

THE UNIVERSITY OF NOTTINGHAM  
SIR PETER MANSFIELD IMAGING CENTRE  
SCHOOL OF PHYSICS AND ASTRONOMY

---

---

**Developing Techniques for Quantitative Renal Magnetic Resonance Imaging**

---

*Author:*  
Alexander James DANIEL MSci

*Supervisor:*  
Prof. Susan FRANCIS

24<sup>th</sup> March, 2020



Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy



## **Abstract**

Renal MRI's primary use in the clinic is to provide basic structural information about the kidneys. There is a wealth of quantitative methodology that has been developed for use in other areas of the body that remains largely unadopted by the clinical renal community. Here, multiple quantitative techniques are transferred to the kidneys and compared to the standard clinical pipeline.

# Contents

<b>1</b>	<b>Validating Renal MRI Measurements using a Nephrectomy Model</b>	<b>1</b>
1.1	Introduction . . . . .	1
1.2	Methods . . . . .	3
1.2.1	Sample Acquisition and Fixation . . . . .	3
1.2.2	$T_1$ Mapping . . . . .	3
1.2.3	$T_2$ Mapping . . . . .	5
1.2.4	$T_2^*$ Mapping . . . . .	7
1.3	Results and Discussion . . . . .	8
1.3.1	Fixation and Protocol Development . . . . .	8
1.3.2	Monitoring Changes in MR Parameters Post Fixation . . . . .	10
1.3.3	Comparing MR and Histological Measures in Aged Kidneys . . . . .	14
1.3.4	Comparison of $T_2$ Mapping Methods . . . . .	16
1.4	Conclusions and Future Work . . . . .	23
1.4.1	Protocol Validation on a Single Sample . . . . .	23
1.4.2	Ex-Vivo Sample Coil . . . . .	23
1.4.3	Human Organs . . . . .	24
<b>2</b>	<b>Applying <math>T_2</math> Relaxation Under Spin Tagging (TRUST) To Assess Renal Oxygenation</b>	<b>25</b>
2.1	Introduction . . . . .	25
2.2	Methods . . . . .	28
2.2.1	Susceptibility-Based Oximetry . . . . .	28
2.2.2	$T_2$ Relaxation Under Spin Tagging . . . . .	29
2.2.3	Inducing Changes in Oxygenation of Blood in the Renal Vein . . . . .	35
2.3	Results and Discussion . . . . .	37
2.3.1	Susceptibility-Based Oximetry . . . . .	37
2.3.2	$T_2$ Relaxation Under Spin Tagging . . . . .	38

2.4	Conclusions and Future Work . . . . .	46
<b>3</b>	<b>Automated Segmentation of Kidneys using Machine Learning</b>	<b>47</b>
3.1	Introduction . . . . .	47
3.2	Methods . . . . .	49
3.2.1	Data Acquisition . . . . .	49
3.2.2	Data Pre-Processing . . . . .	49
3.2.3	Network Architecture . . . . .	50
3.3	Results and Discussion . . . . .	52
3.4	Conclusions and Future Work . . . . .	58
<b>4</b>	<b>Other Work</b>	<b>60</b>
4.1	Layer Based Analysis of the Kidneys . . . . .	60
4.2	Diffusion Tensor Imaging in the Kidneys . . . . .	63
4.3	Designing Novel MR Phantoms using Additive Manufacturing . . . . .	65
<b>5</b>	<b>References</b>	<b>68</b>
5.1	List of Acronyms Used . . . . .	78

# **Chapter 1**

## **Validating Renal MRI Measurements using a Nephrectomy Model**

### **1.1 Introduction**

In the clinic, renal pathologies are currently assessed via blood tests, urine tests or biopsy followed by histological analysis. Given that both blood tests and urine tests are indirect measures of the health of the kidneys this means that biopsy is the most accurate diagnostic method routinely used however it is not without its shortcomings. From a patient experience point of view, collecting the tissue sample is an unpleasant and invasive process; from a diagnostic point of view, the results ascertained are not representative of the entirety of the kidney biopsied (typically the left), let alone the other kidney. Additionally, due to the invasive and destructive nature of the procedure, it is not ideal for longitudinal monitoring of renal health. Given these drawbacks in the current methodology, there has been a recent drive to enable the use of Magnetic Resonance Imaging (MRI) for renal diagnosis as it has the potential to be better for both patients and clinicians. A key aspect in the widespread adoption of MRI in renal clinical practice, is a full understanding of the interplay between the current histological pipeline and the newly developed MRI measurements.

The best way to gain this insight is by analysing the same kidney in multiple ways. By taking samples from patients who are undergoing a nephrectomy as part of their standard treatment, it is possible to scan the kidneys in-vivo, remove the kidney, carry out histology and scan the sample ex-vivo. These three complimentary streams of data allow for the comparison of histology and in-vivo MRI with the ex-vivo scans acting as an intermediary, providing high-quality, high-resolution MR data.

Looking in the literature, there are two works that resemble this paradigm, the first, by Friedli et al, correlated  $T_1$  and Apparent Diffusion Coefficient (ADC) with renal interstitial fibrosis and inflammation in a rat model [1]. Establishing a MRI predictor of interstitial fibrosis is especially important as it is present in the majority of renal pathologies and has been established as an excellent indicator of functional recovery [2]. Friedli was able to establish these correlations by scanning the organ in-vivo, ex-vivo and then carrying out histology. It is this sort of methodology we intend to apply to human samples. Another work by Uribe et al carried out a similar process investigating the diagnostic capabilities of ADC and fractional anisotropy in prostate cancer using human samples [3]. Although there has not been much work correlating human kidney MRI with histology, there has been interesting work on ex-vivo MRI samples using similar methods to those we wish to undertake. A very comprehensive paper by Sengupta et al details their ability to image an 80 mm<sup>3</sup> sample of human occipital lobe generating 60  $\mu\text{m}$  isotropic  $T_2^*$  weighted data and 200  $\mu\text{m}$  isotropic quantitative  $T_2^*$  maps [4]. This was achieved on a 9.4T human scanner using a custom made 16 channel phased array and shows the benefits that custom hardware can bring to ex-vivo imaging on human scanners.

As the purpose of this study is to compare pre-existing histological analysis with newly developed, but previously documented, renal MRI protocols, the largest area in need of development is the ex-vivo scanning of samples, as such this is a large focus of this chapter. Here, the aim is to collect ADC,  $T_1$ ,  $T_2$  and  $T_2^*$  data both ex-vivo and in-vivo for comparison. Established protocols exist within the group for ADC,  $T_1$  and  $T_2^*$  collection but there are multiple methods of renal  $T_2$  mapping which have not been optimised [5–8]. Here  $T_2$  mapping methods are compared before choosing one to include in the scan card to perform on nephrectomy patients. As well as the MRI acquisition, attention needs to be paid to the fixing process. It will not be possible to scan unfixed organs therefore, a knowledge of the effects fixation has upon the kidneys needs to be gained. No literature on this topic exists and thus, this is an area we need to explore ourselves.

## 1.2 Methods

### 1.2.1 Sample Acquisition and Fixation

Initial samples were acquired from a local slaughterhouse, placed into Phosphate-buffered Saline (PBS) for transport to the laboratory before being transferred to ten times the samples volume of 10% Neutral Buffered Formalin (NBF) for twenty four hours. The time between slaughter and transferring the samples to NBF was reduced as much as possible to minimise the effects of ischemic injury. Once the samples had been fixed they were washed in fresh PBS and remained in this solution while being scanned to avoid susceptibility artefacts that would be induced by either scanning the samples immersed in NBF or in air. Where possible, both kidneys were collected from the animal, with one being used for MRI and the other being used to biopsy for histological staining. While the samples acquired from the slaughterhouse were useful for early development work there were consistency issues that will be explained in detail in Section 1.3, for this reason a collaboration between Sir Peter Mansfield Imaging Centre (SPMIC) and The University of Nottingham School of Veterinary Medicine and Science has begun. This allows the procurement of much higher quality samples with a greater degree of control.

### 1.2.2 $T_1$ Mapping

#### Acquisition

Ex-vivo  $T_1$  maps can be acquired at field strengths of 3T and 7T. If the time at which the sample is scanned does not need to be precisely controlled or matched with the time of biopsy then the sample can be scanned at both field strengths successively however if scanning and biopsy needed to be time matched then the sample is only scanned at 3T.

$T_1$  maps are produced using an ultrafast gradient echo scheme. By carrying out multiple scans with different Inversion Time (TI) it is possible to sample the inversion recovery of the tissue and as such, estimate  $T_1$ . An example of the acquisitions at each inversion time is shown in Figure 1.1

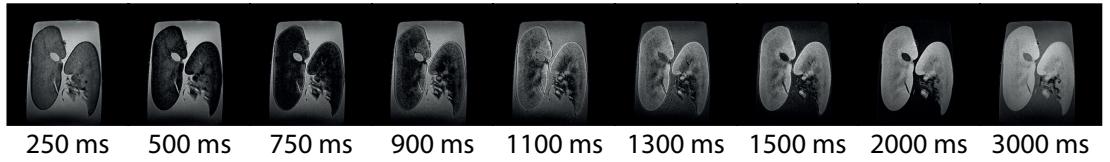


Figure 1.1: Acquisitions at each of the TI at 7T.

The acquisition parameters at 3T are Field Of View (FOV) =  $160 \times 160 \times 50$  mm, voxel size =  $0.7 \times 0.7 \times 1.0$  mm $^3$ , Repetition Time (TR)/Echo Time (TE) = 11 ms/5 ms, Flip Angle (FA) =  $8^\circ$ , Bandwidth = 41.6 Hz/pixel, Turbo Field Echo (TFE) Factor = 64, SENSetivity Encoding (SENSE) Factor = 2.5, Acquisition Time  $\approx$  270 sec per TI collected. The acquisition parameters at 7T are FOV =  $192 \times 170 \times 24$  mm, voxel size =  $0.6 \times 0.6 \times 0.6$  mm $^3$ , TR/TE = 7.2 ms/3.3 ms, FA =  $8^\circ$ , Bandwidth = 56.8 Hz/pixel, TFE Factor = 240, SENSE Factor (P/S) = 2.0/1.5, Acquisition Time  $\approx$  270 sec per TI collected. Initially inversion times of 400 ms, 500 ms, 750 ms, 900 ms, 1100 ms, 1300 ms and 1500 ms were collected at 3T and inversion times of 250 ms, 500 ms, 750 ms, 900 ms, 1100 ms, 1300 ms, 1500 ms, 2000 ms and 3000 ms at 7T however to reduce the scan time of the 3T protocol, this was reduced to inversion times of 400 ms, 500 ms, 750 ms, 900 ms, 1100 ms and 2600 ms. The choice of these inversion times will be elaborated upon later.

## Analysis

The signal recorded at each inversion time is proportional to the modulus of the true longitudinal magnetisation, Figure 1.2. To combat this we can apply polarity correction to the data by saving the phase information and applying the methods of Szumowski et al [9]. This results in a greater dynamic range and thus smaller confidence intervals than using the none polarity corrected data.

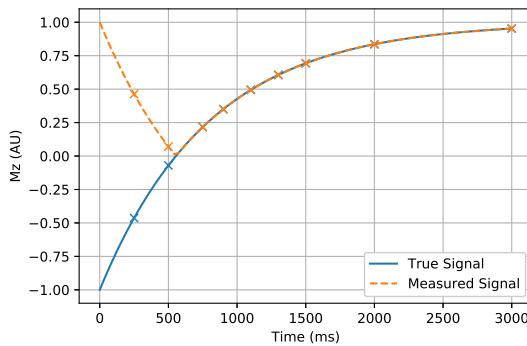


Figure 1.2: A simulation of the true and measured magnetisation for a  $T_1$  of 800 ms. The crosses represent the inversion times at which the inversion recovery is sampled.

Once the data has been polarity corrected a voxel by voxel least squares trust region reflective method is used to fit the data to Equation (1.1) and estimate the  $T_1$  and  $M_0$  of the tissue in that voxel along with an uncertainty in the fit [10]. This data processing is carried out using an in-house Python package. Once the  $T_1$  maps are generated, Region Of Interest (ROI) are defined for the renal medulla and renal cortex and the mean  $T_1$  in these ROI recorded.

$$M_z = M_0 \left( 1 - 2 \cdot e^{-TI/T_1} \right) \quad (1.1)$$

### 1.2.3 $T_2$ Mapping

There are multiple methods for acquiring  $T_2$  maps in the kidneys, Spin Echo (SE), Multi-Echo Turbo Spin Echo (ME-TSE), Gradient Spin Echo (GraSE) and  $T_2$  preparation. We wish to compare each of these methods in-vivo and verify their accuracy using a calibrated phantom. A QalibreMD System Standard Model 130 [11] is used for verification. All  $T_2$  mapping is performed at 3T.

#### Acquisition

**Spin Echo Echo Planar Imaging** A series of volumes are collected at a range of echo times using a multi-slice spin echo acquisition with Echo Planar Imaging (EPI) readout. Acquisition parameters are  $\text{FOV} = 288 \times 288 \times 25 \text{ mm}$ , voxel size =  $3 \times 3 \times 5 \text{ mm}^3$ ,  $\text{TR} = 5000 \text{ ms}$ ,  $\text{FA} = 90^\circ$ ,  $\text{SENSE} = 2.55$ , halfscan = 0.844 and the sequence is respiratory triggered for in-vivo use and has an acquisition time of approximately 9 minutes depending on breathing rate. Volumes are acquired at TE between 20 ms and 70 ms in 5 ms steps with four volumes being acquired at each echo time.

**Multi-Echo Turbo Spin Echo** This method also uses a multi-slice spin echo acquisition however unlike the simple spin echo method, uses a multishot Turbo Spin Echo (TSE) readout. Acquisition parameters for in-vivo scanning are  $\text{FOV} = 288 \times 288 \times 25 \text{ mm}$ , voxel size =  $3 \times 3 \times 5 \text{ mm}^3$ ,  $\text{TR} = 3000 \text{ ms}$ ,  $\text{FA} = 90^\circ$ ,  $\text{SENSE} = 2.55$  and TSE factor = 10; the sequence is respiratory triggered and has an acquisition time of approximately 4 minutes depending on breathing rate. For phantom scanning, the following parameters are modified  $\text{FOV} = 250 \times 250 \times 6 \text{ mm}$ , voxel size =  $0.9 \times 0.9 \times 6 \text{ mm}^3$  and respiratory triggering is removed. Volumes are collected with TE between 13 ms and 130 ms in 13 ms steps.

**Gradient Spin Echo** This method also uses a multi-slice spin echo acquisition but with a GraSE readout. Acquisition parameters for in-vivo scanning are  $\text{FOV} = 288 \times 288 \times 25 \text{ mm}$ , voxel size =  $3 \times 3 \times 5 \text{ mm}^3$ ,  $\text{TR} = 3000 \text{ ms}$ ,  $\text{FA} = 90^\circ$ ,  $\text{SENSE} = 2.55$  and  $\text{TSE factor} = 30$ , startup echoes = 1; the sequence is respiratory triggered and has an acquisition time of approximately 5 minutes depending on breathing rate. For phantom scanning, the following parameters are modified  $\text{FOV} = 250 \times 250 \times 6 \text{ mm}$ , voxel size =  $0.9 \times 0.9 \times 6 \text{ mm}^3$  and respiratory triggering is removed. Volumes are collected with TE between 11.2 ms and 173.8 ms in 5.6 ms steps.

**$T_2$  Preparation** This technique uses a multi-slice Fast Field Echo (FFE) acquisition with a Turbo Field Echo Planar Imaging (TFEPI) readout. Varying degrees of  $T_2$  weighting are applied as a series of  $180^\circ$  preparation pulses for a variable Effective Echo Time (eTE), this is similar to the sequence used in Section 2.2.2. Acquisition parameters for in-vivo scanning are  $\text{FOV} = 288 \times 288 \times 25 \text{ mm}$ , voxel size =  $3 \times 5.65 \times 5 \text{ mm}^3$  (voxel size is limited by the EPI factor),  $\text{TR} = 3000 \text{ ms}$ ,  $\text{TE} = 5.3$ ,  $\text{FA} = 90^\circ$ , EPI factor = 17, SENSE = 3 and halfscan = 0.733; the sequence is respiratory triggered and has an acquisition time of approximately 6 minutes depending on breathing rate. As the voxel size is already at its minimum, the only modifications made for scanning the phantom are to decrease the FOV to  $250 \times 250 \times 6 \text{ mm}$  and remove respiratory triggering. eTEs of 0, 40, 80 and 160 ms are used with three volumes acquired at each eTE

## Analysis

The data is fit on a voxel by voxel basis using a least squares trust region reflective method to fit the data to Equation (1.2) to estimate  $T_2$  and  $S_0$  with an uncertainty in the fit [10]. For methods where multiple volumes are acquired at an TE, individual volumes are used for the fit e.g. four points at each TE for the SE-EPI method rather than taking the mean of the volumes for each TE, this makes potential data discarding easier. This post-processing is performed by an in-house Python package. Once the  $T_2$  maps have been generated, ROI can be defined for different tissue types or phantom compositions.

$$S(t) = S_0 \cdot e^{-t/T_2} \quad (1.2)$$

### 1.2.4 $T_2^*$ Mapping

#### Acquisition

Ex-vivo  $T_2^*$  maps can be acquired at both 3T and 7T using a multi-slice FFE sequence with scans being performed at a range of different echo times. An example of the acquisition at each echo time is shown in Figure 1.3

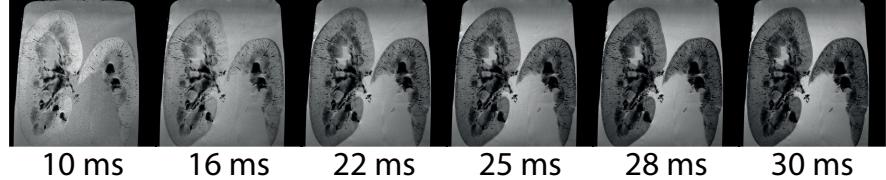


Figure 1.3: Acquisitions at each of the TE at 7T.

