean signal across TEs. Figure 12 confirms our

expectations, highlighting for the second patient a decay curve that intersects the one for the third patient despite starting at a higher value. Clearly, the results are influenced by the particular ROI chosen. With the aim of illustrating this, we compute $T2^*$ values for three more ROIs (Figure 24) per patient, and compare all the estimates with the results obtained by the scanner's software. As expected, Table 2 describes significant fluctuations in the results between different ROIs. For example, the "left" ROI often seems to cause the highest error.

Table 2. Comparison of mean estimated $T2^*$ and scanner results at different ROIs. The intensity of the red colour of a cell is proportional to the amount of the total absolute error caused by the estimate in that cell.

	Patient 1	Patient 2	Patient 3	Patient 4
Scanner result	9.1323	4.2624	16.2949	10.7012
Right ROI	10.4905	3.8968	16.3604	10.9868
Vasc ROI	10.8184	5.2993	16.9020	10.9755
Edge ROI	8.1741	4.9750	17.1849	9.9705
Left ROI	11.0596	4.6903	14.1894	18.1106

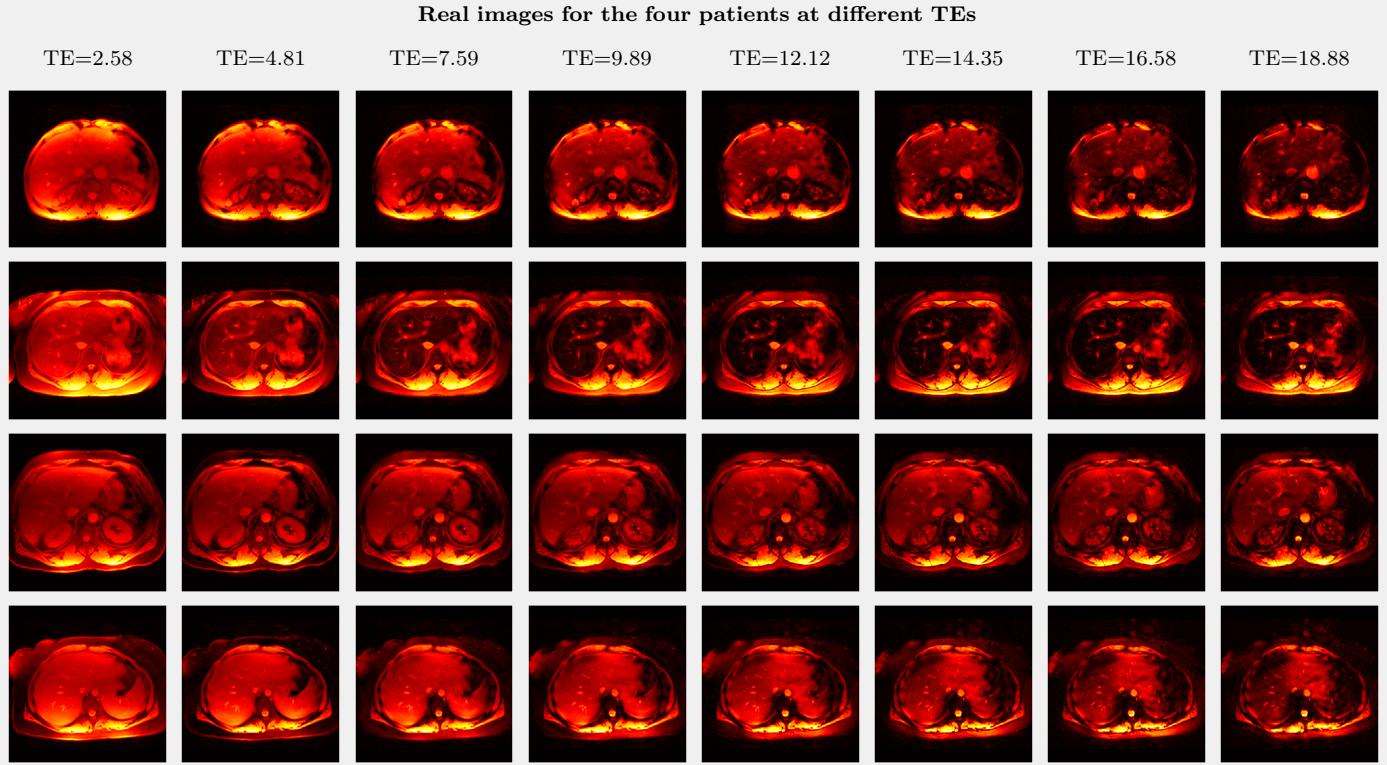


Fig. 8: All the real images considered. Each row corresponds to a patient. Each column corresponds to a different TE. The second patient (second row) seems to be the one for which the signal decays faster. This may suggest that this patient has a severe iron overload.

Maps of estimates of $T2^*$ for the four patients

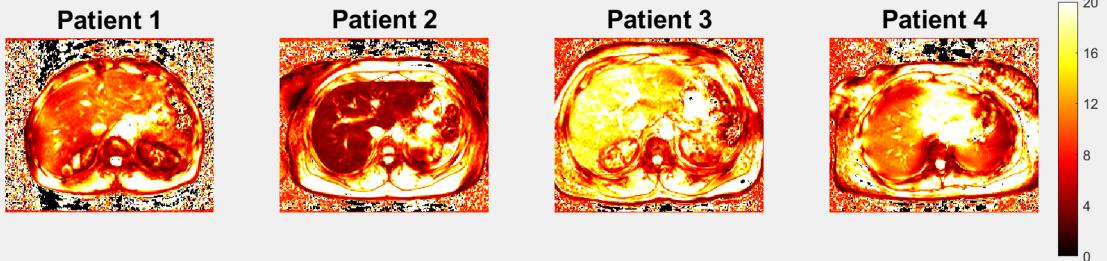


Fig. 9: Parameter $T2^*$ for the four patients. In particular, for the second patient the estimated $T2^*$ assumes the lowest values, indicating the highest iron concentration. The main hyperparameters were used with default values $\lambda_1 = 1e-5$, $\lambda_2 = 1e-5$, $\eta = 1e-8$. Different constants for regularisation were tested in Figure 21 and Table 4.

Maps of estimates of a for the four patients

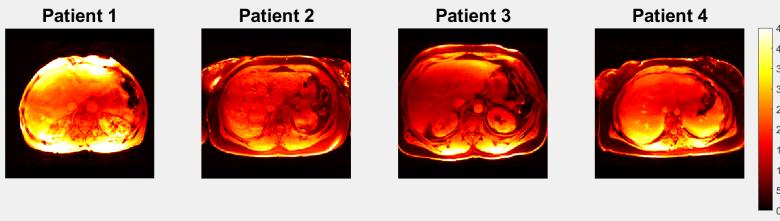


Fig. 10: Parameter a for the four patients. The highest intercept appears to be the one for patient 1.

Patient 1: example of ROI

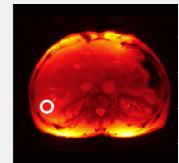


Fig. 11: Illustration of the "right" ROI for patient 1.

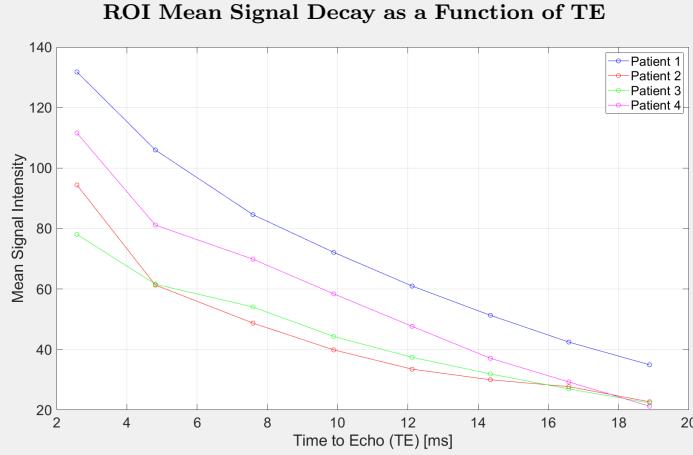


Fig. 12: The figure illustrates the mean signal decay curves for the four patients. The red line refers to the second patient and shows a fast decay. This is particularly evident in comparison with the third patient (in green). Indeed, the red line starts higher but intersects the green line.

Using Levenberg Marquardt

Finally, for comparison purposes, we repeated the experiments using the Levenberg–Marquardt method instead of ADMM. Similarly to ADMM, Levenberg–Marquardt is designed for nonlinear least squares problems, but in its basic form it does not include regularisation terms. Therefore, there is no explicit noise handling or encoding of global information. Mathematically, we use it to minimise the following simplified objective:

$$\mathcal{L}_{LM}(\mathbf{a}, \mathbf{r}) := \frac{1}{2} \sum_{i=1}^N \|\mathbf{s}_i - a_i e^{-r_i t}\|_2^2 \quad (14)$$

More precisely, LM solves this problem by interpolating between the Gauss–Newton and gradient descent methods. In our case, for every voxel n , it defines the parameter update rules at iteration step k as:

$$a_n^{(k+1)} = a_n^{(k)} - \left(\frac{\partial^2 \mathcal{L}_{LM}}{\partial^2 a_n} + \eta \right)^{-1} \frac{\partial \mathcal{L}_{LM}}{\partial a_n}; \quad (15)$$

$$r_n^{(k+1)} = r_n^{(k)} - \left(\frac{\partial^2 \mathcal{L}_{LM}}{\partial^2 r_n} + \eta \right)^{-1} \frac{\partial \mathcal{L}_{LM}}{\partial r_n}. \quad (16)$$

Here, $\eta > 0$ is the hyperparameter that controls the balance between gradient descent and Gauss–Newton behavior. When η is large the update resembles gradient descent, while a small η leads to a Gauss–Newton behaviour (Table 5). Some of the results obtained with this method find interpretations similar to those already described, and are therefore reported in the Appendix.

The experimental results on phantom images are illustrated in Figures 22 and 23, which show similarities with what we observed for ADMM. Nonetheless, we can observe that in the presence of high noise, the absence of regularisation terms in LM leads to lower robustness and flexibility compared to ADMM (Figure 13). Subsequently, we performed the experiments on real images, obtaining the parameter maps for the four patients (Figures 14 and 15). The insights achieved with ADMM, especially regarding the second patient, appear to be confirmed by the LM method. Table 3 displays the best mean estimates achieved across the two models for the different ROIs, with blue cells indicating that the corresponding best estimate was obtained with ADMM. For completeness, we include all the mean estimates for LM in Table 6. We observe how the ADMM yields the best estimate almost twice as frequently as LM. ADMM likely owes its improved

performance to the inclusion of Total Variation terms. Indeed, these regularisation terms help suppress noise while also capturing the spatial correlation between neighbouring voxels. On the other hand, the total computational time needed for obtaining the estimates for the four patients appears reduced by LM, indicating that this algorithm may be a good option when efficiency is a priority (Figure 16).