The acquisition parameters at 3T are  $\text{FOV} = 145 \times 145 \times 15 \text{ mm}$ , voxel size =  $0.6 \times 0.6 \times 1.5 \text{ mm}^3$ ,  $\text{TR} = 697 \text{ ms}$ ,  $\text{FA} = 38^\circ$ , SENSE Factor = 2.0, Acquisition Time  $\approx 180 \text{ sec}$  per TE collected. The acquisition parameters at 7T are  $\text{FOV} = 145 \times 145 \times 10 \text{ mm}$ , voxel size =  $0.5 \times 0.5 \times 1.0 \text{ mm}^3$ ,  $\text{FA} = 38^\circ$ , SENSE Factor = 2.0, Acquisition Time  $\approx 180 \text{ sec}$  per TE collected. Initially echo times were 15 ms, 20 ms, 25 ms, 30 ms, 35 ms, 40 ms and 50 ms at 3T and 10 ms, 16 ms, 22 ms, 25 ms, 28 ms and 30 ms at 7T however to reduce acquisition times, the 3T echo times were reduced to 15 ms, 20 ms, 25 ms, 40 ms and 50 ms.

#### Analysis

The data is fit voxel by voxel using a weighted echo time fit from the log of the exponential signal decay (Equation (1.3)) to generate the  $T_2^*$  maps [12]. This data processing is carried out using an in-house Python package. The ROI were defined using the  $T_1$  weighted data if available as it has a greater cortical medullary contrast at shorter times from fixation. The mean  $T_2^*$  in these ROI is recorded.

$$S(t) = S_0 \cdot e^{-TE/T_2^*} \quad (1.8)$$

## **1.3 Results and Discussion**

### **1.3.1 Fixation and Protocol Development**

As was alluded to in Section 1.2.1, there was significant variability in the quality of the samples collected from the slaughterhouse. This was largely due to the legislation surrounding animals slaughtered to enter the human food chain. If any part of the animal is destined for human consumption then the carcase must be thoroughly inspected before any tissue can be released. This can cause two issues. As part of the inspection, the kidneys need to be examined for parasites, this is done by making an incision in the organ, however, the quality of this incision can vary massively with some samples having a 20 mm slice cut into them while others are roughly cut in half. The second issue is caused by the variable time between slaughter and the tissue being released after inspection. For these reasons kidneys began to be procured from Veterinary Science collaborating with Prof David Gardner. The animals slaughtered there are not destined for human consumption and as such the kidneys can be placed into PBS and subsequently NBF far quicker and the kidneys do not need to be sliced open for inspection. The difference in the samples from these two sources can clearly be seen in Figure 1.4. This collaboration also enables the procurement of kidneys from a greater range of animals including different ages of pigs and therefore different degrees of fibrosis and inducing Acute Kidney Injury (AKI) in the animals prior to scanning and histology. Veterinary Science can also carry out the histology in house, thus streamlining the protocol by avoiding transporting one kidney to Derby for histology and the other to SPMIC for scanning.

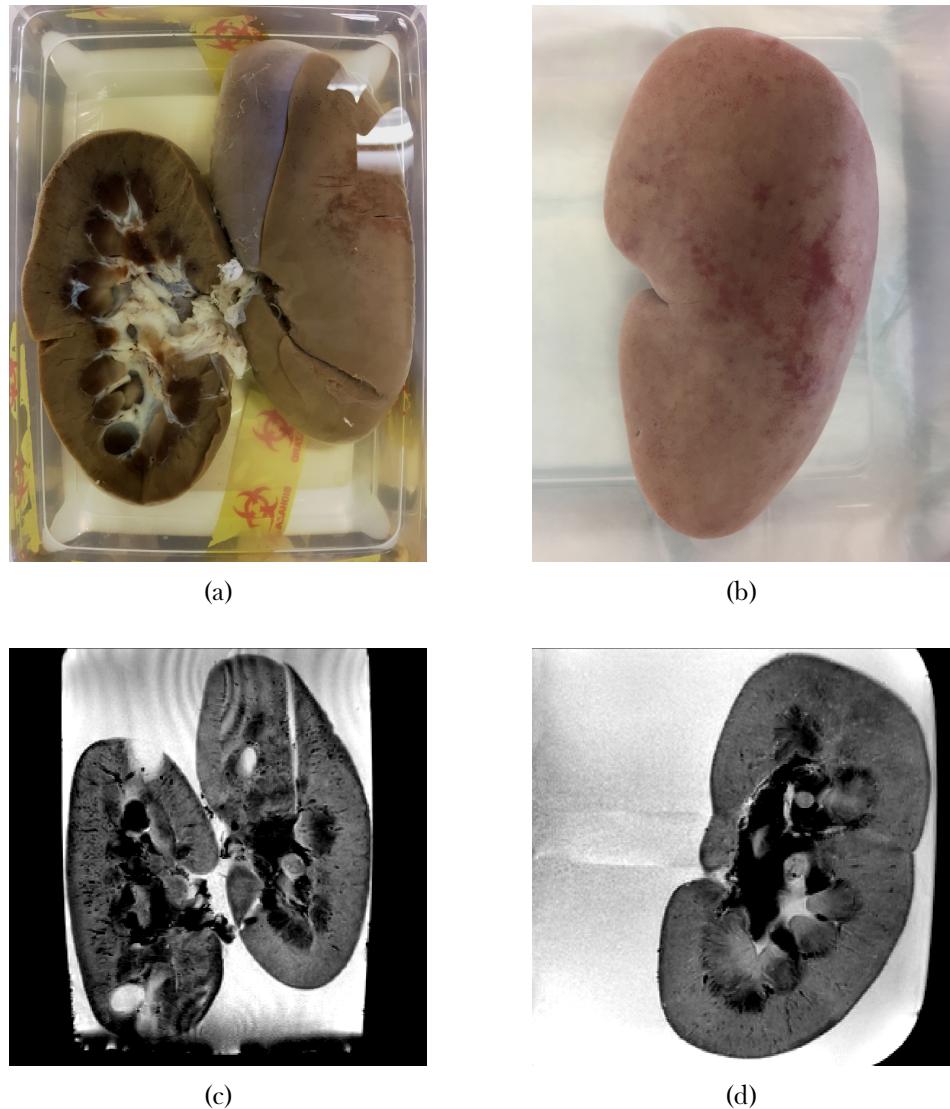


Figure 1.4: (a) An example of a sample procured from the slaughterhouse after it has been fixed. The left hand kidney has been sliced in half; the right hand kidney has the incisions from the meat inspector clearly visible. (b) An example of a sample procured from Veterinary Science post fixing. (c) An example of a  $T_2$  weighted FFE with TE = 40 ms of a kidney procured from the slaughterhouse. (d) An example of a  $T_2$  weighted FFE with TE = 40 ms of a kidney procured from Veterinary Science.

Once samples had been fixed and transferred to PBS they could be placed into the scanner. Using the protocols outlined above it was possible to generate maps as shown in Figure 1.5.

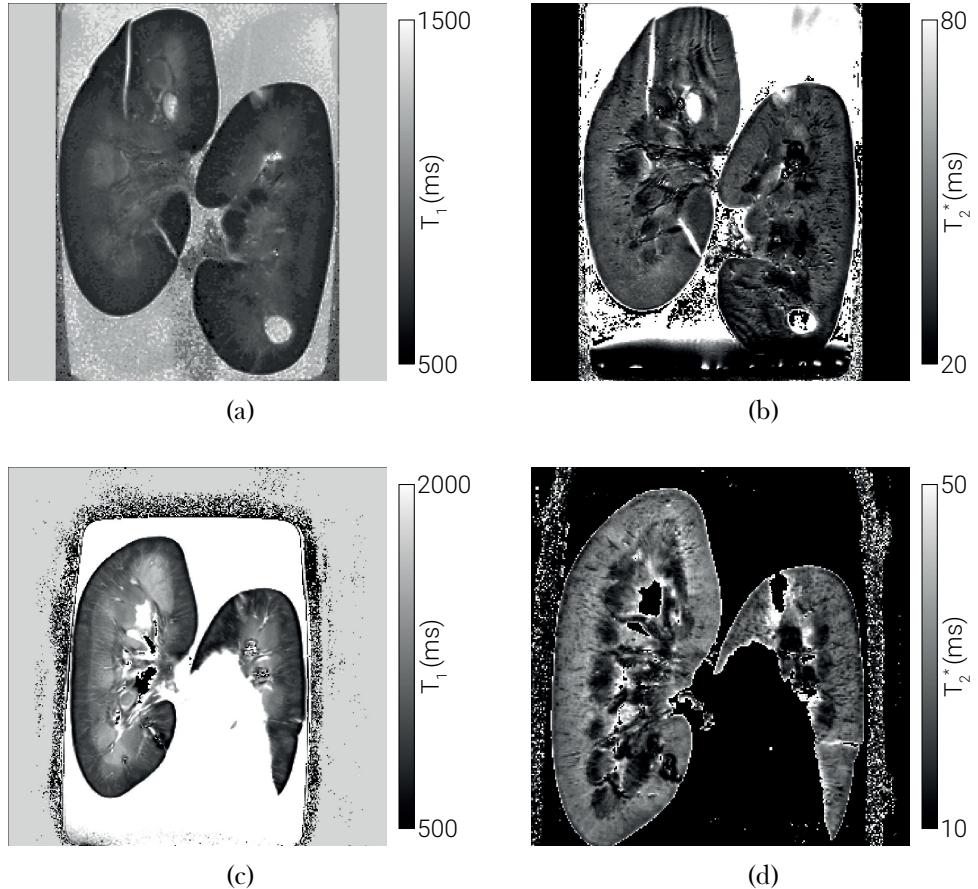


Figure 1.5: (a)  $T_1$  map of a kidney twenty four hours after fixation at 3T (b)  $T_2^*$  map twenty four hours after fixation at 3T (c)  $T_1$  map twenty four hours after fixation at 7T (d)  $T_2^*$  map twenty four hours after fixation at 7T. The sample shown in the 3T and 7T figures is different.

### 1.3.2 Monitoring Changes in MR Parameters Post Fixation

To study the effects of the fixation process upon MR measurements, a kidney was fixed as per the method in Section 1.2.1 and scanned at both field strengths of 3T and 7T. Collecting a  $T_1$  and  $T_2^*$  map took approximately 90 minutes per field strength. Once the maps had been generated, ROI were defined and the mean  $T_1$  and  $T_2^*$  for the cortex and the medulla were calculated. The sample was monitored for ten weeks. The variation in  $T_1$  and  $T_2^*$  over time can be seen in Figure 1.6.

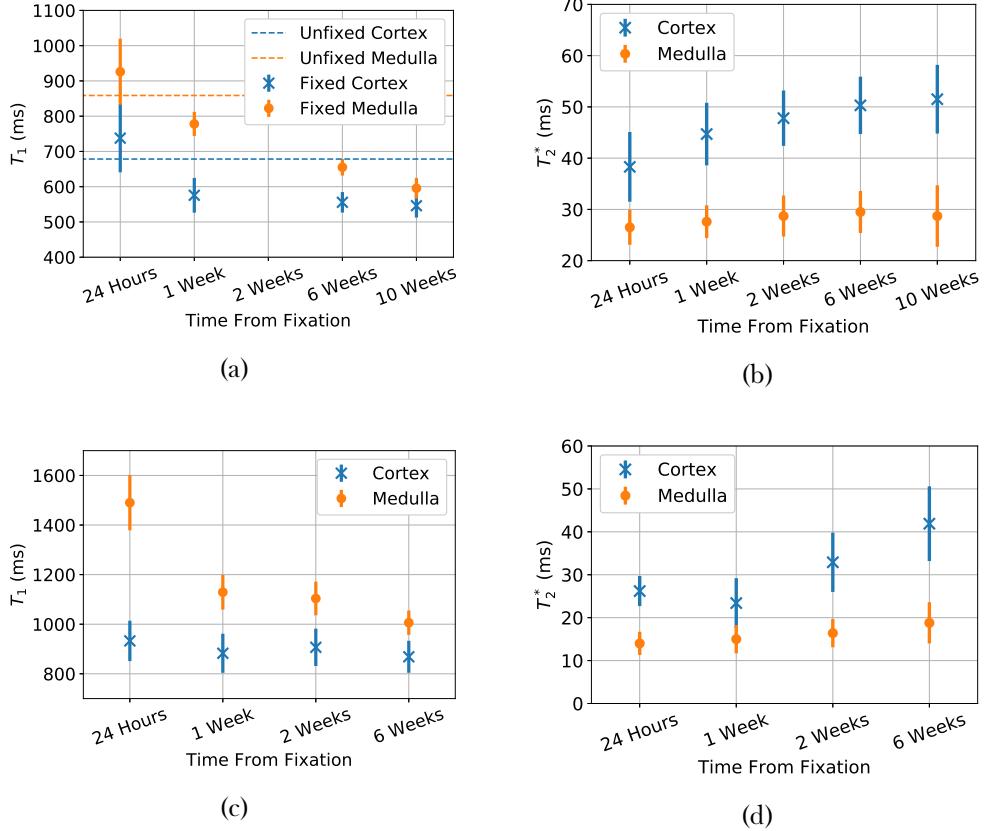


Figure 1.6: (a) Variation in  $T_1$  as a function of time after fixation measured at 3T (b) Variation in  $T_2^*$  as a function of time after fixation measured at 3T (c) Variation in  $T_1$  as a function of time after fixation measured at 7T (d) Variation in  $T_2^*$  as a function of time after fixation measured at 7T.

Unfortunately due to technical scanner issues, we were not able to scan the sample at 7T at ten weeks and the quality of the 3T  $T_1$  acquisition at two weeks was significantly inferior; as such these data points have been omitted. It can be seen that the largest changes in  $T_1$  and  $T_2^*$  occur between twenty four hours and one week after fixation, after that there is a general trend that the  $T_1$  of the cortex and medulla converge while the  $T_2^*$  of each tissue type diverges, one could argue that the  $T_2^*$  of the cortex measured at 3T is plateauing. This means that, although the samples will reach a steady state, in the first few weeks after fixation, their  $T_1$  and  $T_2^*$  will have a dependence on time. This necessitates the need to standardise the protocol, specifically the time at which the samples are scanned. It would have been useful to know the  $T_1$  and  $T_2^*$  of unfixed porcine kidneys and as such, a fresh, unfixed kidney was scanned using the same protocol. Unfortunately, due to the difference in stiffness between fixed and unfixed kidneys, the same protocol did not deliver usable  $T_2^*$  data as the unfixed kidney vibrated too much while floating in the PBS. This problem could potentially be reduced by either vibration insulation between the sample and the scanner as per Dawe et al [13] or by embedding the sample in an agarose medium rather than allowing it to float

in PBS as per Kolk et al[14]. The  $T_1$  of the unfixed kidney was seen to be between that of the fixed kidney between 24-hours and one week.

To investigate this time dependence over a shorter time scale, a pair of kidneys were acquired and fixed as before. It will be possible to scan most human samples within 24 hours of fixation, as such it is desirable to see how much  $T_1$ ,  $T_2^*$  and histology change over this period. Scanning was only carried out at 3T as more frequent measurements were preferable to measurements at different field strengths. For this reason the number of inversion times and echo times used to generate the  $T_1$  and  $T_2^*$  maps was reduced to TI = 400 ms, 500 ms, 750 ms, 900 ms, 1100 ms, 2600 ms and TE = 15 ms, 20 ms, 25 ms, 40 ms, 50 ms respectively. The choice of these inversion/echo times was arrived at empirically by carrying out the analysis pipeline on a single slice of data using each combination of six and five of the previously used TIs and TEs respectively. The results from each combination of TI/TE were compared to those generated when using the full complement of inversion/echo times and the combination delivering the minimum difference chosen. This reduction in the number of inversion times resulted in a mean error per voxel in the kidney of  $20.9 \pm 12.8$  ms if the full complement of inversion times was taken as the ground truth, the reduction in echo times resulted in a mean error of  $0.3 \pm 1.2$  ms.