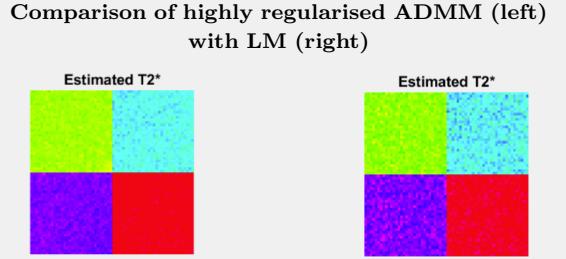


Fig. 13: Estimates of $T2^*$ in the presence of severe noise ($\sigma^2 = 100$). ADMM used λ_1 and λ_2 equal to 100, offering greater robustness to noise compared to LM.

Table 3. Best mean $T2^*$ estimates between ADMM and LM, compared with the scanner result. Every cell reports the best estimate. The cell is blue if this estimate came from ADMM, yellow if it came from LM, and white if both models returned the same estimate or if the cell refers to the scanner's results.

	Patient 1	Patient 2	Patient 3	Patient 4
Scanner result	9.1323	4.2624	16.2949	10.7012
Right ROI	10.4905	3.8968	16.3604	10.9868
Vasc ROI	10.8184	5.2674	16.9020	10.9755
Edge ROI	8.1741	4.9463	17.1849	9.9708
Left ROI	11.0596	4.6689	14.1895	18.1106

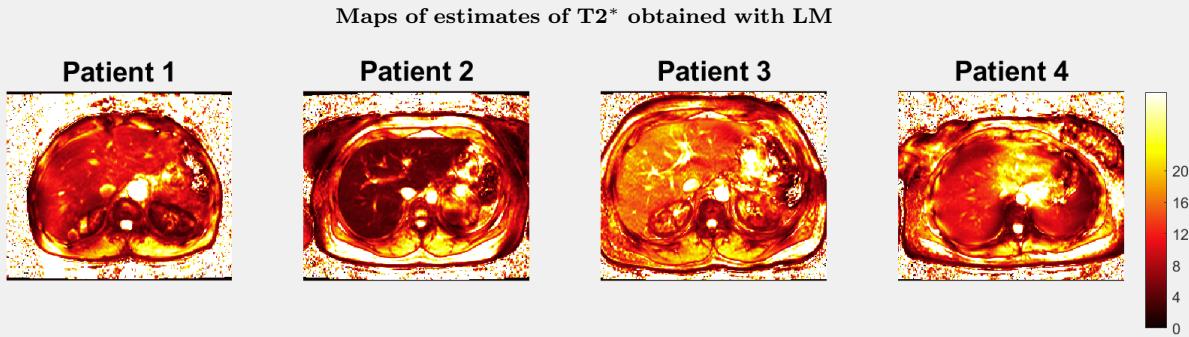


Fig. 14: Parameter $T2^*$ for the four patients obtained using LM. Confirming previous observations, patient 2 shows the lowest $T2^*$ values. The maps appear visually similar to those obtained earlier. As also described by Table 3, LM generally shows a worse performance than ADMM. Here we used starting point [100,0.2], Figure 25 shows the results obtained with a different initial point.

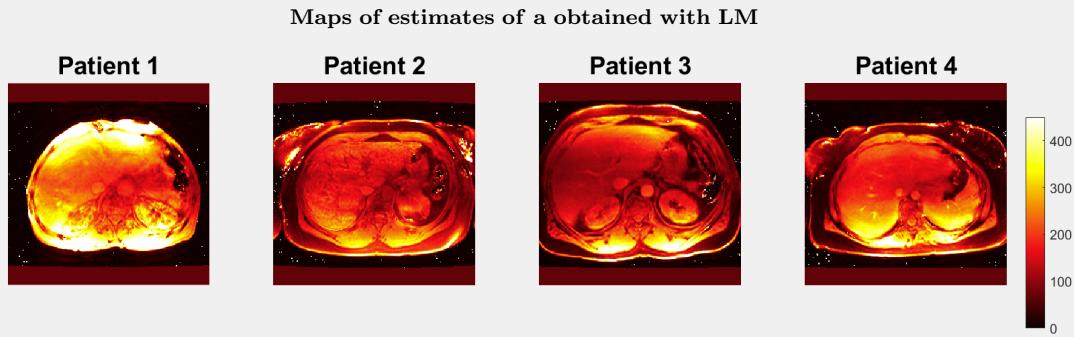


Fig. 15: Parameter α for the four patients obtained with LM. Also in this case, patient 1 displays the highest intercept.

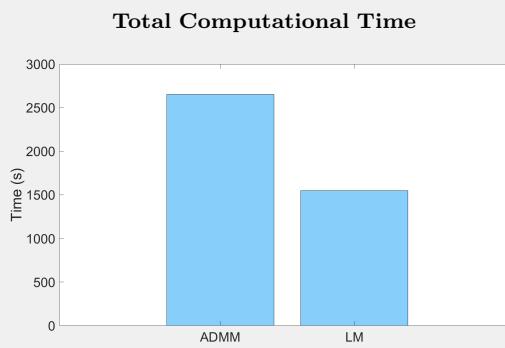


Fig. 16: Total computational time (in seconds) for obtaining all the parameter estimates. The simpler approach proposed by LM offers an increase in efficiency.

Conclusions and future work

Thalassemias represent a major challenge for public health and well-being. The treatment with regular transfusions causes iron overload and therefore needs to be complemented by chelation therapy. Quantification of the amount of surplus iron is crucial for a correct initiation and monitoring of this treatment. The current gold standard for this quantification is liver biopsy, which however presents several disadvantages, including its invasiveness. In the study presented here, no biopsy sample was considered. Therefore, we were not

equipped with any true ground truth of LIC for the patient images. Consequently, after carefully explaining the steps of ADMM, we began our experiments on phantom data, that could be created arbitrarily choosing ground truth values. The estimates obtained in these experiments were promising for both ADMM and LM. The main difference between the two methods emerged in the presence of high noise, where only ADMM appeared to maintain its performance, due to its inclusion of regularisation terms. Subsequently, we considered the patient images, obtaining parameter maps of α and $T2^*$. For both ADMM and LM, the resulting $T2^*$ values of the second patient appeared particularly low, providing strong evidence for a severe condition of iron overload. The final experiment focused on ROIs to compute the mean parameter estimate of the corresponding pixels. The results returned by the scanner were used as a proxy for a ground truth. The experiments revealed how the heterogeneous iron deposit pattern in the liver can drastically influence the results and their accuracy. Indeed, the estimates were subject to significant fluctuations across ROIs, with the "left" ROI leading to the worst predictions. In terms of absolute error, ADMM demonstrated superior performance compared to LM, and its advantages were probably again related to the inclusion of regularisation terms. On the other hand, LM has proven its utility in increasing estimation efficiency.

This study presents some limitations. First of all, the reduced sample size does not provide us with solid guarantees on the generalisation of our results. Furthermore, while we only employed a single exponential model, the application of other alternatives may also yield interesting results. In addition, the lack of a true ground truth mines the validity of our interpretations on the quality of the estimates. Moreover, as noted

during the experiments, the arbitrary selection of ROIs is a delicate passage of the pipeline for estimating iron overload. The subjectivity introduced by this step may compromise the generalisability of the conclusions.

Therefore, further studies could begin by considering a larger sample size, or a richer set of models. Then, developing a method for automatically optimising the choice of the ROIs would represent an important milestone in the field. This could reduce the dependence of results on the human operator, and could be particularly beneficial in the diagnosis of borderline patients. Similarly, methods aimed at correcting the estimates for the presence of fat hold the promise to enhance the reliability of results. Moreover, the studies presented in this paper may be complemented by retrieving a calibration curve, offering a direct relation between $T2^*$ and LIC. Furthermore, the hyperparameters of ADMM (such as the learning rate for r and the constants multiplying the Total Variation terms) could be tuned for additional performance improvements (Figure 21 and Table 4). While Total Variation is a standard choice of regularisation in this field, similar alternatives such as Laplacian regularisation are probably worth exploring, in order to observe the effects on noise handling. In addition, machine learning has delivered impressive breakthroughs in the field of medical MRI, and investigating approaches on the same lines as Positano et al. [36] may offer insightful observations. For example, Positano et al. describe the need of employing data augmentation for their models. However, this represents an expensive operation. Designing an architecture able to generalise over scanner parameters without the need of data augmentation could lead to improvements in both efficiency and accuracy. In addition, because of the complexity of CNNs, works focusing on enhancing their interpretability may generate fruitful insights. Finally, we notice that the typical noise distribution for these tasks is Rician, and this can be taken into account by introducing ad-hoc initial denoising layers in the CNNs, similar to what Sedlar et al. have already proposed for analyses of brain dMRI [41].

References

1. Sherif RZ Abdel-Misih and Mark Bloomston. Liver anatomy. *Surgical Clinics*, 90(4):643–653, 2010.
2. LJ Anderson, S Holden, B Davis, E Prescott, CC Charrier, NH Bunce, DN Firmin, B Wonke, J Porter, JM Walker, et al. Cardiovascular T2-star ($T2^*$) magnetic resonance for the early diagnosis of myocardial iron overload. *European heart journal*, 22(23):2171–2179, 2001.
3. Emanuele Angelucci, Gary M Brittenham, Christine E McLaren, Marta Ripalti, Donatella Baronciani, Claudio Giardini, Maria Galimberti, Paola Polchi, and Guido Lucarelli. Hepatic iron concentration and total body iron stores in thalassemia major. *New England Journal of Medicine*, 343(5):327–331, 2000.
4. Bruce R Bacon and Anthony S Tavill. Role of the liver in normal iron metabolism. In *Seminars in liver disease*, volume 4, pages 181–192. © 1984 by Thieme Medical Publishers, Inc., 1984.
5. Wajihah Bano, Gian Franco Piredda, Mike Davies, Ian Marshall, Mohammad Golbabaei, Reto Meuli, Tobias Kober, Jean-Philippe Thiran, and Tom Hilbert. Model-based super-resolution reconstruction of T2 maps. *Magnetic resonance in medicine*, 83(3):906–919, 2020.
6. Kenneth P Batts. Iron overload syndromes and the liver. *Modern Pathology*, 20(1):S31–S39, 2007.
7. Antonin Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1–2):89–97, 2004.
8. Athanasios Christoforidis, Vassilios Perifanis, George Spanos, Efthimia Vlachaki, Marina Economou, Ioanna Tsatra, and Miranda Athanassiou-Metaxa. MRI assessment of liver iron content in thalassamic patients with three different protocols: comparisons and correlations. *European journal of haematology*, 82(5):388–392, 2009.
9. Yves Deugnier and Bruno Turlin. Pathology of hepatic iron overload. *World journal of gastroenterology: WJG*, 13(35):4755, 2007.
10. Som Dev and Jodie L Babitt. Overview of iron metabolism in health and disease. *Hemodialysis International*, 21:S6–S20, 2017.
11. Ahmed Karam Eldaly and Ayman M Khalifa. A general framework for hepatic iron overload quantification using MRI. *Digital Signal Processing*, 137:104048, 2023.
12. Qing Feng, Jixing Yi, Tao Li, Bumin Liang, Fengming Xu, and Peng Peng. Narrative review of magnetic resonance imaging in quantifying liver iron load. *Frontiers in Medicine*, 11:1321513, 2024.
13. Maciej W Garbowski, John-Paul Carpenter, Gillian Smith, Michael Roughton, Mohammed H Alam, Taigang He, Dudley J Pennell, and John B Porter. Biopsy-based calibration of $T2^*$ magnetic resonance for estimation of liver iron concentration and comparison with R2 ferriscan. *Journal of cardiovascular magnetic resonance*, 16(1):40, 2014.
14. Mohammad Golbabaei, Guido Buonincontri, Carolin M Pirkl, Bjoern H Menzel, Marion I and Menze, Mike Davies, and Pedro A Gómez. Compressive MRI quantification using convex spatiotemporal priors and deep encoder-decoder networks. *Medical image analysis*, 69:101945, 2021.
15. Jane S Hankins, M Beth McCarville, Ralf B Loeffler, Matthew P Smeltzer, Mihaela Onciu, Fredric A Hoffer, Chin-Shang Li, Winfred C Wang, Russell E Ware, and Claudia M Hillenbrand. R2* magnetic resonance imaging of the liver in patients with iron overload. *Blood, The Journal of the American Society of Hematology*, 113(20):4853–4855, 2009.
16. Taigang He, Peter D Gatehouse, Gillian C Smith, Raad H Mohiaddin, Dudley J Pennell, and David N Firmin. Myocardial t measurements in iron-overloaded thalassemia: An in vivo study to investigate optimal methods of quantification. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 60(5):1082–1089, 2008.
17. B Henninger, H Zoller, S Rauch, A Finkenstedt, M Schocke, W Jaschke, and C Kremser. R2* relaxometry for the quantification of hepatic iron overload: biopsy-based calibration and comparison with the literature. In *RöFo-Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*, volume 187, pages 472–479. © Georg Thieme Verlag KG, 2015.
18. Diego Hernando, Ruiyang Zhao, Qing Yuan, Mounees Aliyari Ghasabeh, Stefan Ruschke, Xinran Miao, Dimitrios C Karampinos, Lu Mao, David T Harris, Ryan J Mattison, et al. Multicenter reproducibility of liver iron quantification with 1.5-t and 3.0-t MRI. *Radiology*, 306(2):e213256, 2022.
19. C. Hershko. Iron overload in transfusion-dependent thalassemia. *International Journal of Hematology*, 101(6):758–759, 2015. Accessed: 07-04-2025.
20. Shuai Huang, James J. Lah, Jason W. Allen, and Deqiang Qiu. Fast nonconvex t_2^* mapping usingADMM, 2020.
21. Shuai Huang, James J Lah, Jason W Allen, and Deqiang Qiu. Accelerated model-based T1, $T2^*$ and proton density mapping using a bayesian approach with automatic hyperparameter estimation. *Magnetic Resonance in Medicine*, 93(2):563–583, 2025.