Scanning sessions started at 1.5 hours, 2.5 hours, 4 hours, 5.5 hours, 19 hours and 22 hours after the sample was removed from the NBF. Due to the potential for the properties to change relatively quickly, especially  $T_1$ , it was decided to randomise the order in which the inversion/echo times were collected, this way any change in  $T_1/T_2^*$  over the 30/20 minute acquisition period would manifest itself as non-systematic noise and thus will increase the uncertainty in the fit rather than affecting the predicted value. At the start of each scanning session, a biopsy was performed on the kidney not being scanned. Masson's trichrome and Haematoxylin and Eosin (H and E) staining was performed on these samples.

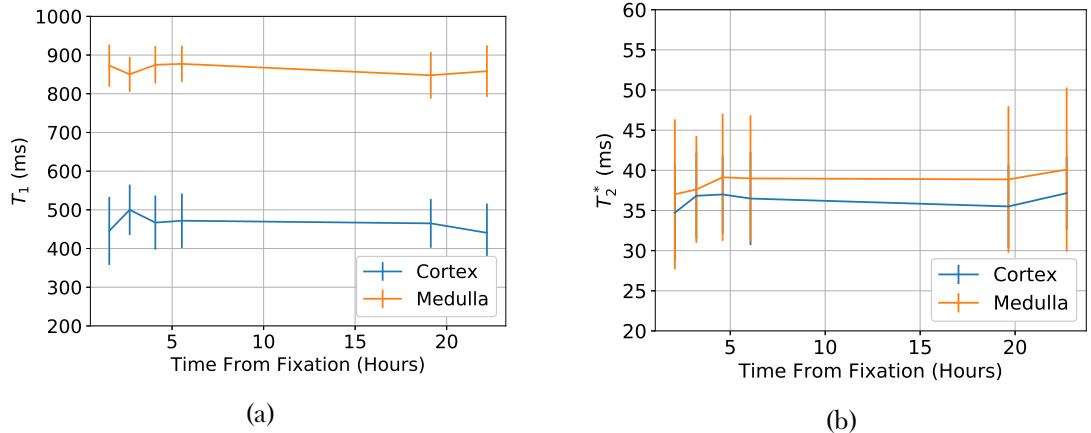


Figure 1.7: (a) Variation in  $T_1$  as a function of time after fixation measured at 3T (b) Variation in  $T_2^*$  as a function of time after fixation measured at 3T.

No significant change in  $T_1$  or  $T_2^*$  was observed over the period the sample was monitored. This is promising as it means that when this protocol is applied to human samples, there will be a relatively large time window in which the ex-vivo scan can be carried out, making the experimental procedure simpler. The corresponding histology results showed no change in the cortex over this period however there was a noticeable inflammatory response in the medulla.

It was noted that the sample used for this experiment was not of especially high quality, the kidney has two slices in it, one that almost bifurcated the sample along the coronal plane, another cut down one half of the sample along the sagittal plane, visible in Figure 1.8a. These cuts meant that air became trapped within the sample causing it to float in the PBS. If not corrected this would cause large susceptibility artefacts where the sample came into contact with the air at the top of the PBS; to remedy this the sample was entirely bifurcated. Despite the best efforts of investigators, air bubbles remained in the sagittal slice, causing the aforementioned artefacts, especially visible in Figure 1.8b. Given these concerns over sample quality, it was decided to scan the sample again one week after fixation, to match the time period in the previous experiment (Figure 1.6).

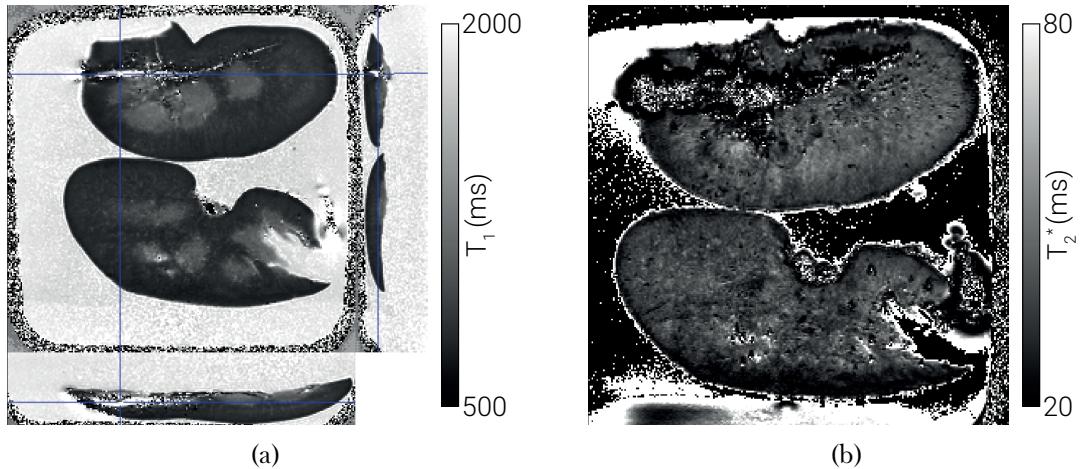


Figure 1.8: (a) An example of the  $T_1$  map collected from the short time scale kidney (b) An example of the  $T_2^*$  map collected from the short time scale kidney.

It was expected that upon repeating the measurements one week later, the  $T_1$  of both cortex and medulla would decrease and the  $T_2^*$  of the cortex would increase. This was not the case, there was a slight decrease in the  $T_1$  of the medulla but otherwise, no change was observed, Figure 1.9. This lead us to conclude that the large cuts in the sample had lead to a different level of fixation. Subsequent to this, samples were to be procured from Veterinary Science as they are more consistent and therefore more akin to the human samples that will be used.

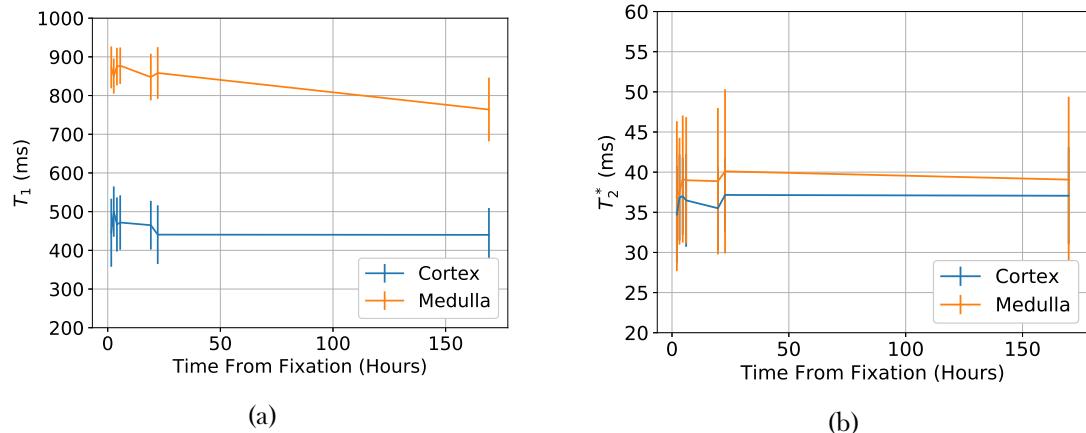


Figure 1.9: (a) Variation in  $T_1$  as a function of time after fixation measured at 3T (b) Variation in  $T_2^*$  as a function of time after fixation measured at 3T.

### 1.3.3 Comparing MR and Histological Measures in Aged Kidneys

To verify the correlation of MR measurements with histology, kidneys were collected from a 0.5 year old and 2.5 year old pig. These different ages were expected to have

differing levels of renal inflammation and fibrosis. Figure 1.10 shows example MRI data collected from these samples, Figure 1.11 shows the quantitative differences in  $T_1$  and  $T_2^*$  between the two samples.

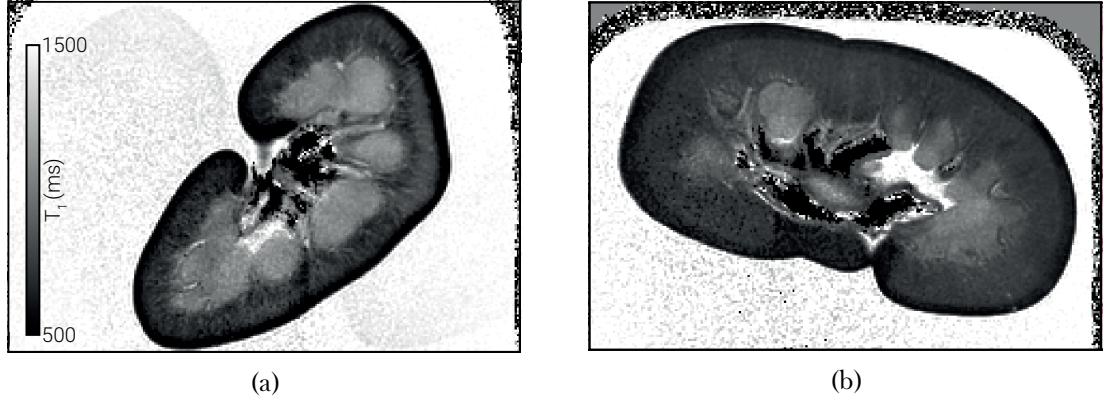


Figure 1.10: (a)  $T_1$  map of a 0.5 year old pig kidney. (b)  $T_1$  map of a 2.5 year old pg kidney.

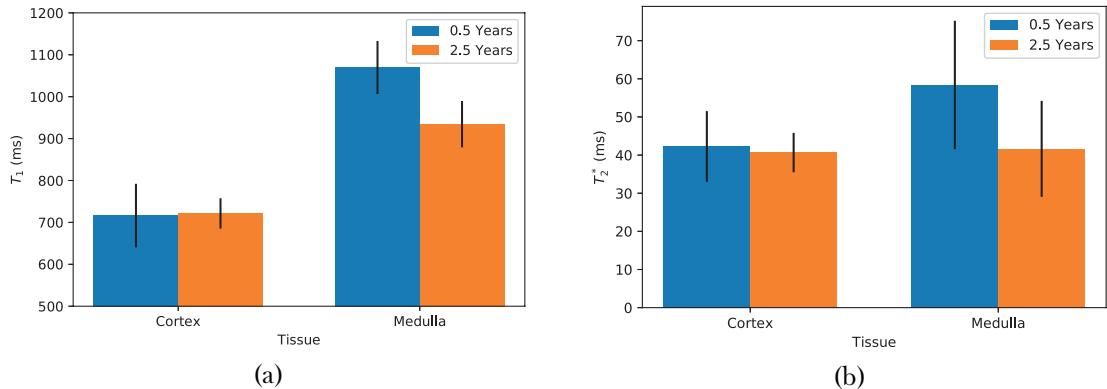


Figure 1.11: (a) The  $T_1$  of the renal cortex and medulla of the two samples. (b) The  $T_2^*$  of the renal cortex and medulla of the two samples.

No significant change is observed in the  $T_1$  or  $T_2^*$  of cortex the two kidneys. There is however a decrease in  $T_1$  seen in the medulla of the older kidney. Cortical samples were removed from the same animals for histological analysis. These samples were stained using H and E and Masson's trichrome to enable the evaluation of levels of fibrosis, these micrographs are shown in Figure 1.12.

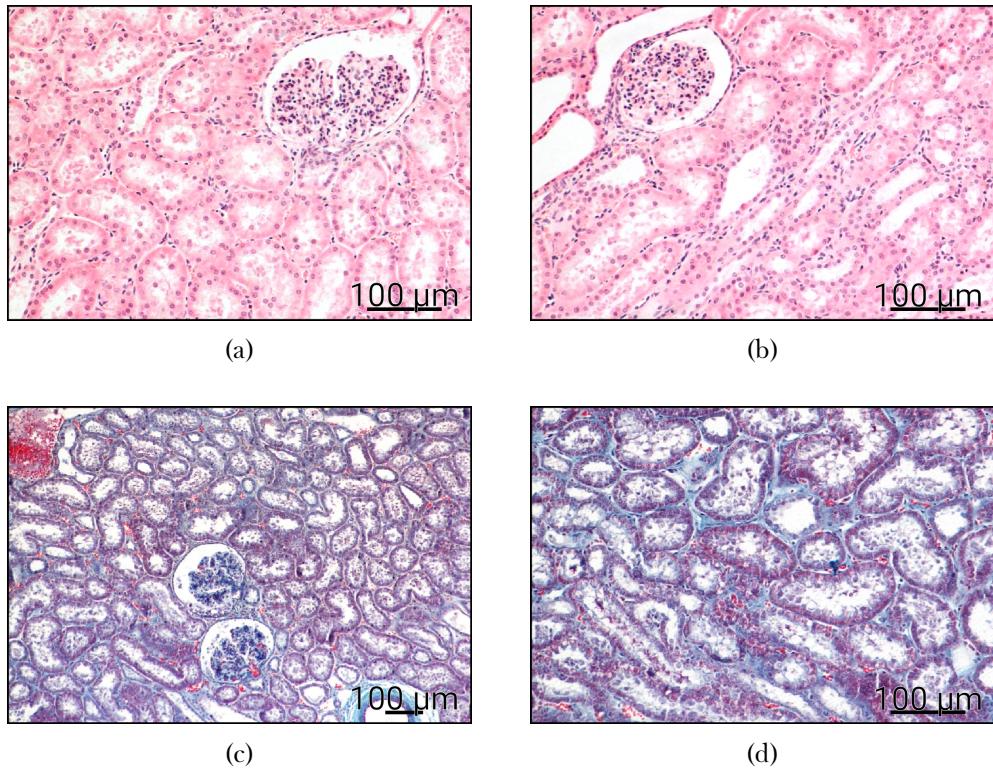


Figure 1.12: (a) A sample of renal cortex from a 0.5 year old pig stained with H and E. (b) A sample of renal cortex from a 2.5 year old pig stained with H and E. (c) A sample of renal cortex from a 0.5 year old pig stained with Masson's trichrome. (d) A sample of renal cortex from a 2.5 year old pig stained with Masson's trichrome.

No significant difference is seen between the histology of these cortical samples. This means that MRI and histology are in agreement. Unfortunately no samples were taken from the renal medulla, the area which showed a change in MR measurements. In future samples with a larger difference in age should be used as these will have a greater difference in fibrosis and samples should be taken from the medulla for histological analysis.

#### 1.3.4 Comparison of $T_2$ Mapping Methods

##### Quantitative Phantom

We began by comparing the quantitative accuracy of each of the proposed  $T_2$  mapping methods (spin echo EPI, multi-echo TSE, GraSE and  $T_2$  preparation). The QalibreMD System Standard Model 130 phantom used has fourteen spheres with  $T_2$  between 5.35 ms and 645.8 ms spanning the range of  $T_2$  expected in the kidneys, Figure 1.13. This phantom was scanned using each of the methods outlined in 1.2.3 and a ROI defined for each sphere.

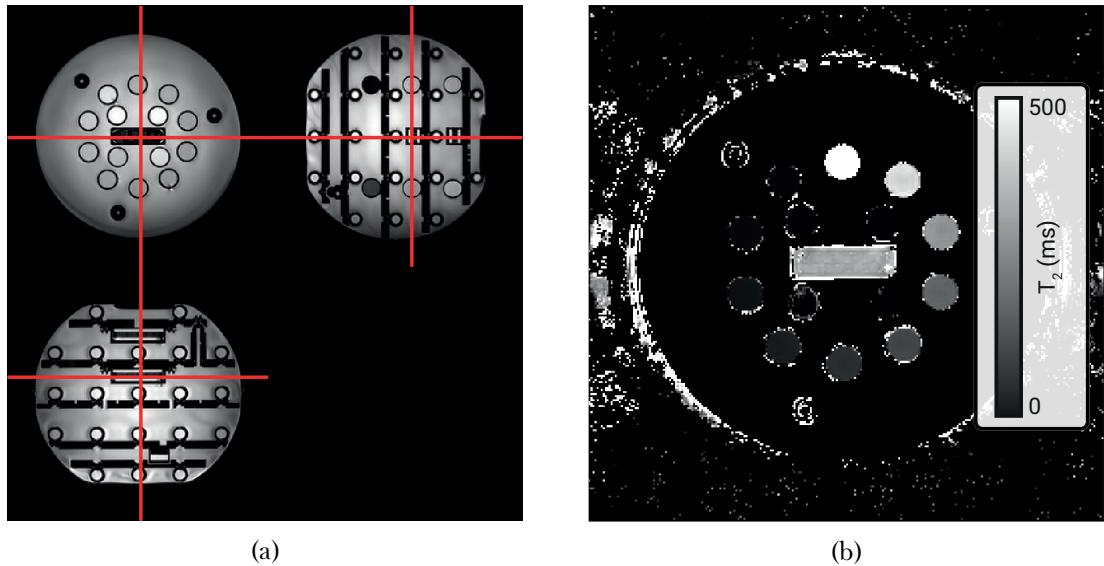


Figure 1.13: (a) The  $T_2$  spheres inside the phantom. (b) An example  $T_2$  map, in this case generated using the GraSE method.

The spin echo method produced vastly over-estimated readings for the spheres of  $T_2$  less than 20 ms (Figures 1.14b) due to the longer minimum TE compared to the other methods. This can be seen in Figure 1.14a where the signal from the shortest  $T_2$  spheres has already mostly decayed. This method did however deliver accurate measurements for the remaining spheres.