22. El-Sayed H Ibrahim, Ayman M Khalifa, and Ahmed K Eldaly. MRI T2* imaging for assessment of liver iron overload: study of different data analysis approaches. *Acta Radiologica*, 57(12):1453–1459, 2016.
23. Roxanne Labranche, Guillaume Gilbert, Milena Cerny, Kim-Nhien Vu, Denis Soulières, Damien Olivié, Jean-Sébastien Billiard, Takeshi Yokoo, and An Tang. Liver iron quantification with mr imaging: A primer for radiologists. *RadioGraphics*, 38(2):392–412, 2018. PMID: 29528818.
24. Kenneth Levenberg. A method for the solution of certain nonlinear problems in least squares. *Quarterly of Applied Mathematics*, 2(2):164–168, 1944.
25. Thomas G Maris, Olympia Papakonstantinou, Vasiliki Chatzimanoli, Alekos Papadakis, Konstantinos Pagonidis, Nickolas Papanikolaou, Apostolos Karantanas, and Nicholas Gourtsoyiannis. Myocardial and liver iron status using a fast t quantitative MRI (t qMRI) technique. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 57(4):742–753, 2007.
26. Donald W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.
27. Delphine Meynard, Jodie L Babitt, and Herbert Y Lin. The liver: conductor of systemic iron balance. *Blood, The Journal of the American Society of Hematology*, 123(2):168–176, 2014.
28. Herbert L Muncie Jr and James S Campbell. Alpha and beta thalassemia. *American family physician*, 80(4):339–344, 2009.
29. Novartis. Thalassemia & iron overload, 2017. Accessed: 09-04-2025.
30. World Health Organization. Anaemia, 2024. Accessed: 08-04-2025.
31. Anita Paisant, Gaspard d'Assignies, Elise Bannier, Edouard Bardou-Jacquet, and Yves Gandon. MRI for the measurement of liver iron content, and for the diagnosis and follow-up of iron overload disorders. *La Presse Médicale*, 46(12):e279–e287, 2017.
32. Sisir K. Patra. Beta thalassemia - symptoms, causes, treatment. Accessed: 07-04-2025.
33. Alessia Pepe, Massimo Lombardi, Vincenzo Positano, Eliana Cracolici, Marcello Capra, Roberto Malizia, Luciano Prossomariti, Daniele De Marchi, Massimo Midiri, and Aurelio Maggio. Evaluation of the efficacy of oral deferiprone in β -thalassemia major by multislice multiecho T2. *European journal of haematology*, 76(3):183–192, 2006.
34. Antonello Pietrangelo. Pathogens, metabolic adaptation, and human diseases—an iron-thrifty genetic model. *Gastroenterology*, 149(4):834–838, 2015.
35. Kanokwan Pinyopornpanish, Adisak Tantiwarawit, Apinya Leerapun, Atiwat Soontornpun, and Satawat Thongsawat. Secondary iron overload and the liver: a comprehensive review. *Journal of Clinical and Translational Hepatology*, 11(4):932, 2023.
36. Vincenzo Positano, Antonella Meloni, Maria Filomena Santarelli, Laura Pistoia, Anna Spasiano, Liana Cuccia, Tommaso Casini, Maria Rita Gamberini, Massimo Allò, Pier Paolo Bitti, et al. Deep learning staging of liver iron content from multiecho MR images. *Journal of Magnetic Resonance Imaging*, 57(2):472–484, 2023.
37. Vincenzo Positano, Benedetta Salani, Alessia Pepe, Maria Filomena Santarelli, Daniele De Marchi, Anna Ramazzotti, Brunella Favilli, Eliana Cracolici, Massimo Midiri, Paolo Cianciulli, et al. Improved T2* assessment in liver iron overload by magnetic resonance imaging. *Magnetic resonance imaging*, 27(2):188–197, 2009.
38. Eliezer A Rachmilewitz and Patricia J Giardina. How I treat thalassemia. *Blood, The Journal of the American Society of Hematology*, 118(13):3479–3488, 2011.
39. Scott B Reeder, Takeshi Yokoo, Manuela França, Diego Hernando, Ángel Alberich-Bayarri, José María Alústiza, Yves Gandon, Benjamin Henninger, Claudia Hillenbrand, Kartik Jhaveri, et al. Quantification of liver iron overload with MRI: review and guidelines from theESGAR andSAR. *Radiology*, 307(1):e221856, 2023.
40. Sandwell West Birmingham Hospitals NHS Trust. Chelation therapy for iron overload, 2014. Accessed: 09-04-2025.
41. Sara Sedlar, Abib Alimi, Théodore Papadopoulos, Rachid Deriche, and Samuel Deslauriers-Gauthier. A spherical convolutional neural network for white matter structure imaging via dmri. In *Medical Image Computing and Computer Assisted Intervention-MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part III 24*, pages 529–539. Springer, 2021.
42. Leonard B Seeff, Gregory T Everson, Timothy R Morgan, Teresa M Curto, William M Lee, Marc G Ghany, Mitchell L Shiffman, Robert J Fontana, Adrian M Di Bisceglie, Herbert L Bonkovsky, et al. Complication rate of percutaneous liver biopsies among persons with advanced chronic liver disease in the holt-c trial. *Clinical gastroenterology and hepatology*, 8(10):877–883, 2010.
43. Ali T Taher and Antoine N Saliba. Iron overload in thalassemia: different organs at different rates. *Hematology 2014, the American Society of Hematology Education Program Book*, 2017(1):265–271, 2017.
44. Leida Tandara and Ilza Salamunić. Iron metabolism: current facts and future directions. *Biochimia medica*, 22(3):311–328, 2012.
45. Caroline Thomas and Andrew B Lumb. Physiology of haemoglobin. *Continuing Education in Anaesthesia, Critical Care & Pain*, 12(5):251–256, 2012.
46. Anne-Cathrine S Vogt, Tasneem Arsiwala, Mona Mohsen, Monique Vogel, Vania Manolova, and Martin F Bachmann. On iron metabolism and its regulation. *International journal of molecular sciences*, 22(9):4591, 2021.
47. Mark Westwood, Lisa J Anderson, David N Firmin, Peter D Gatehouse, Clare C Charrier, Beatrix Wonke, and Dudley J Pennell. A single breath-hold multiecho T2* cardiovascular magnetic resonance technique for diagnosis of myocardial iron overload. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 18(1):33–39, 2003.
48. John C Wood, Cathleen Enriquez, Nilesh Ghugre, J Michael Tyzka, Susan Carson, Marvin D Nelson, and Thomas D Coates. MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients. *Blood*, 106(4):1460–1465, 2005.
49. John C Wood, Pinggao Zhang, Hugh Rienhoff, Walid Abi-Saab, and Ellis Neufeld. R2 and R2* are equally effective in evaluating chronic response to iron chelation. *American journal of hematology*, 89(5):505–508, 2014.

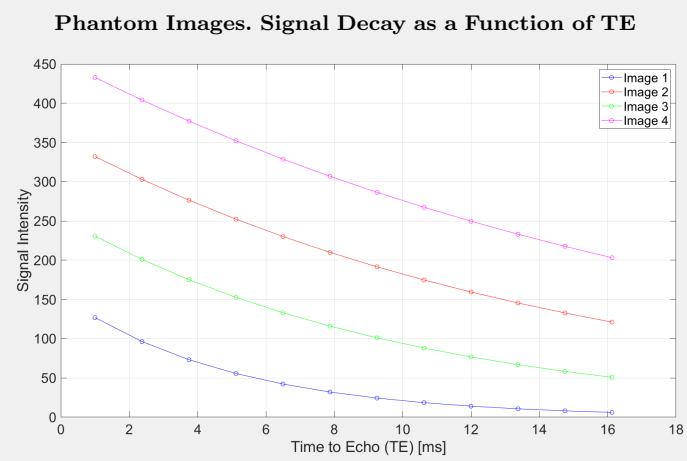
Appendix**Additional Figures and Tables**

Fig. 17: Decay of the signal in phantom images as a function of TE.

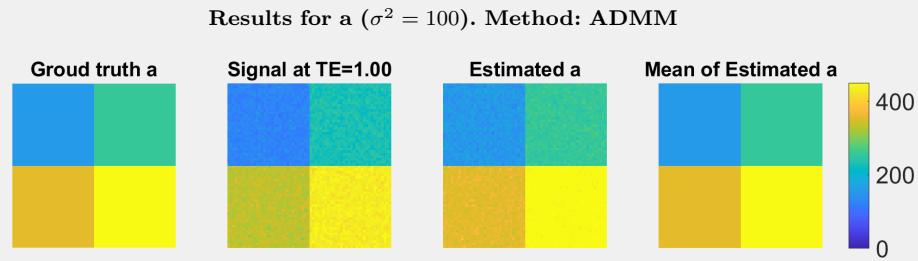


Fig. 18: The interpretation of the figures is similar to before (Figure 5), but we have added more noise ($\sigma^2=100$) to better highlight the results.