More accurate results were generated for shorter  $T_2$  spheres using the ME-TSE method. The raw data (Figure 1.14c) is more noisy, with a sawtooth pattern visible, this additional noise manifests itself as inaccuracies in the longer  $T_2$  measurements where the dynamic range over the TE sampled is smaller.

The GraSE method produced the most accurate measurements although still struggled to measure the sphere with a  $T_2$  of 5.35 ms, Figure 1.14f. The large range of TE and number of volumes collected means this method produced the most accurate results. It also has the benefit of being able to be performed at high resolutions with voxel sizes of 0.9 x 0.9 mm unlike the SE-EPI and  $T_2$  prep methods; this makes it well suited to both in-vivo and ex-vivo measurements.

The data collected using the  $T_2$  prep method (Figures 1.14g and 1.14h) did not fit well due to its small number of data points and large degradation in image quality.

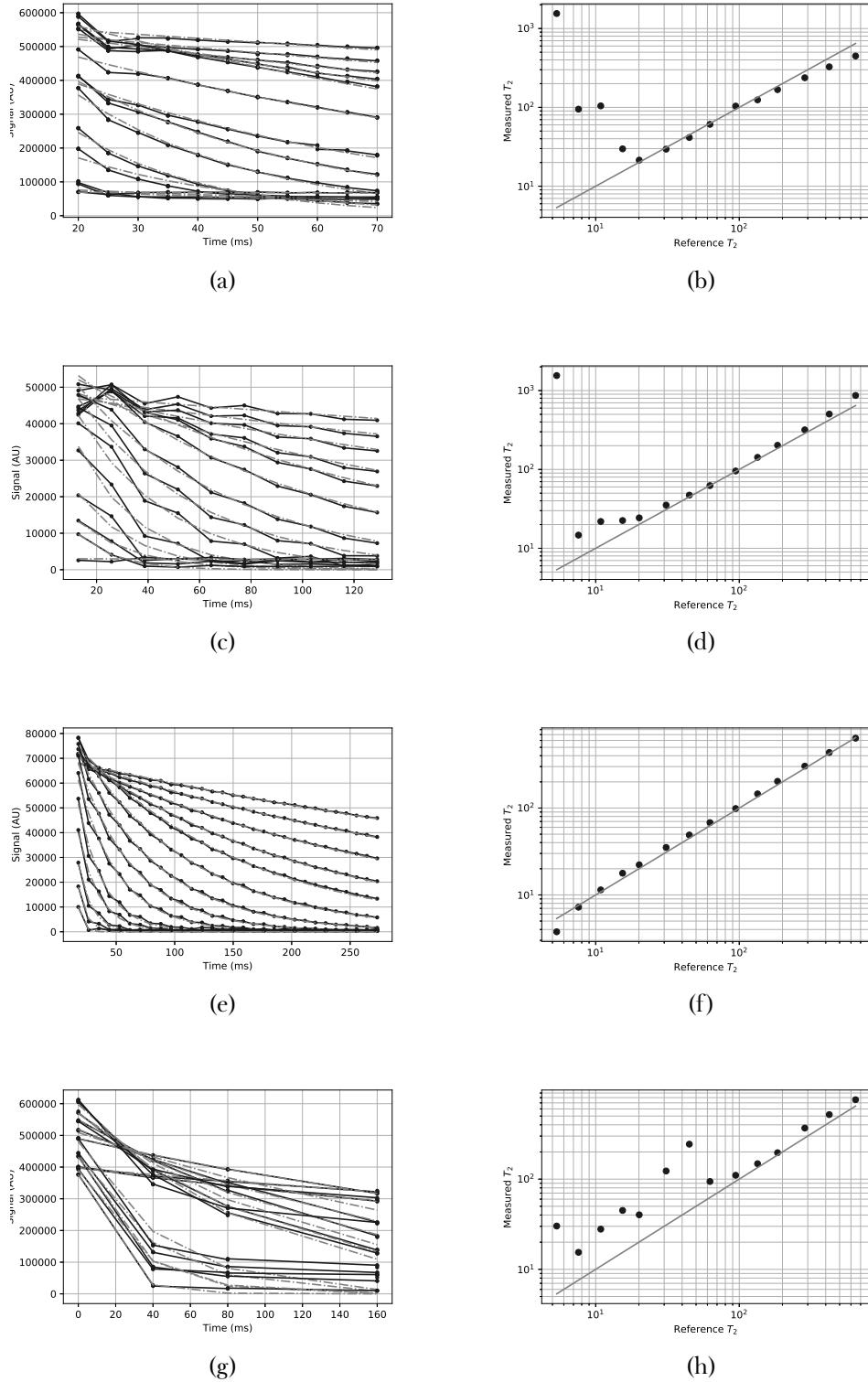


Figure 1.14: Figures (a), (c), (e) and (g) show the raw signal decay for each of the fourteen spheres and the fit decay. Figures (b), (d), (f) and (h) show how the fit  $T_2$  compares to the literature value. Figures (a) and (b) show the results from the SE-EPI method, (c) and (d) show the results form the ME-TSE method, (e) and (f) show the results form the GraSE method and (g) and (h) show the results from the  $T_2$  prep method.

## In-Vivo

$T_2$  maps using all four methods were collected on the same subject in the same scanning session to allow for a direct comparison of the in-vivo data.

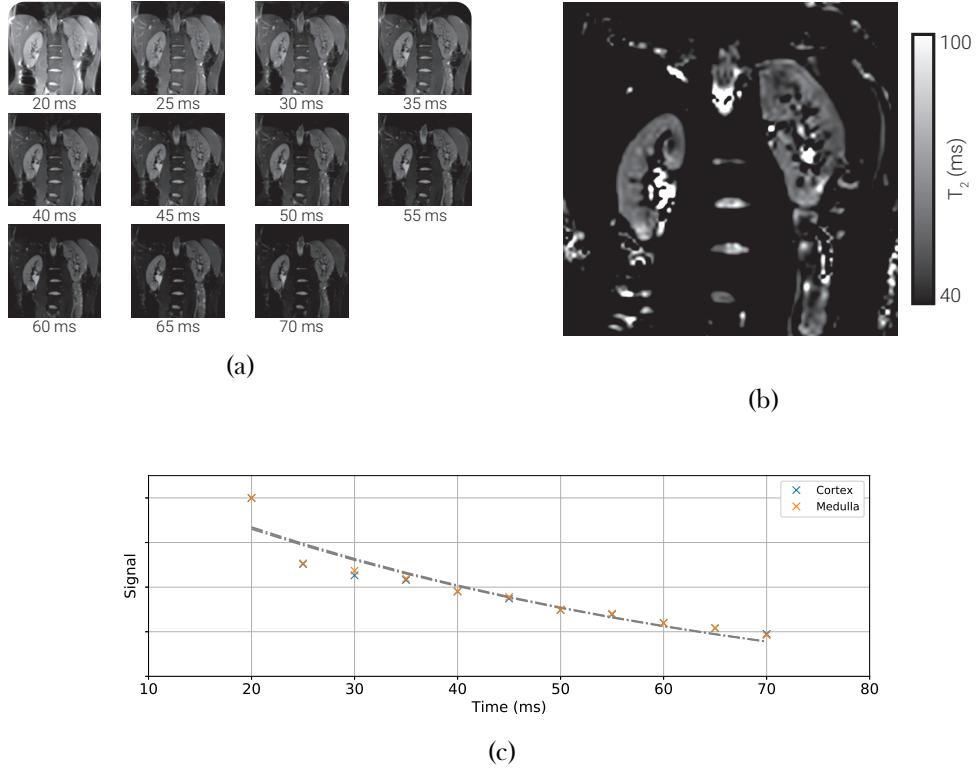


Figure 1.15: (a) The raw data used to generate the SE-EPI  $T_2$  map. (b) An example slice from the SE-EPI  $T_2$  map. (c) The signal decay for the renal cortex and medulla.

The SE-EPI method (Figure 1.15) generated maps with little blurring however there is also a lack of differentiation in  $T_2$  between the renal cortex and medulla. The data collected at TE of 20 ms appears to be artificially high and leads to a reduction in fit  $T_2$ . This sequence is the most susceptible of the methods to patient motion due to the acquisition method of a series per TE, this increase in motion is clear when scrolling through TE.

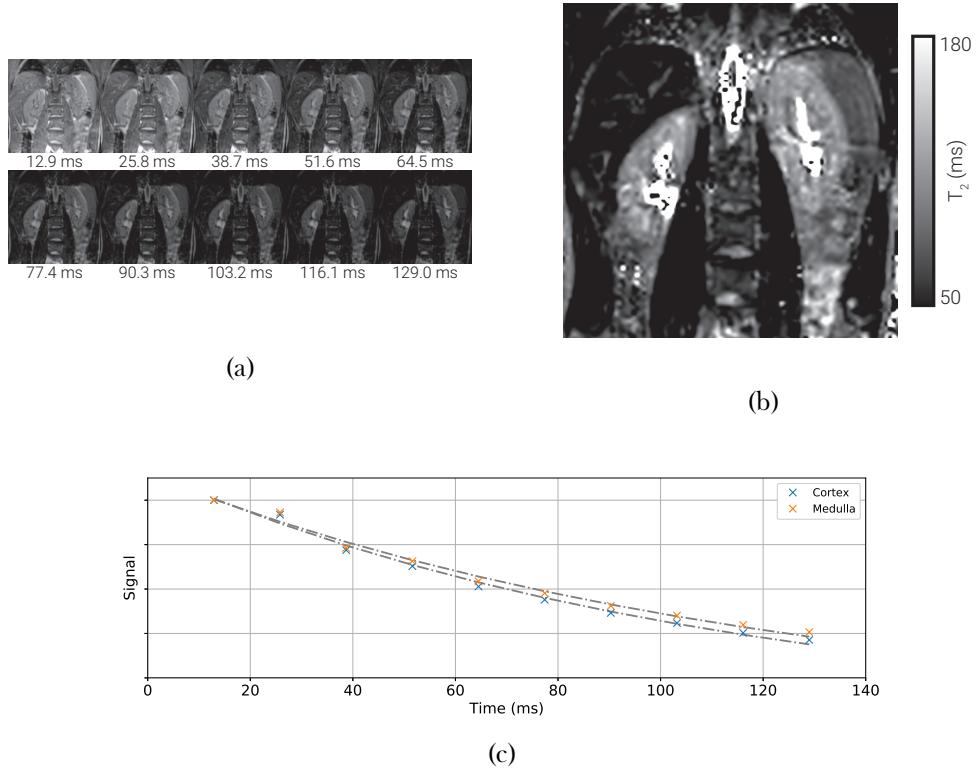


Figure 1.16: (a) The raw data used to generate the ME-TSE  $T_2$  map. (b) An example slice from the ME-TSE  $T_2$  map. (c) The signal decay for the renal cortex and medulla.

The map generated by the ME-TSE method (Figure 1.16) suffers from a large amount of blurring due to the relatively long echo train length. The number of echoes acquired is limited to the TSE factor therefore to acquire ten echoes, a TSE factor of ten needs to be used. This blurring leads to structures being obscured in the map and only a very small differentiation between cortex and medulla.

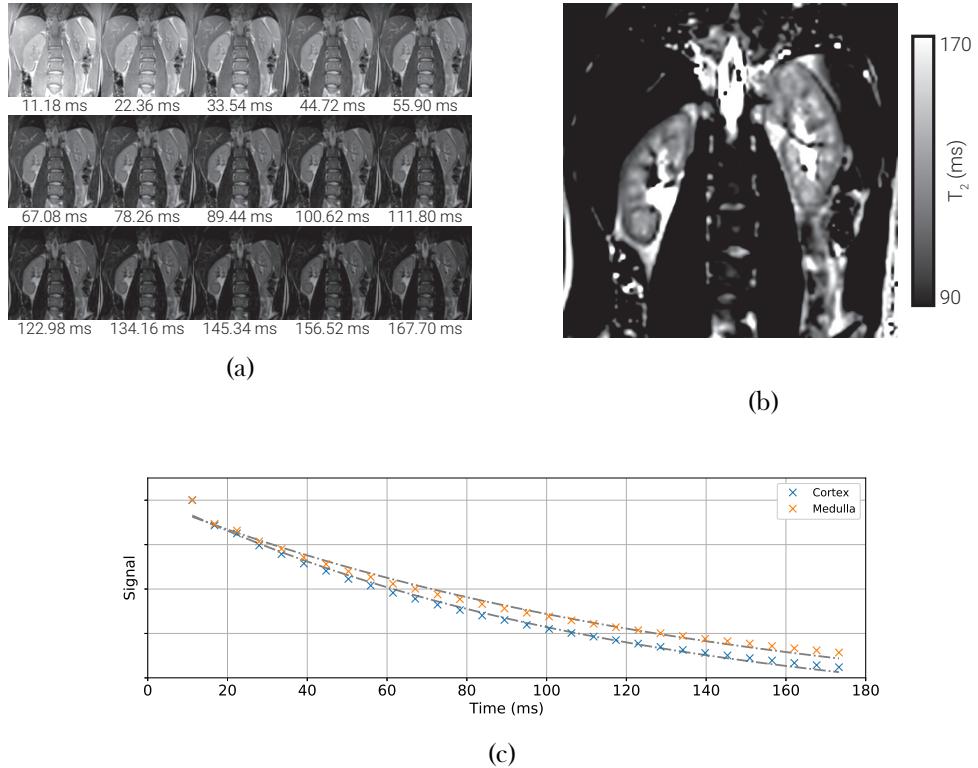


Figure 1.17: (a) The raw data used to generate the GraSE  $T_2$  map. (b) An example slice from the GraSE  $T_2$  map. (c) The signal decay for the renal cortex and medulla.

Using the GraSE method the data in Figure 1.17 was collected. There is a clear difference between cortical and medullary  $T_2$  and the data fits well to a  $T_2$  decay (Figure 1.17c). The signal from the first echo in Figure 1.17c is too intense, this effect was even more pronounced when no startup echoes were used. For tissues with a longer  $T_2$  using two startup echoes would be preferable however this makes measurements of tissues with a short  $T_2$  more inaccurate, as such a compromise of a single startup echo was used. The short echo-spacing made possible by GraSE means more TE can be sampled and therefore leads to a more accurate fit.

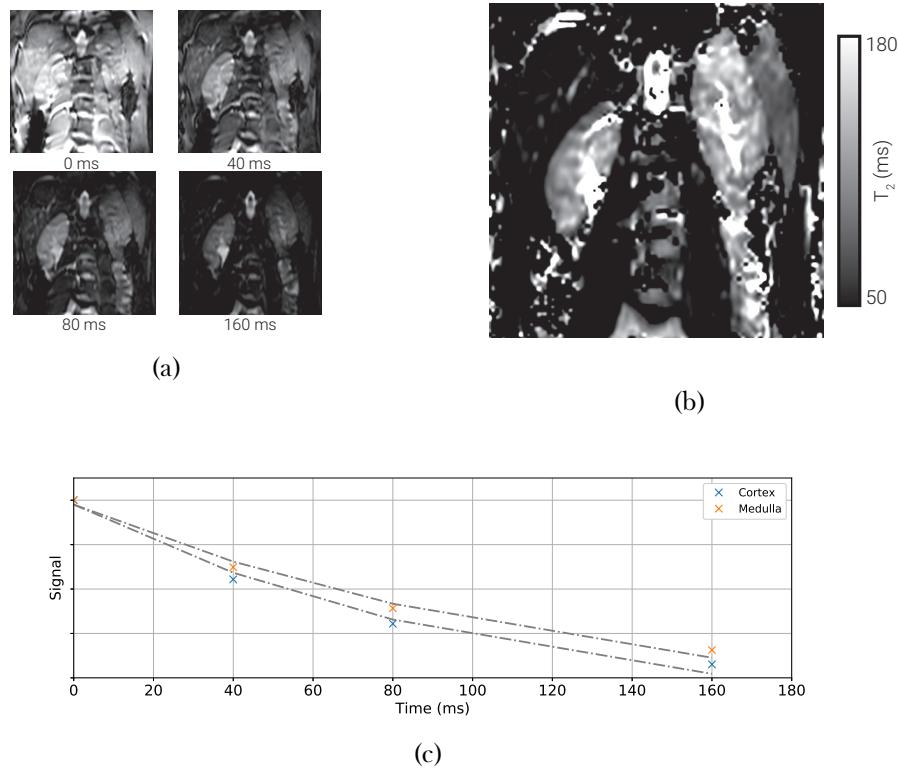


Figure 1.18: (a) The mean data at each eTE used to generate the  $T_2$  preparation  $T_2$  map. (b) An example slice from the  $T_2$  preparation  $T_2$  map. (c) The signal decay for the renal cortex and medulla.

The map made using the  $T_2$  preparation method (Figure 1.18) suffers from noise in the raw data, this is despite there being three acquisitions at each eTE. When comparing Figure 1.18b to Figure 1.17b it's possible to see that some of the areas of greater  $T_2$  do match with the medulla, however the degree of noise in 1.18b means it is un-useable on its own. The small number of eTE collected means the uncertainty in the fit  $T_2$  is higher for this method.