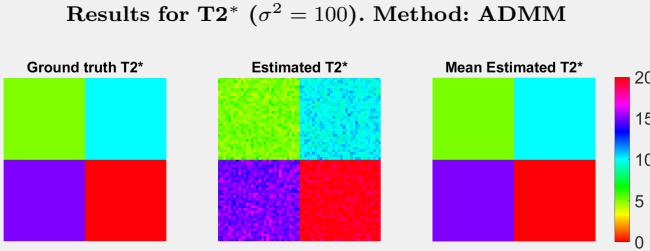


Fig. 19: The interpretation of the figures is similar to before (Figure 7), but we have added more noise ($\sigma^2=100$) to better highlight the results.

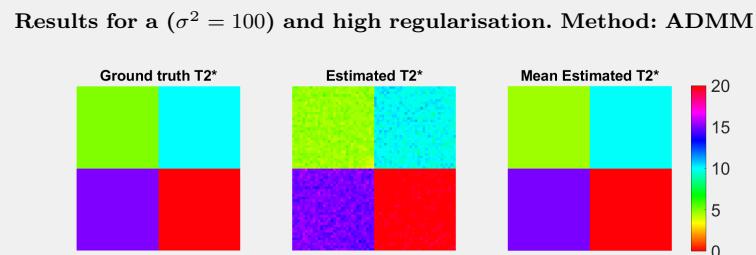


Fig. 20: Results for the estimation of $T2^*$ using ADMM on phantom images with severe noise ($\sigma^2 = 100$) and high regularisation (λ_1 and λ_2 equal to 100).

Estimates with λ_1 and λ_2 equal to 10. Method: ADMM

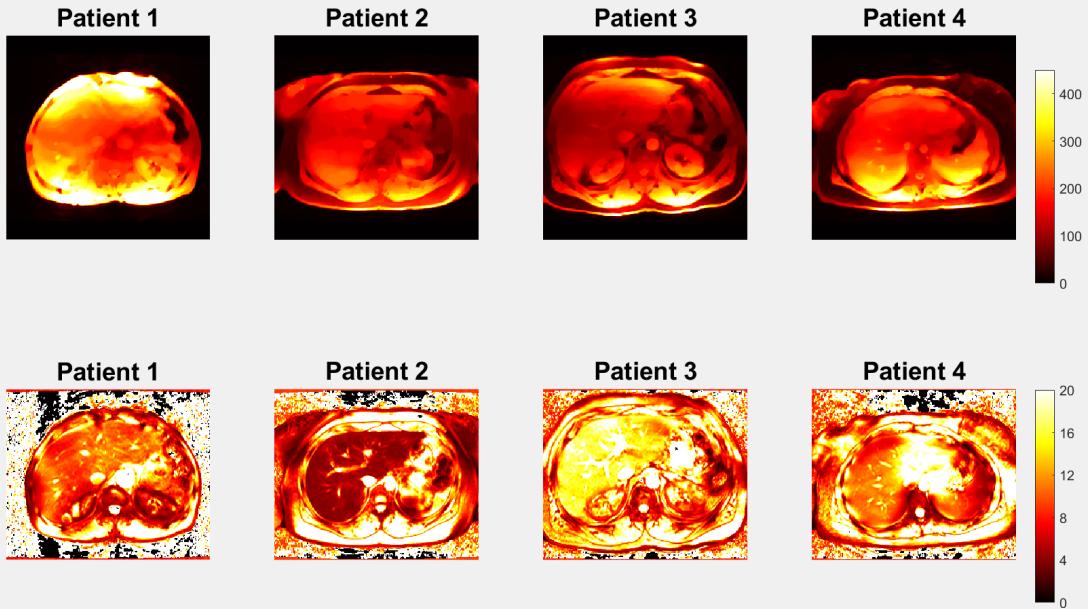


Fig. 21: Parameter maps obtained with ADMM with higher regularisation ($\lambda_1 = 10, \lambda_2 = 10$). We observe smoother regions, less impacted by noise.

Results for a . Method: LM

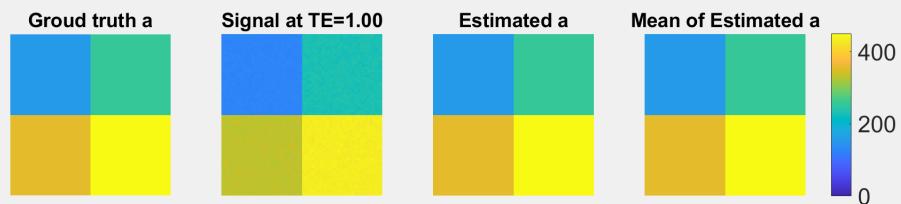


Fig. 22: Results for a obtained using LM. Visually, the estimates seem once again close to the ground truths.

Results for $T2^*$. Method: LM



Fig. 23: Results for $T2^*$ obtained using LM. Visually, the estimates seem once again close to the ground truths.