The two methods that have delivered the highest image quality, SE-EPI and GraSE, produce substantially different values of  $T_2$  in-vivo. Even when the data from the 20 ms volume is omitted from the SE-EPI fit, the  $T_2$  is far lower. This is surprising give that when deployed on the phantom, this protocol delivered accurate results over the range of  $T_2$  we see in the kidneys. This disparity is due to the additional confounding factors of diffusion and flow that are present in the body. These factors do not affect the GraSE sequence to the same degree as the SE-EPI sequence.

Of the methods explored, the GraSE sequence produced the most accurate results on the phantom and superior image quality in-vivo, we will use this sequence in  $T_2$  mapping

going forward.

## 1.4 Conclusions and Future Work

This chapter shows progress towards correlating renal MRI measurements with histology. We are able to acquire high resolution  $T_1$ ,  $T_2$ ,  $T_2^*$  maps. We have also developed protocols to carry out simultaneous biopsy for histology and MRI acquisition. These protocols have shown that the  $T_1$  and  $T_2^*$  of the kidneys are not constant after fixation however there is a window of 24-hours after fixation in which scanning is optimum. We have also shown that these measures agree with histology. Below are listed some of the directions in which future work could explore.

### 1.4.1 Protocol Validation on a Single Sample

As each protocol, including those in Section 4.2, has been developed separately, they have not been carried out on the same sample, as the intention is to use all the protocols outlined in Section 1.2 on each nephrectomy sample, it would be useful to collect all protocols on a single sample. This could be coupled with a repeat investigation into the effects of ageing by collecting data from kidneys with a larger difference in ages.

### 1.4.2 Ex-Vivo Sample Coil

Sengupta demonstrated the benefits of using custom made ex-vivo sample coils in human scanners [4]. Currently scanning uses the standard head coils however this results in a relatively large distance between sample and coils as seen in Figure 1.19. There would certainly be improvements in data quality if a coil specifically designed for small sample imaging at 7T were fabricated.



Figure 1.19: A sample sat within the 82 channel 3T head coil. A bespoke ex-vivo sample coil would have less space between the coil and the sample.

#### 1.4.3 Human Organs

All work thus far has been using porcine kidneys. While these provide an excellent model for protocol development due to their similarities to human kidneys, the utility of this investigation will be enhanced massively when human organs are studied. To this end, once the development work has been completed and protocols finalised, samples will begin to be procured from subjects undergoing a nephrectomy as part of their standard clinical care.

Another source of human organs are those rejected for transplant. Due to the relatively small time window in which a transplant centre has between an organ donation being made and the organ losing its transplant viability, a large number of human kidneys are unable to be successfully donated. While not suitable for transplant any more, these organs would still be useful in providing ex-vivo MRI data and histology in the healthy population. There are pre-existing agreements enabling failed transplant tissue to be used in scientific research, as such, this would be an interesting avenue to explore.

## Chapter 2

# Applying $T_2$ Relaxation Under Spin Tagging (TRUST) To Assess Renal Oxygenation

### 2.1 Introduction

Sufferers of Chronic Kidney Disease (CKD) can have abnormalities in kidney structure or reduced urine function. More quantitatively CKD can be assessed clinically by Glomerular Filtration Rate (GFR), the rate at which fluid is filtered through the kidneys, with a value below 60 ml/min/1.73m<sup>2</sup> of body surface area being diagnostic or the presence of albumin, the main protein in blood plasma, in the patients urine [15–17]. An estimated 5–11% of the global population suffer from CKD [18–22] making it a significant public health concern. Late referral of renal disorders results in an increase in mortality rate and treatment costs [23–25]. Given that in 2013/2014 renal services cost the United Kingdom’s National Health Service £586 million [26] there are clear health and economic advantages to an early diagnosis of CKD.

The current methods available for CKD diagnosis are not ideal for a variety of reasons. Histological samples are the gold standard for diagnosis however collecting them is an invasive process and as such they are not suitable for monitoring the progress of a patient’s condition on a regular basis. This coupled with the fact that a small sample is not representative of the entirety of both kidneys means that this method has large drawbacks. Ultrasound can be used to gather structural information about the kidneys non-invasively, however, it suffers from low spatial resolution and the images being difficult to interpret.

The most common method of diagnosis is to estimate GFR from the creatinine content in a blood sample however this measure does not allow for the individual assessment of each kidney and is an indirect measure of kidney tissue damage.

MRI is a flexible non-invasive tool that can be used to collect a wealth of information about the kidneys. A current research interest at the SPMIC is to use multi-parametric renal MRI to assess and predict CKD. This protocol assesses haemodynamics, oxygenation and microstructure in a single 45 minute scanning session and shows significant changes in certain combinations of parameters in subjects with CKD as opposed to healthy volunteers [12]. Currently oxygenation is assessed using Blood Oxygen Level Dependent (BOLD)  $T_2^*$  maps to measure oxygenation of different tissues within the kidney, predominately the separation in mean  $T_2^*$  between the renal cortex and medulla, an example of which is shown in Figure 2.1. These BOLD  $T_2^*$  maps are, however, affected by other factors such as susceptibility effects, shimming and baseline blood flow and thus may be limited in their ability to draw quantitative conclusions despite their widespread use [17, 27].



Figure 2.1: An example  $T_2^*$  map. A clear difference can be seen between the renal medulla and cortex.

A welcome addition to this multi-parametric model would be the assessment of Renal Metabolic Rate of Oxygen ( $\text{RMRO}_2$ ); a measure analogous to the Cerebral Metabolic Rate of Oxygen ( $\text{CMRO}_2$ ) [28]. This measure can be calculated via Equation (2.1)

$$\text{RMRO}_2 = (Y_a - Y_v) \times \text{RBF} \times [\text{Hct}] \quad (2.1)$$

where  $Y_a$  and  $Y_v$  are arterial and venous oxygen saturation respectively, RBF is renal blood flow (in ml/min) and Hct is the ratio of the volume of erythrocytes to the volume of the rest of the blood, known as haematocrit. Renal Blood Flow (RBF) can be measured rel-

atively easily using Phase Contrast (PC)-MRI [29] and Hct is usually taken to be 0.41 for healthy adults but can be measured from a simple blood test [30, 31] or using the correlation between  $T_1$  of blood and its haematocrit [32]. This means that only a measurement of blood oxygen saturation via a non-invasive protocol is required to generate a quantitative value of RMRO<sub>2</sub>.

Blood oxygen saturation can be measured precisely via the insertion of catheters into the patient, however this is clearly an invasive process [33]. There are currently two well established methods of measuring blood oxygenation via MRI however these have only been used in the brain thus far. These methods are  $T_2$  Relaxation Under Spin Tagging (TRUST) [34–37] and susceptibility-based oximetry [38–41]. TRUST builds on the ideas of an Arterial Spin Labelling (ASL) sequence in the fact that by subtracting control images from labelled images only blood is imaged. However, instead of labelling a slab of tissue in the neck and imaging a superior slice, when implementing TRUST the imaging plane is inferior to the labelled slab. By collecting a series of pairs of labelled and control images with different  $T_2$  weightings it is possible to fit the data from the sagittal sinus to a  $T_2$  relaxation and use a calibration curve to convert the value of  $T_2$  to venous oxygenation [42]. Susceptibility-based oximetry is based upon the differences in magnetic susceptibility between the blood and the surrounding tissue. Using a phase map it is possible to model this difference in susceptibility and using the known difference in susceptibility between fully oxygenated blood and fully deoxygenated blood, venous oxygenation can be calculated.

Here both of the above techniques are applied to study oxygenation in the renal vein in young healthy individuals to assess the technicalities of transferring these protocols from the brain to the body. Given that these techniques have already been used in the brain with a number of studies in the literature, the sequences are first implemented on the brain to assess oxygenation in the superior sagittal sinus, then adapted to work within the more challenging environment of below the neck applications. These adapted sequences are compared to the results gained using the established techniques in the brain before testing on the renal vein. An oxygen challenge is carried out to verify that changes in oxygenation can be measured in the renal vein. If proved successful these sequences will be incorporated into the multi-parametric renal MRI protocol.

## 2.2 Methods

Imaging was performed on a whole body 3 Tesla MRI scanner (Ingenia, Philips Medical Systems, The Netherlands) using a 32 channel head or body coil. Studies were carried out according to the principles of the Declaration of Helsinki and approved by either the Local Ethics Committee or the East Midlands Research Ethics Committee. Written informed consent was obtained from all subjects.

### 2.2.1 Susceptibility-Based Oximetry

#### MRI Protocol

The principle behind susceptibility-based oximetry is based on the fact that there is a difference in magnetic susceptibility between the blood within a vessel and the tissue surrounding it [43]. As outlined by Jain, if a blood vessel is modelled as a long paramagnetic cylinder, it is possible to calculate the oxygenation of the blood by knowing the phase difference between blood in the vessel and the surrounding tissue, the angle of the vessel to the  $B_0$  field, the echo time of the scan and the subject's haematocrit [38]. This relationship is shown in Equation (2.2).

$$Y_v = \left[ 1 - \frac{2|\Delta\phi|}{\gamma TE \Delta\chi_{do} B_0 (\cos^2 \theta - 1/3) Hct} \right] \times 100 \quad (2.2)$$

where  $\Delta\phi$  is the average phase difference between the blood in the vessel and the surrounding tissue,  $\gamma$  is the gyromagnetic ratio of a proton, TE is the echo time,  $\Delta\chi_{do}$  is the susceptibility difference between fully deoxygenated and fully oxygenated blood ( $4\pi \times 0.27$  p.p.m) [44, 45],  $B_0$  is the static field strength,  $\theta$  is the angle of the vessel to the  $B_0$  field and Hct is the subjects haematocrit. Given haematocrit can be assumed or is measured with a blood test or by measuring the  $T_1$  of the blood, this means that from a simple phase map it is possible to calculate  $Y_v$ . The optimum phase map for this purpose was produced using a 2D  $T_1$  weighted FFE sequence with a flip angle of  $25^\circ$ , flow compensation, coil homogeneity correction and flyback. The FOV was  $230 \times 184 \times 29$  mm, matrix size of  $400 \times 300$ , TR of 12 ms, TE of 7.5 ms and three signal averages. This led to a total acquisition time of 9 seconds and as such could be completed in a single breath hold if required.

## Analysis

Once the phase map has been acquired, a ROI containing the superior sagittal sinus was defined. This mask was then dilated with concentric shells to generate the two ROI shown in Figure 2.2, note that the outer ROI has been constrained to within the brain during its dilation. There were no occurrences of phase wrapping in or immediately surrounding the superior sagittal sinus observed due to its small size and the high field homogeneity within the head and of the 3T scanner used. Any occurrences of phase wrapping could easily be corrected using Phase Region Expanding Labeller for Unwrapping Discrete Estimates (PRELUDE), a tool within fMRIB Software Library (FSL) (fMRIB, The University of Oxford) [46]. The average values of phase within these two ROI along with the angle of the vessel to the  $B_0$  field, as calculated from the localisation scans can then be used with Equation (2.2) to calculate  $Y_v$ .

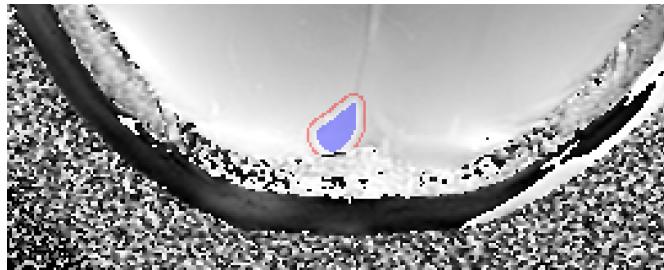


Figure 2.2: The region of interest averaged to find the intra-vascular phase (blue) and the region of interest used to find the phase of the surrounding tissue (red).

### 2.2.2 $T_2$ Relaxation Under Spin Tagging

#### MRI Protocol

The protocol for the TRUST MRI sequence in the brain involves the acquisition of a series of paired images using the pulse sequence shown in Figure 2.3a. A series of presaturation pulses using the Water suppression Enhanced through  $T_1$  effects (WET) scheme are applied to the imaging slice, shown in Figure 2.3b, to reduce the signal from static tissue and reduce contamination of the magnetisation in the imaging slice by an imperfect labelling slab profile [47, 48]. In the first of each image pair, a labelling pulse is applied consisting of two successive slice-selective  $90^\circ$  Radio Frequency (RF) pulses to generate a  $180^\circ$  label. The next image in the sequence has a control pulse applied to it instead of a labelling pulse, in this image the second of the  $90^\circ$  pulses is applied  $180^\circ$  out of phase to give zero net effect. As such any effects of magnetisation transfer related signal in the sta-

tionary tissue can be cancelled out because the net RF effect on the macromolecular spin magnetization is identical for both the labelling pulse and control pulse. This method of labelling is known as Transfer Insensitive Labelling Technique (TILT) and is widely used in literature for labelling in TRUST in the brain [49]. A series of non-selective  $T_2$  preparation pulses are then applied to minimise the blood outflow effect and modulate the  $T_2$  weighting of the image, the time between the application of the labelling pulse and the  $T_2$  preparation is known as the Post Label Delay (PLD). Finally a 90° excitation pulse is applied followed by a standard EPI readout at time TE later [35]. If the control image is subtracted from the labelled image then only the venous blood that flowed from the labelled slab to the imaging slice will be visible, as shown in Figure 2.4.

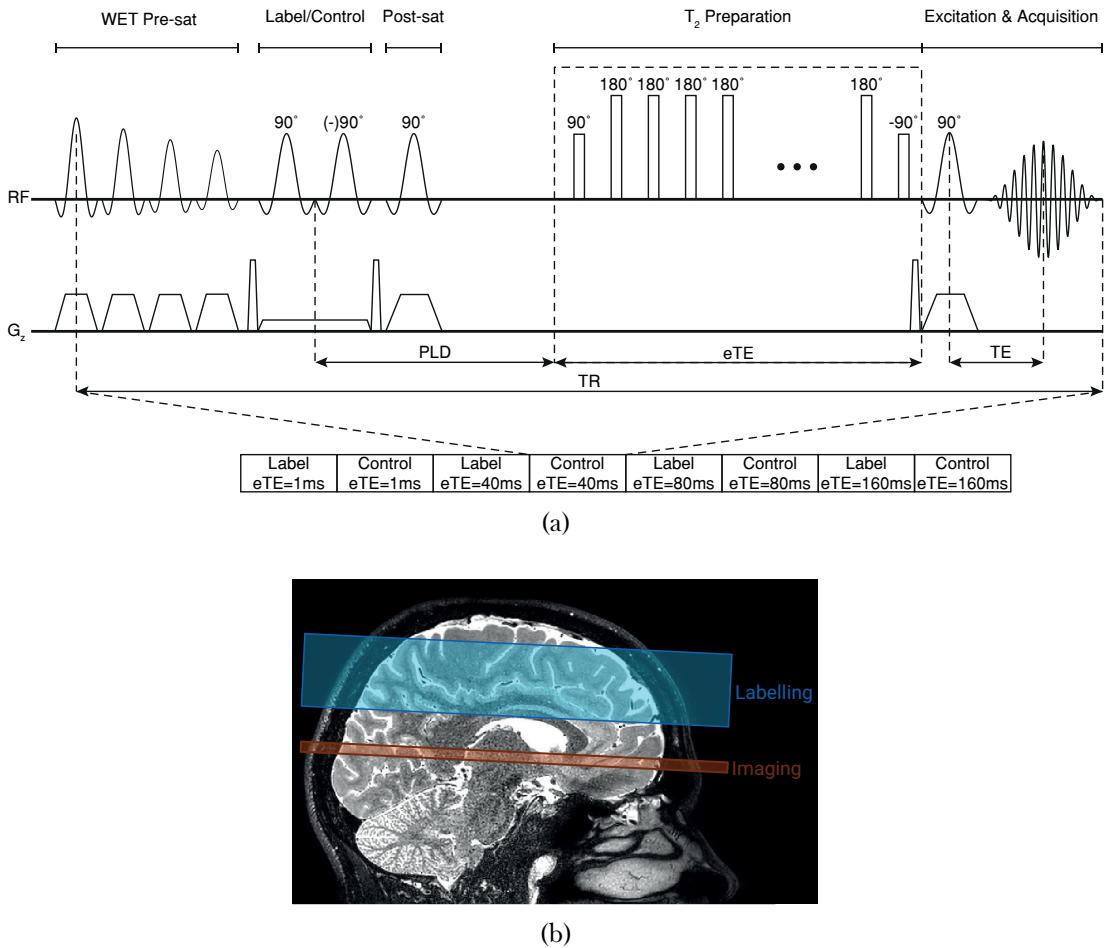


Figure 2.3: (a) The pulse sequence for TRUST MRI using the TILT labelling sequence. (b) The labelling and imaging volumes used for TILT tagging within the brain.