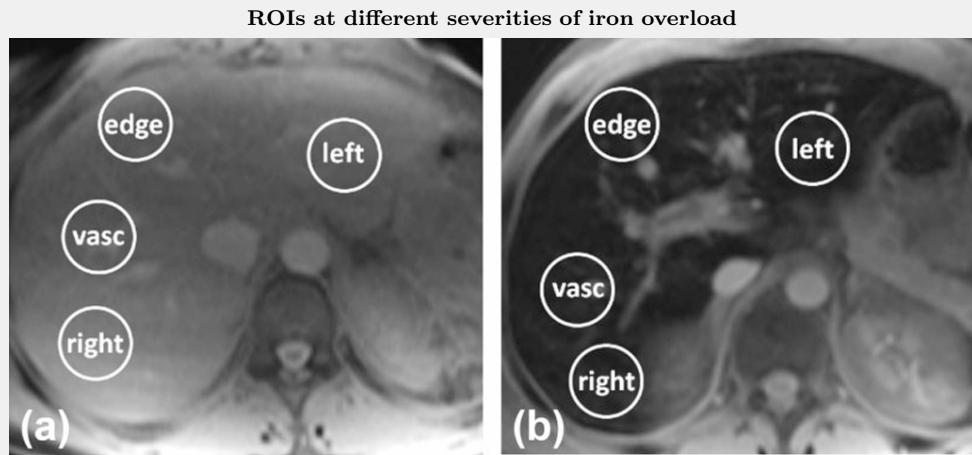


Fig. 24: Example of ROIs for a patient with light (left) and severe (right) iron overload. During the experiments our attempt was to consider ROIs similar to these ones. Image taken from [22]

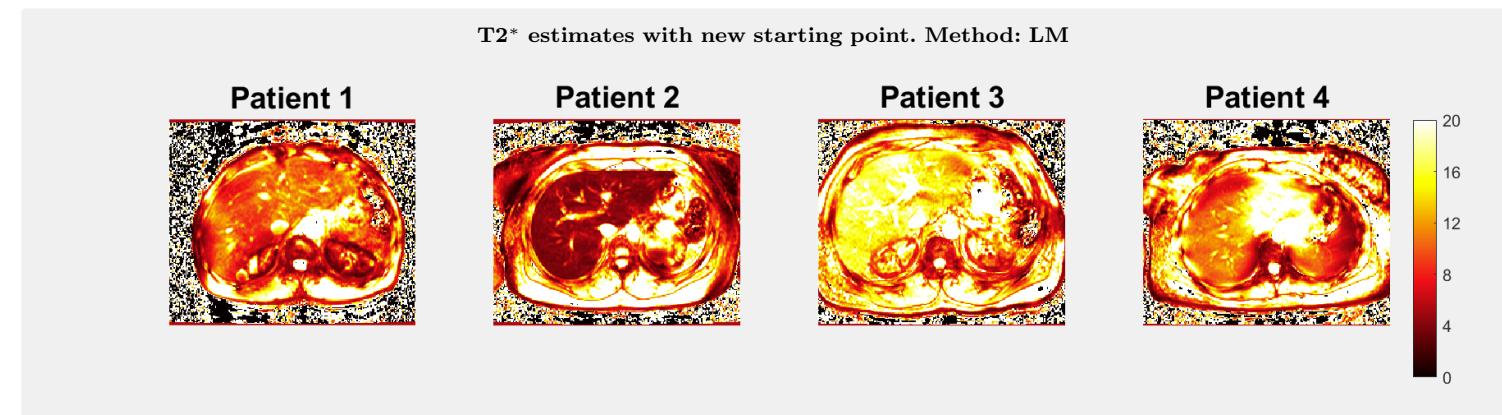


Fig. 25: T2^{*} estimates obtained with LM and starting point [50,0.05]. The different starting point leads to observable differences in the maps.

Table 4. Comparison of scanner results and mean estimated T2* obtained by ADMM with $\lambda_1 = 10$ and $\lambda_2 = 10$ on real images. Green cells indicate estimates that have improved with respect to those obtained with the default hyperparameters ($\lambda_1 = 1e-5$ and $\lambda_2 = 1e-5$). This is the case for 10 out of 16 cells.

	Patient 1	Patient 2	Patient 3	Patient 4
Scanner result	9.1323	4.2624	16.2949	10.7012
Right ROI	10.4761	4.0115	16.6239	11.0310
Vasc ROI	10.6564	5.2298	16.7156	11.0000
Edge ROI	8.1844	4.8931	17.1673	9.9374
Left ROI	10.9702	4.6956	14.2112	18.1914

Table 5. Effect of η in the Levenberg–Marquardt update

Large η (Gradient Descent behavior)	Small η (Gauss–Newton behavior)
$a_n^{(k+1)} = a_n^{(k)} - \frac{1}{\eta} \frac{\partial \mathcal{L}_{LM}}{\partial a_n}$	$a_n^{(k+1)} = a_n^{(k)} - \left(\frac{\partial^2 \mathcal{L}_{LM}}{\partial a_n^2} \right)^{-1} \frac{\partial \mathcal{L}_{LM}}{\partial a_n}$
$r_n^{(k+1)} = r_n^{(k)} - \frac{1}{\eta} \frac{\partial \mathcal{L}_{LM}}{\partial r_n}$	$r_n^{(k+1)} = r_n^{(k)} - \left(\frac{\partial^2 \mathcal{L}_{LM}}{\partial r_n^2} \right)^{-1} \frac{\partial \mathcal{L}_{LM}}{\partial r_n}$

Table 6. Comparison of scanner results and mean estimated T2* with LM on real images. The intensity of the red colour of a cell is proportional to the amount of the total absolute error caused by the estimate in that cell.

	Patient 1	Patient 2	Patient 3	Patient 4
Scanner result	9.1323	4.2624	16.2949	10.7012
Right ROI	10.4905	3.8744	16.3613	10.9869
Vasc ROI	10.8185	5.2674	16.9149	10.9766
Edge ROI	8.1739	4.9463	17.1872	9.9708
Left ROI	11.0596	4.6689	14.1895	18.1107