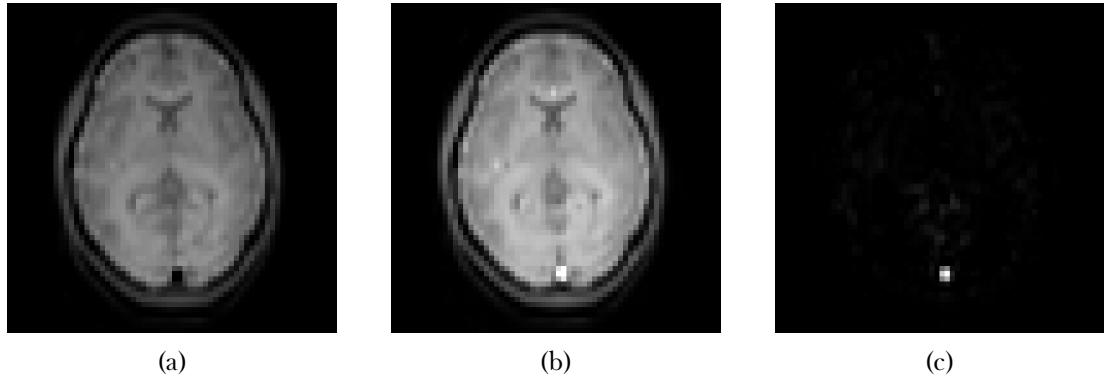


Figure 2.4: The control image, (a), is subtracted from the labelled image, (b), to generate a difference image, (c), of only the tagged blood.

This process is then repeated for another pair of images, however, this time the duration of the  $T_2$  preparation is increased to a larger eTE, this applies a  $T_2$  weighting to the image in addition to the constant weighting caused by the regular TE. Three label/control image pairs were acquired with each eTE of 1 ms, 40 ms, 80 ms and 160 ms.

The resulting signal in the superior sagittal sinus of the difference between the labelled image and control image,  $\Delta S$ , is defined by Equation (2.8)

$$\begin{aligned}
 \Delta S &= S_{\text{label}} - S_{\text{control}} \\
 &= S_{\text{blood label}} - S_{\text{blood control}} \\
 &= S_0 e^{eTE(1/T_1 - 1/T_2)} \tag{2.8}
 \end{aligned}$$

where  $S_0 = 2e^{-T_1/T_1 - T_E/T_2^*}$  and;  $T_1$ ,  $T_2$  and  $T_2^*$  are the relaxation constants of blood. If it is assumed that  $T_1$  of blood is approximately 1624 ms [50] then it is possible to fit the collected data to a mono-exponential function and find an estimate of  $T_2$ . It is deemed acceptable to use a mean value of  $T_1$  as it will always be much greater than the value of  $T_2$  and thus the possible small changes in  $T_1$  due to blood oxygenation and haematocrit become negligible when fitting the  $T_2$  curve.

The final step in this procedure is to convert the value of  $T_2$  into one of  $Y_v$ . The relationship between  $T_2$  and  $Y_v$  is relatively well known and as such a simple empirically derived calibration curve can be used for this conversion, Figure 2.5 [31, 51, 52].

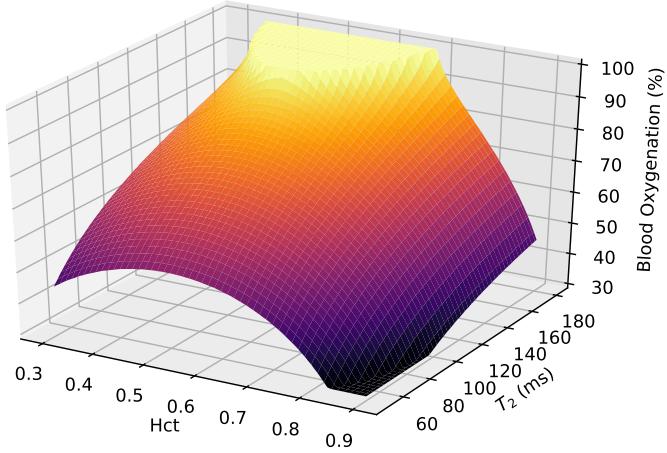


Figure 2.5: The calibration curve used to convert between  $T_2$  and  $Y_v$  for a given haematocrit [58].

The parameters used in the brain TILT TRUST sequence were as follows: label slab thickness = 100 mm, imaging slice thickness = 5 mm, distance between centre of imaging slice and centre of labelling slice = 75 mm, FOV =  $220 \times 220 \times 5$  mm, matrix size =  $64 \times 64$ , voxel size =  $3.44 \times 3.44$  mm, SENSE = 3, EPI factor = 15,  $T_1$  = 1624 ms, PLD = 1022 ms, the choice of this value will be explored later, TR = 3000 ms, TE = 2.9 ms, eTE = 1 ms, 40 ms, 80 ms and 160 ms with three pairs of images acquired at each. This led to a total scan duration of approximately 84 seconds.

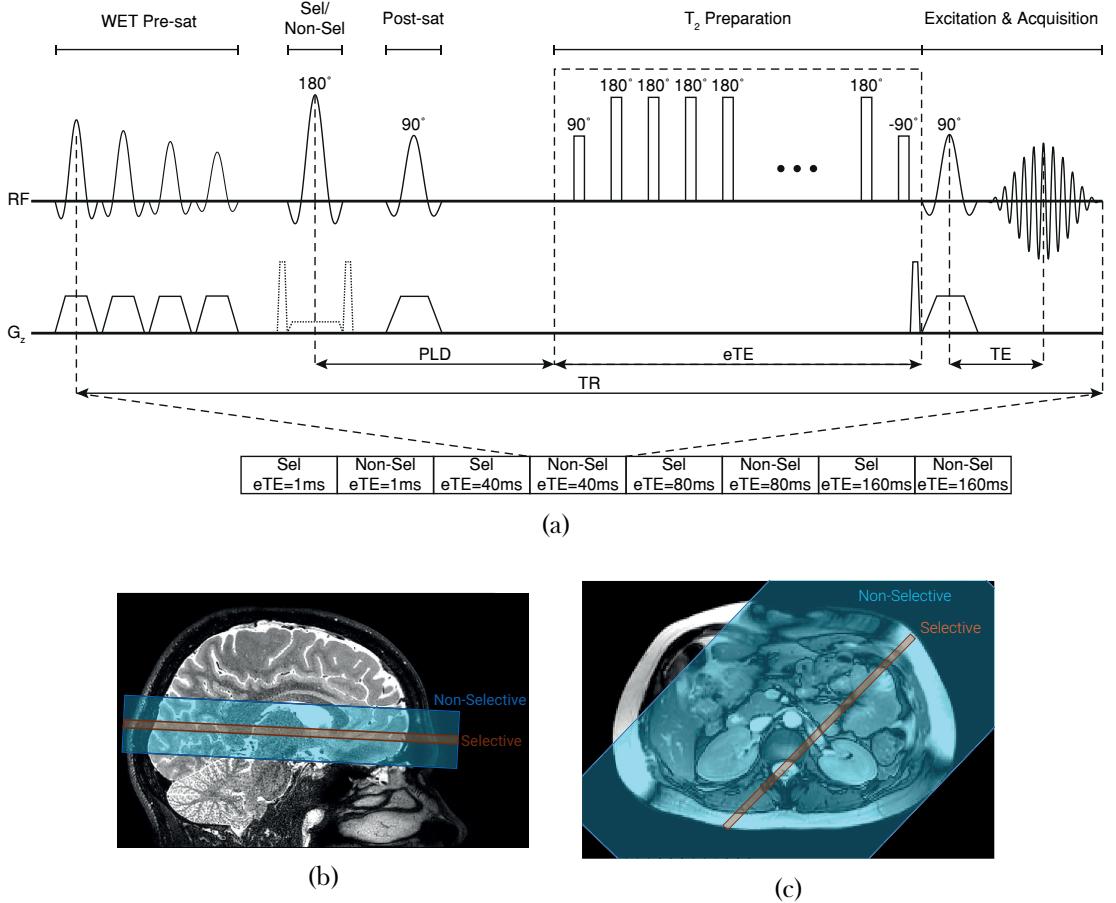


Figure 2.6: (a) The pulse sequence for TRUST MRI using the FAIR labelling sequence. (b) The selective and non-selective volumes used for tagging via FAIR in the brain. (c) The selective and non-selective volumes used for tagging via FAIR in the kidneys.

The main hurdle to be overcome when moving TRUST to the body is the inhomogeneity in the magnetic field caused by the far less homogeneous tissue susceptibilities within the body compared to the brain. These inhomogeneities mean that it is not possible to use TILT as the labelling method, instead the Flow-sensitive Alternating Inversion Recovery (FAIR) labelling scheme will be used [54], a diagram of this pulse sequence is shown in Figure 2.6a. In the FAIR labelling scheme a selective inversion pulse is applied with slice selective gradients turned on followed by  $T_2$  preparation and acquisition to generate the first image in the pair, a non-selective inversion pulse is then applied with a lower slice selective gradient followed by  $T_2$  preparation and then acquisition to generate the second image. An example of the raw images produced is shown in Figure 2.7. A schematic of the selective and non-selective slices in the brain and the renal vein are shown in Figures 2.6b and 2.6c respectively. This sequence also has the advantage of being far easier to plan, in the brain having a separate labelling and imaging slice is relatively trivial however the flow of blood in the body is far less ordered and as such, the use of a selective slab within a non-selective

slab yields far better results. Movement is a much greater problem in the body. Given the long acquisition time of TRUST it is impossible to carry out the scan in a breath hold, as such the sequence is respiratory triggered via a respiratory belt applied around the subjects chest. The total scan time is therefore dependent upon respiratory rate. Depending on the subject, a delay can be applied between the respiratory trigger and the labelling pulse to acquire images while the subject has fully exhaled.

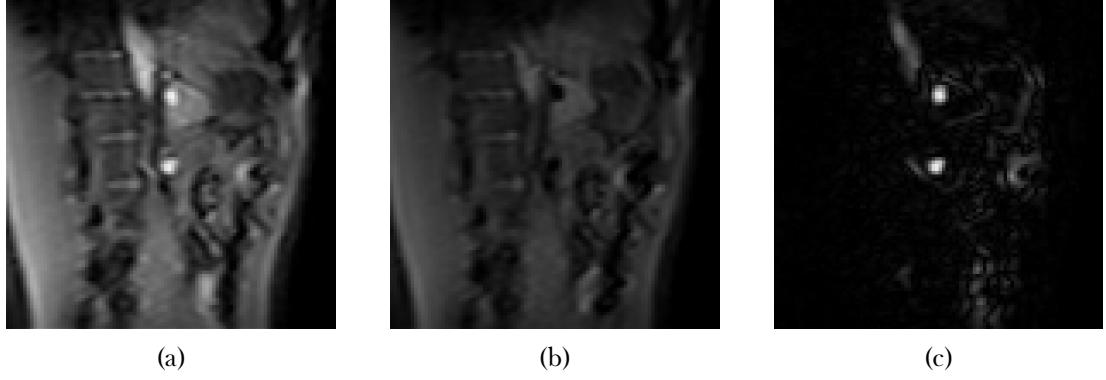


Figure 2.7: The raw images generated when using the FAIR labelling sequence on the kidneys. The non-selective image, (a), is subtracted from the selective image, (b), and generates (c), an image of only the untagged blood. The raw FAIR images from the brain are omitted as they are very similar to those seen in Figure 2.4.

When using the FAIR labelling sequence on the brain the following parameters were used: selective slab thickness = 25 mm, non-selective slab thickness = 400 mm, FOV =  $220 \times 220 \times 5$  mm, matrix size =  $64 \times 64$ , voxel size =  $3.44 \times 3.44 \times 5$  mm, SENSE = 3, EPI factor = 15,  $T_1$  = 1624 ms, PLD = 800 ms, TR = 7276 ms, TE = 2.9 ms, eTE = 1 ms, 40 ms, 80 ms and 160 ms with three pairs of images acquired at each. When used on the body, the parameters were as follows: selective slab thickness = 25 mm, non-selective slab thickness = 400 mm, FOV =  $244 \times 244 \times 5$  mm, matrix size =  $96 \times 96$ , voxel size =  $3.44 \times 3.44 \times 5$  mm, SENSE = 8, EPI factor = 15,  $T_1$  = 1624 ms, PLD = 1000 ms, the choice of this value will be explored later, TR = 8076 ms, TE = 2.9 ms, eTE = 1 ms, 40 ms, 80 ms and 160 ms with three pairs of images acquired at each.

## Analysis

The analysis of the data collected using the above protocol was carried out using custom MATLAB (MathWorks, Natick, MA) software based upon code written by Liu and modified to work with data collected using the FAIR labelling method by Cox [55]. This

software loads the data and carries out the subtraction of each image pair then presents a difference image to the user so the vessel can be drawn around. At this point the voxels with the greatest intensity within the vessel, four voxels when calculating  $Y_v$  for the superior sagittal sinus and nine voxels when working on the renal vein, are averaged, as are the intensities of each repeat eTE. These mean signals are then fit to Equation (2.3) to compute a value of  $T_2$  with confidence bounds. The value of  $Y_v$  can then be found using the aforementioned calibration curve. Once the software has finished, it saves all outputs and intermediary variables to a file on the computer for later analysis.

### 2.2.3 Inducing Changes in Oxygenation of Blood in the Renal Vein

In order to assess the ability of these methods to measure a change in renal oxygenation, a method of inducing such a change in the kidneys needed to be devised. Looking at literature that has carried out similar studies, it is suggested that changes in renal oxygenation can be induced by either varying the subjects sodium intake, water intake or inspired oxygen level [56, 57].

Due to the challenges associated with controlling subjects diet for two weeks as was performed in Priijm [58], the use of sodium intake was discounted. From previous work we know that applying a large water load to subjects during the scanning session, as in Tumkur and Prasad [59, 60], can cause undesired effects on the resultant shim as assessed by  $B_0$  maps due to the large susceptibility change adding such a large quantity of water to the abdomen can cause, as such, this method was also discounted leaving us to pursue an oxygen challenge.

This method consisted of localisers and anatomical images being collected followed by alternating BOLD  $T_2^*$  and TRUST scans while the subject was breathing room air to record a baseline. Pure oxygen was then delivered to the subject at 15  $\ell/\text{min}$  via a gas mask and, after a two minute wash in period, the BOLD  $T_2^*$  and TRUST scans were repeated. A visual representation of this protocol can be seen in Figure 2.8. The BOLD  $T_2^*$  scans had a slice thickness of 5 mm, 12 echoes with an initial TE of 5 ms and subsequent echo spacing of 3 ms, the flip angle was 30°. The total scan time was approximately 17 seconds and was acquired during a single breath hold. The TRUST scans were conducted as per Section 2.2.2.

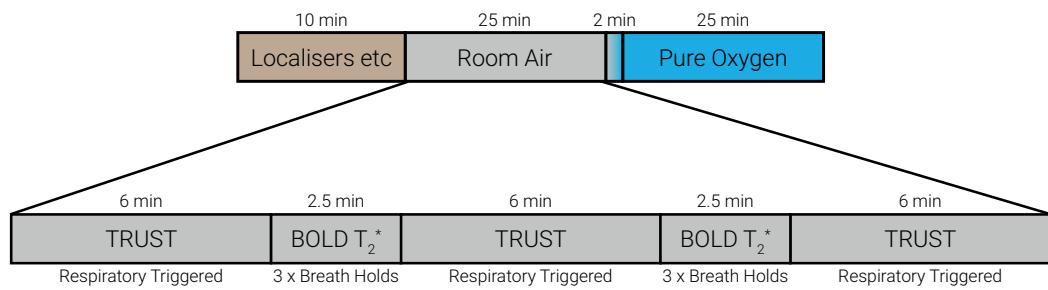


Figure 2.8: The protocol used to induce changes in renal oxygenation.

## 2.3 Results and Discussion

### 2.3.1 Susceptibility-Based Oximetry

#### Susceptibility-Based Oximetry in the Brain

Having collected data using the method outlined in 2.2.1 it was possible to use Equation (2.2) to estimate  $Y_v$  in the superior sagittal sinus to be  $63 \pm 2.1\%$ . This is consistent with the value reported by Liu of  $61.1 \pm 1.4\%$  found in a multi centre TRUST trial with 250 participants over a wide range of ages and ethnicity distribution [37].

#### Susceptibility-Based Oximetry in the Renal Vein

Having calculated an acceptable result in the brain that agreed with literature it was possible to move onto applying techniques to assess oxygenation in the renal vein. A set of three phase maps were collected along with three localisers, one along each plane. If  $\Delta\phi$  is plotted against  $\theta$  for a typical  $Y_v$  of 85%, Figure 2.9 is produced. It can be seen that, for an expected  $Y_v$ , the phase difference is greatest if the vessel runs parallel to the  $B_0$  field. No part of the renal vein is located parallel to the  $B_0$  field, typically the angle is in the region of  $75^\circ$  (there is a large degree of variability in vasculature geometry between subjects) and as such delivers a very small phase difference. This coupled with the fact that the gradient of this function at these angles is large, meaning that the uncertainty in angle corresponds to a larger uncertainty in  $Y_v$  means it will unfortunately not be possible to use susceptibility-based oximetry to accurately measure  $Y_v$  within the renal vein.

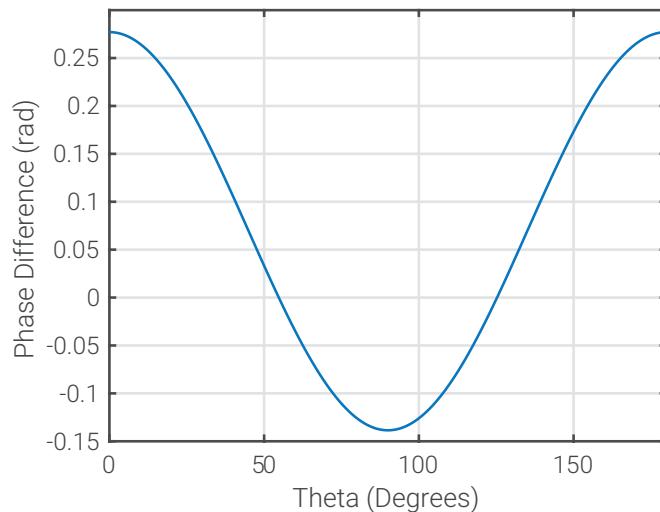


Figure 2.9: For a typical  $Y_v$  of 85% the phase difference produced by a vessel at a range of angles to  $B_0$ .

This technique would perhaps be better suited to use in the liver to assess oxygenation in the portal vein. This vessel runs at a much smaller angle to the  $B_0$  field and as such the model will still be valid with reasonable errors, Figure 2.10. This would potentially work much better than TRUST here as the sequence is much quicker and therefore will be less susceptible to movement.

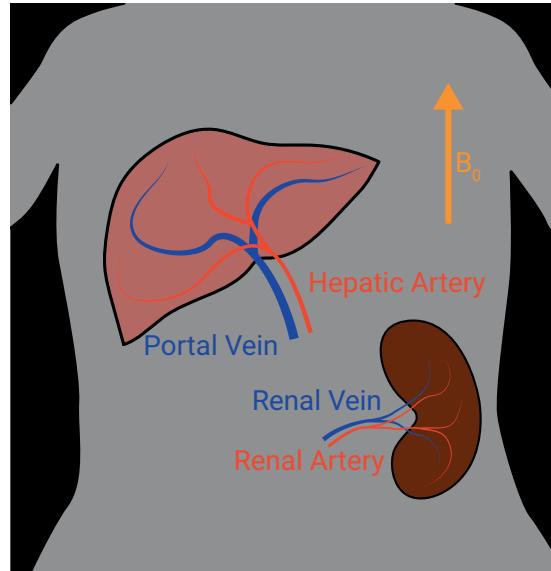


Figure 2.10: A schematic of the portal and renal veins entering the liver and left kidney respectively in relation to the  $B_0$  field.

### 2.3.2 $T_2$ Relaxation Under Spin Tagging

#### TRUST in the Brain

To test if the FAIR labelling sequence delivered the same signal decay as the TILT sequence both labelling schemes were performed sequentially on the superior sagittal sinus with a PLD of 800 ms. The resulting normalised signals are shown in Figure 2.11.

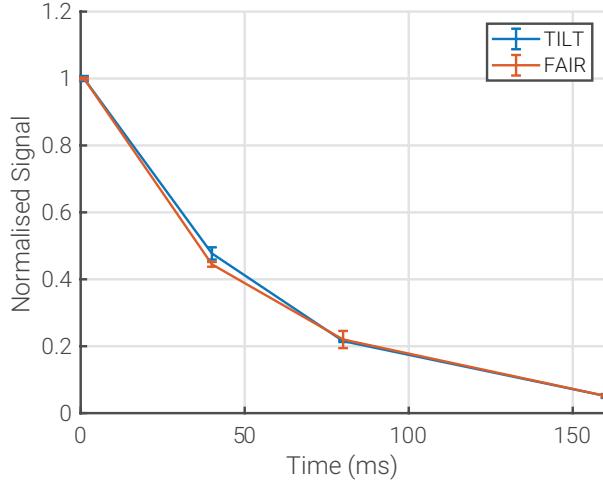


Figure 2.11: The signal decay within the superior sagittal sinus found using TRUST with both TILT and FAIR labelling sequences scaled by their initial signal intensities at eTE = 1 ms.

As can be seen these signals are in excellent agreement with the TILT sequence producing a  $T_2$  of  $52 \pm 1$  ms and the FAIR sequence producing a  $T_2$  of  $50 \pm 2$  ms, therefore in agreement within the bounds of error. This means that FAIR can be directly substituted for TILT in the TRUST sequence to measure  $Y_v$  in the superior sagittal sinus and can subsequently be used for the renal TRUST measurements.

To find the dependence PLD has upon the signal measured, scans were carried out at a range of delays from 400 ms to 1400 ms while using the FAIR labelling sequence. The signal from eTE=1 ms was then plot against label delay. Figure 2.12 shows the signal from the difference images. The maximum signal is observed with a PLD of 800 ms. This value is reached due to the balance between  $T_1$  relaxation of the non-selective blood and inflow of unlabelled blood. This maximum in signal agrees with literature using the TILT labelling scheme [34]. By carrying out scans with this PLD the maximum Signal to Noise Ratio (SNR) will be achieved.

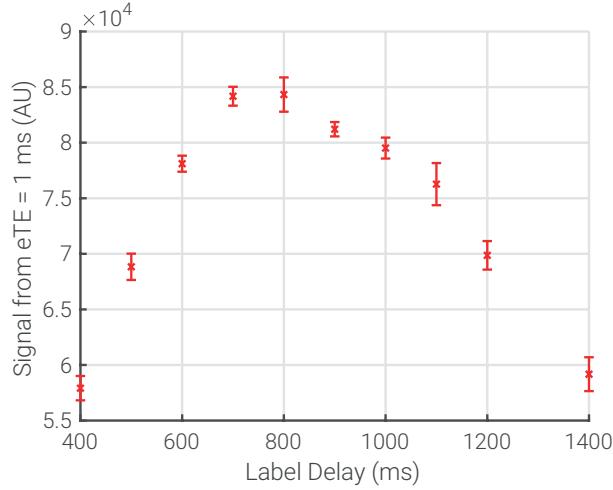


Figure 2.12: The mean signal from the first echo of each difference image over a range of PLD times.

$T_2$  should have no dependence upon PLD given the signal from the difference image will have the same decay in time, it will just be a lower intensity for non-optimal PLD thus leading to a larger confidence interval. To confirm this the fit values of  $T_2$  were plotted against PLD, Figure 2.13.

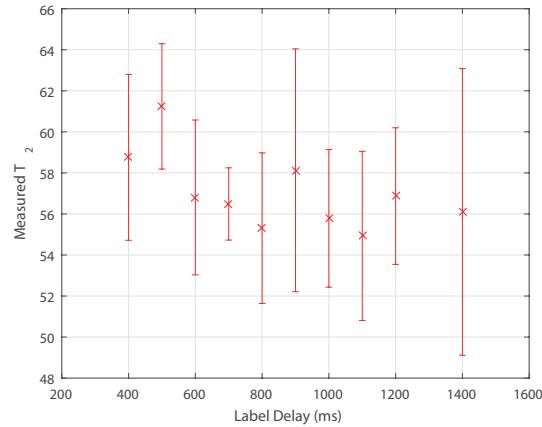


Figure 2.13: The dependence of  $T_2$  upon PLD.

It can be seen that, as predicted, there is no relationship between  $T_2$  and PLD. An increase in error with label delay was not observed, this effect may only show itself at larger values of PLD however for our purposes, simply confirming there is no large increase in error around our chosen PLD is sufficient. This means that if there is a variation in the optimum PLD between subjects due to the larger range in RBF compared to Cerebral Blood Flow (CBF) then this will not have an affect upon the value of  $T_2$  and thus  $Y_v$ .

When the analysis is carried out on the images, the four brightest voxels of the differ-

ence image are averaged before the fitting occurs. This number of voxels is chosen due to the average size of the superior sagittal sinus however, for some subjects more voxels could be included, potentially yielding better results. To assess the variability in  $T_2$  measurements with the number of voxels averaged, the analysis was run multiple times with one to twelve voxels included in the calculation. Multiple TRUST scans were performed on the same subject and averaged generating Figure 2.14a.

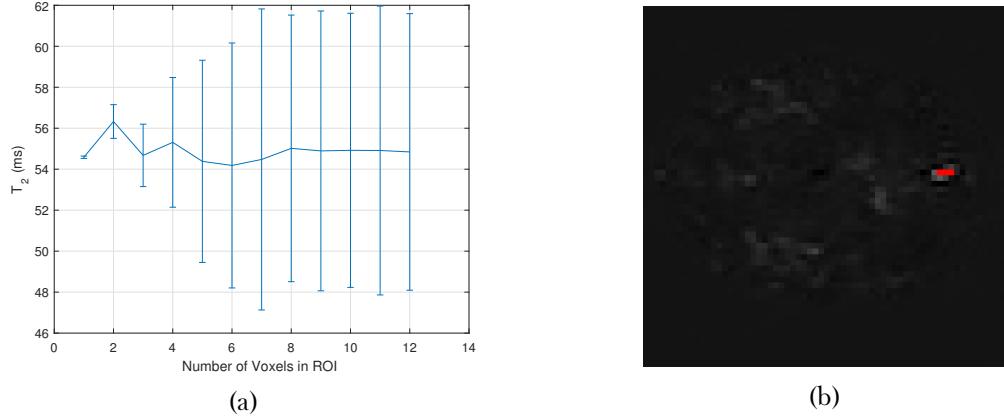


Figure 2.14: (a) The value of  $T_2$  computed for the superior sagittal sinus with different numbers of voxels included in the calculation. (b) The difference image of the superior sagittal sinus with a three voxel ROI shown. This is already covering most of the vessel, hence the noise going up as more voxels are added to the calculation.

Although from Figure 2.14a it would appear that it would be best to only use the brightest voxel in the calculation due to its very small error and that it has the same value of  $T_2$  as the results with far more voxels; this would not be a very robust method. It is fairly easy to conceive a greater than average level of noise being recorded on a single pixel in the relaxation and as such skewing the output of the calculation. The confidence interval is so large above six voxels because by this point the calculations are simply including the noise around the vessel rather than the signal from the blood within the sagittal sinus. Given these results, using four voxels in the calculation seems to be a reasonable balance between uncertainty and robustness.

To assess the repeatability of this measure, the optimised scan was repeated ten times on a single subject during one scanning session. This yielded a  $Y_v$  of  $69.5 \pm 0.6\%$ , a value consistent with literature [33, 37]. Given the success of the modified sequence on the superior sagittal sinus, it was possible to attempt to measure  $Y_v$  in the renal vein.

## TRUST in the Body

Ideal vessels to test the TRUST sequence within the body are the portal vein and hepatic artery as these vessels are large, have different oxygen saturations and can easily be imaged at the same time. Using the modified TRUST sequence the  $T_2$  and oxygen saturation of the portal vein was found to be  $109 \pm 5$  ms and  $79.9 \pm 0.8$  % respectively; the  $T_2$  and oxygen saturation of the hepatic artery was found to be  $157 \pm 10$  ms and  $100 \pm 1$  % respectively. This means that, as expected, the oxygen saturation in the hepatic artery is measured as greater than that of the portal vein and therefore the TRUST protocol is working as expected. Although normally the analysis would simply be based upon the mean of the brightest voxels in the difference image as outlined in Section 2.2.2, in Figure 2.15 a voxel by voxel analysis has been carried out for illustrative purposes.

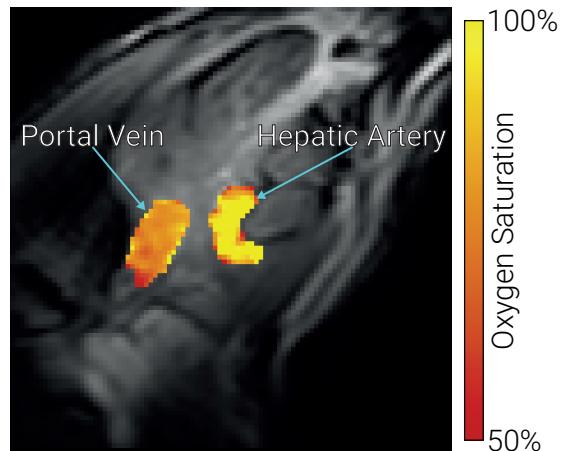


Figure 2.15: The oxygen saturation of the portal vein and hepatic artery measured using TRUST.

To assess if the PLD that generates the greatest signal is the same in the renal vein as in the superior sagittal sinus, a series of scans were collected with PLD ranging from 400 ms to 1400 ms and the signal from  $eTE = 1$  ms recorded.

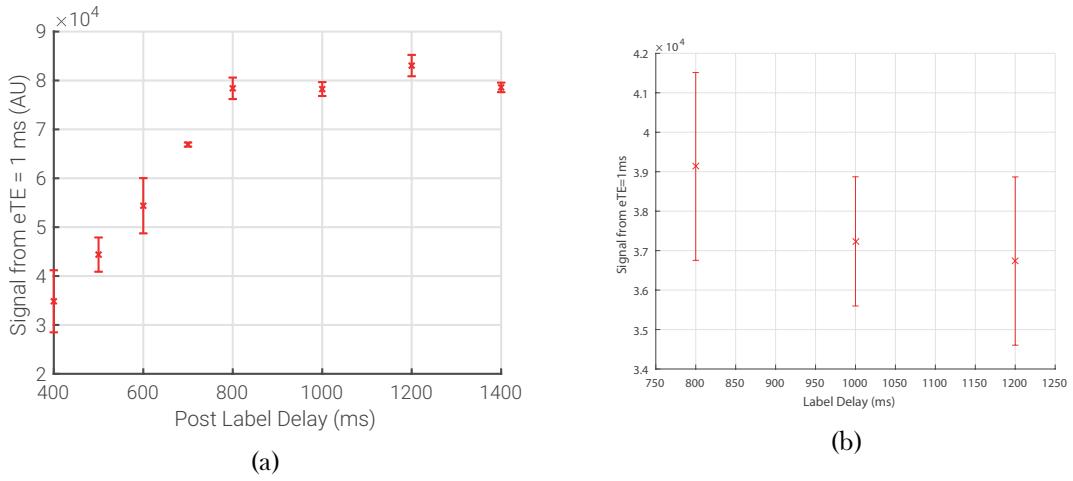


Figure 2.16: (a) The mean signal from the first echo of each difference image of the renal vein over a range of PLD. (b) Mean signal from the first echo versus PLD from a different subject.

As seen in Figure 2.16a the PLD producing the largest signal in the difference image of the renal vein is indeed different to that of the superior sagittal sinus. This is most likely due to differences in blood flow through each of these vessels,  $413 \pm 136$  ml/min in the renal vein [12] and  $285 \pm 19$  ml/min in the superior sagittal sinus [29]. Given the much larger uncertainty in blood flow in the renal vein, a different subject was scanned over a smaller range of PLD to ascertain if the PLD delivering the maximum signal varies much between subjects, Figure 2.16b.

The maximum signal for the first subject was achieved at a PLD of 1200 ms whereas for the second subject the maximum is at a PLD of 800 ms. Given that these subjects had a RBF either side of the mean and that there is little dependence of  $T_2$  upon PLD it seems appropriate to use a PLD of 1000 ms for optimum signal in most subjects.

Given the larger size of the renal vein compared to the superior sagittal sinus, it would be better to include more voxels in the calculations when fitting to find a value of  $T_2$ . Multiple scans were completed on a single subject and the value of  $T_2$  found for each using one to twelve voxels in the fitting process. The results were averaged and plot in Figure 2.17a.

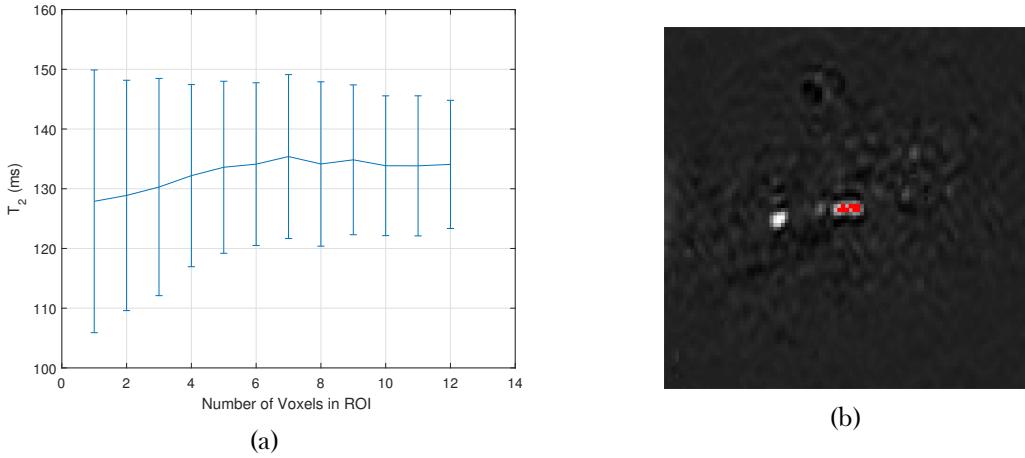


Figure 2.17: (a) The value of  $T_2$  calculated for the renal vein with different numbers of voxels included in the calculation. (b) The difference image of the renal vein with a nine voxel ROI shown.

Unlike the results when this process was carried out on the superior sagittal sinus in Figure 2.14a, here the error decreases as more voxels are added to the calculation. This uncertainty comes from the large variation in  $T_2$  for one voxel rather than a large error on the fit i.e. the error is coming from the differences between scans rather than the robustness of each scans results, this is precisely the concern that was raised with using a single voxel when discussing the superior sagittal sinus. As more voxels are added the error decreases until approximately six voxels are included, at this point the value of  $T_2$  stops increasing and stays approximately constant. Once again, given the large variation in renal veins, it would be advisable to include slightly more than six voxels but not so many that in the cases of small vessels the algorithm is sampling surrounding tissue. Nine voxels seems to be a suitable middle ground as to work effectively with both small and large vessels.

To assess the repeatability of the measurements within the kidney, the same scan was repeated ten times in a single session with the optimised renal parameters. This yielded a  $T_2$  of  $135 \pm 5$  ms corresponding to a  $Y_v$  of  $89 \pm 2\%$ . The variation in measurements of  $Y_v$  in the renal vein are relatively substantial and show no dependence upon time so are therefore not likely due to physiological changes. The value of  $Y_v$  in the renal vein is much higher than in the sagittal sinus however is within the range found by Nielsen [61].

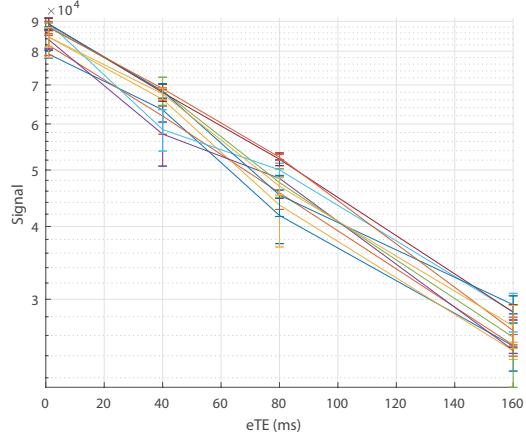


Figure 2.18: The  $T_2$  relaxation curves of ten scans repeated on a single subject.

To compare the abilities of BOLD  $T_2^*$  maps and TRUST to measure changes in oxygenation in the kidneys, a hyperoxia challenge was conducted. In Figure 2.19a, no systematic, bulk change in  $T_2^*$  can be seen indicating that the change in  $T_2^*$  caused by the introduction of pure oxygen is dominated by other confounding factors. This is confirmed when ROI are defined for the renal cortex and renal medulla with the mean change in  $T_2^*$  found to be  $-2 \pm 8$  ms and  $-1 \pm 6$  ms respectively. When TRUST is used to measure the oxygen saturation in the renal vein an increase of  $16 \pm 3$  % is observed, Figure 2.19b. This shows that it is possible to measure changes in renal oxygenation using TRUST that would be undetectable using the current standard, BOLD  $T_2^*$  mapping.

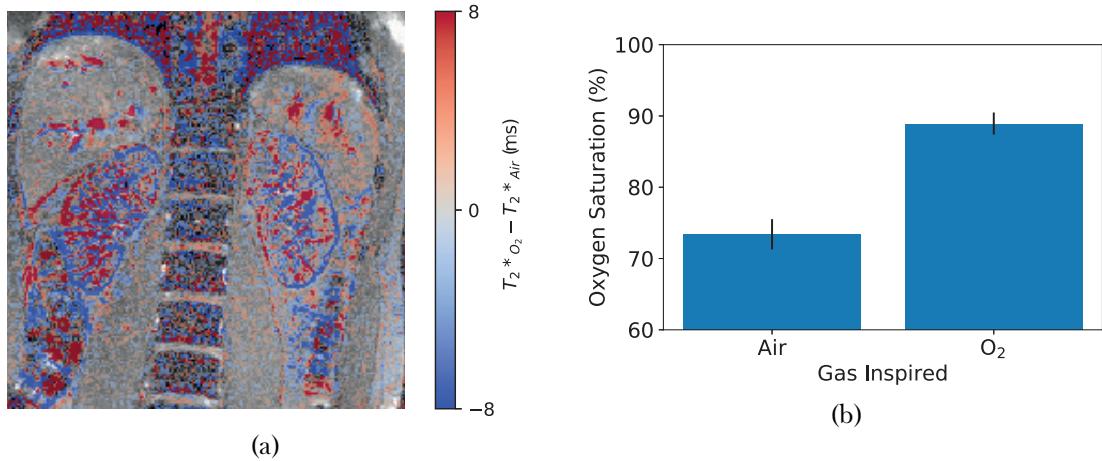


Figure 2.19: (a), The difference in  $T_2^*$  measured between baseline and the hyperoxia state. (b) The difference in  $Y_v$  measured using TRUST.

## 2.4 Conclusions and Future Work

This work shows promising results for the use of a modified TRUST sequence to measure oxygenation of blood within the body. The existing TRUST sequence was modified to be respiratory triggered and use the FAIR labelling scheme making it suitable for use in the body. Once these modifications had been carried out, parameters such as PLD and the number of voxels used in the ROI were optimised. The ability of TRUST to measure a change in renal oxygenation was successfully verified via a hyperoxia challenge which was able to measure an increase of  $16 \pm 3\%$  where the current standard measurement of renal oxygenation, BOLD  $T_2^*$  maps, recorded no significant change.

Looking forward this work could be expanded by carrying out the hyperoxia challenge on more subjects. Although a small number of measurements were gathered on the hepatic vessels, further work could be undertaken to compare the use of susceptibility based oximetry and TRUST to measure oxygenation in the portal vein in response to a hyperoxia challenge as conducted for the kidneys here. In the current protocol, haematocrit is assumed to be an average value of 0.41 unless a blood test has recently been undertaken. As stated above, there is a correlation between  $T_1$  of blood and its haematocrit, this means that a measurement of the subjects haematocrit could be taken while they are in the scanner, thus leading to a more accurate measurement of oxygenation with only a small increase in scan time.

## Chapter 3

# Automated Segmentation of Kidneys using Machine Learning

### 3.1 Introduction

Segmentation of the kidneys in MRI images is a vital, yet time consuming, aspect of many renal studies [12, 62, 63]. Total Kidney Volume (TKV) is used as a biomarker for a variety of renal pathologies; autosomal dominant polycystic disease is characterised by an increase in TKV [64, 65], while a decrease in TKV is associated with a decrease in renal function [66]. In addition to TKV measurements, renal segmentation is an important first step for many other processing pipelines, be that to increase accuracy of automated cortical-medullary segmentations or reduce computation times by only carrying out calculations for a relevant ROI. The gold standard of segmentation is manual ROI definition by an experienced and skilled professional, this manual process is highly time consuming and difficult due to similar signal intensities between the kidneys and other organs, anatomical differences between subjects and imaging artefacts. These factors mean that developing a fully automated method of renal segmentation is highly desirable. Such methods have been proposed with varied successes [67, 68] however the techniques used differ between diseases. The fact that each technique is highly optimised for a specific dataset means that it needs to be re-written to be applied to different a pathology, another time consuming and highly skilled process.

Machine learning allows for a single method to be written and then trained on different datasets. This means that as more data becomes available the algorithm can become more accurate and generalised, without a need to rewrite the methods, thus making it a

better choice for long term development. Such methods have already been applied to segmentation in other areas of medical imaging, especially successful have been U-Nets. An example of a Fully Convolutional Network (FCN), these algorithms convolve the image with a series of filters to extract features from the input data and thus generate a voxel by voxel classification. The weights of each pixel in these convolution kernels is honed through an optimisation process where the filters are applied to the training data and their performance is evaluated against the manually segmented data using a loss function such as mean square error or dice overlap score. Changes to each filter are then back propagated through the network and the process starts again. After many iterations, the filters become tuned to detecting the feature labelled in the manually segmented data.

To avoid the network becoming too specific and, for example, just learning to detect the specific kidneys in the data the network has been trained on rather than all kidneys, the data is divided into three categories, training, testing and validation. The training data is used for adjusting the filter weights over a short time scale, usually after a few tens of images. Once all the training data has been processed by the network and filter weights adjusted, all the test data is run through the network without any further weight adjustments and the performance evaluated using the loss function, this train, test cycle is known as an epoch. If the network has become too specialised and finely adjusted to the training data then it will not perform well on the test data. To stop this over specialisation, or over-fitting, the weights used at the start of each epoch are those that delivered the best performance on the test data. The validation data is never used to influence the filter weights and is used to validate the performance of the network on unseen data.

Similar methods have been applied to segment other areas of anatomy [69, 70], however this has not been successfully applied to segment the kidneys. Here we propose a FCN to segment the kidneys from a standardised MRI protocol.

## 3.2 Methods

### 3.2.1 Data Acquisition

All data is acquired on a 3T Philips Ingenia system using a  $T_2$ -weighted Half-Fourier Single-shot Turbo spin Echo (HASTE) sequence (TE = 60, TR = 1800 or 1300 ms, FOV =  $350 \times 350$  mm $^2$ , voxel size =  $1.46 \times 1.46 \times 5$  mm $^3$  with enough slices to cover the entire kidney, usually 12-14, SENSE = 2.5), the sequence is carried out in a single breath hold. Parameters have been optimised to deliver the maximum contrast between the kidneys and surrounding tissue. Training and test data is a single volume per subject; validation datasets are composed of five volumes acquired on the same subject in the same scanning session with the subject being removed from the scanner, asked to move, then positioned back in the scanner between each acquisition. The scanner operator attempted to vary the acquisition geometry between repeats while still acquiring the full kidney volume. These validation datasets allow the consistency of the networks ability to measure TKV to be assessed. Manual binary masks are generated for every volume to allow the network to train and its accuracy to be investigated. A summary of the data collected can be seen in Table 3.1, to make the algorithm as generalisable as possible, healthy volunteers and patients with CKD are scanned.

Dataset	Number of Subjects	Number of Volumes
Healthy kidneys	26	26
CKD kidneys	23	23
Validation healthy kidneys	5	25
Validation CKD kidneys	3	15

Table 3.1: Number of subjects and volumes in each dataset type.

The accuracy of the network will increase as it is trained on a greater volume of data. As such, this protocol is still being run on as many subjects as possible to further increase the accuracy of the algorithm.

### 3.2.2 Data Pre-Processing

All training and test data is loaded and the order of the volumes randomised i.e. healthy volunteers and patients are mixed. Each slice is resampled to a matrix size of  $256 \times 256$  and voxel intensities scaled between 0 and 1 where black is set to the mean voxel value minus 0.5 times the standard deviation of that slice and white is set to the mean voxel value plus 4 times the standard deviation of the slice; values outside this range are clipped to 0 or

1. This windowing leads to a clear contrast between kidneys and surrounding tissue while negating the effects of bulk signal changes between subjects.

This resampling and windowing is also applied to the validation data before it is processed by the network. Once a prediction of the renal mask has been generated by the algorithm, the mask is resampled back to the original image volume dimensions.

The ratio of training data to test data is eighty to twenty. No data augmentation is performed in this architecture.

### 3.2.3 Network Architecture

A summary of the network architecture is shown in Figure 3.1. Each volume is split into two-dimensional slices before being processed by the network. Convolution and deconvolution layers use a  $3 \times 3$  kernel. Activation layers use a Rectified Linear Unit (ReLU). Max pooling with a stride 2 is used on the encoding half of the network.

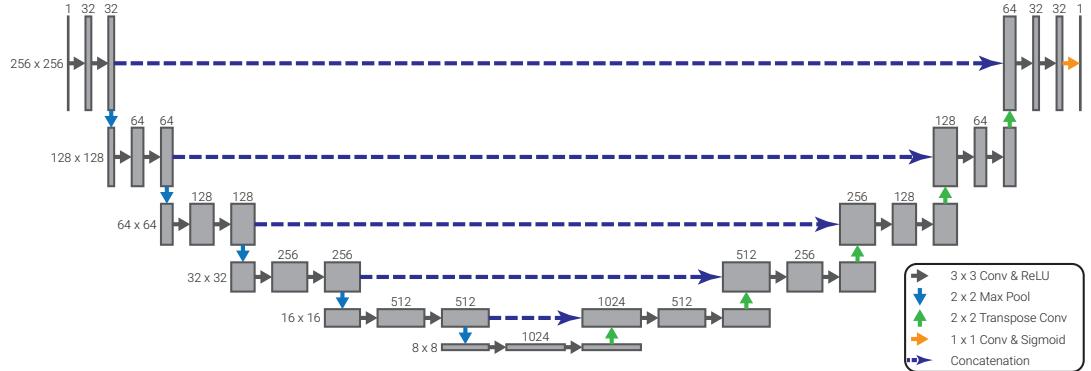


Figure 3.1: The architecture of the network used.

The network uses a dice score, defined by Equation (3.1), as its loss function; this function is ideal for renal segmentation as it doesn't weight true negatives which represent the majority of voxels input to the network. Training is carried out over 150 epochs using stochastic gradient decent with a learning rate of 0.01 to optimise the networks approximately thirty million trainable parameters.

$$D(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (3.1)$$

The network is implemented using Keras [noauthor\_keras\_2019] with a TensorFlow [noauthor\_tensorflow\_2019] back-end and is trained on an NVIDIA Titan Xp Graphical Processing Unit (GPU). Train-

ing takes approximately forty minutes for the 150 epochs and predicting a renal mask from a thirteen slice volume takes approximately 9 seconds when executed on a computer with no GPU (as it would be in most end user cases).

### 3.3 Results and Discussion

Initial data was collected with a TR of 1800 ms leading to a breath hold of approximately 28 seconds. Some subjects struggled to hold their breath for this long on expiration, therefore the effects of reducing the TR of the sequence were investigated. As can be seen in Figure 3.2, there is no degradation in image quality from the image with TR or 1800 ms to that with at TR of 1300 ms, the differences between these images are mainly due to the small movements between volumes, as can be seen in the difference data where the largest differences are seen around the periphery of the kidneys and in the gut. Moving forward, the TR was reduced to 1300 ms leading to a sequence with a breath hold of approximately 17 seconds.

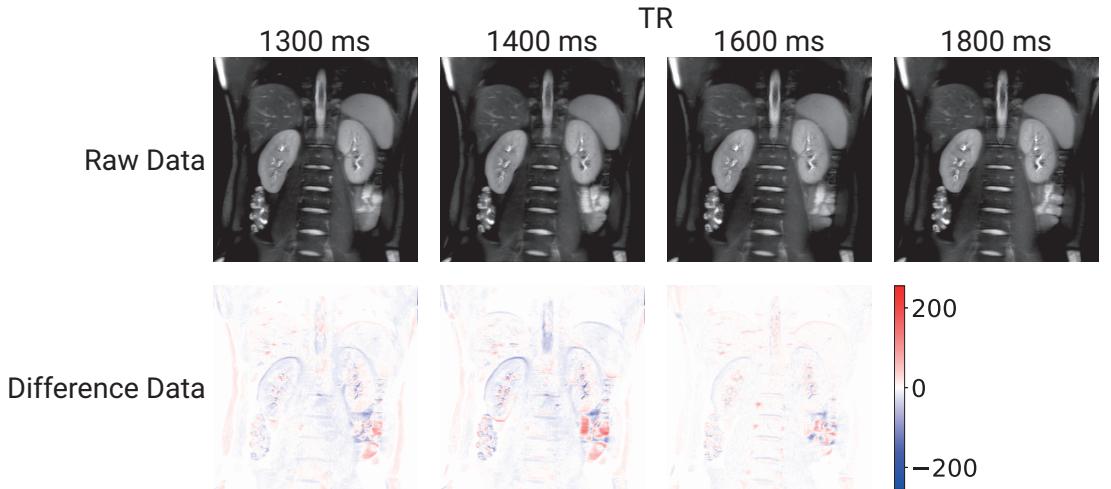


Figure 3.2: The effects of changing the TR of the sequence.

To verify that the trained network is behaving as expected saliency maps were produced, Figure 3.3, this is especially important given the black box nature of machine learning methods. This map shows the areas the network is using most in its classification [71]. It verifies that the networks is using the outside areas of the kidney to make its prediction with areas of a similar intensity receiving some attention to distinguish them from the kidney. While this is precisely what is expected of the algorithm, it is important to check this as it is possible for such a method to have learnt a slightly different mechanism for the segmentation, one that is more prone to errors if new data is presented to it.

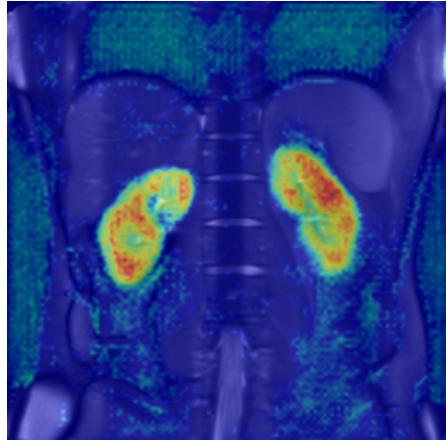


Figure 3.3: An example saliency map of the areas the network uses most when segmenting the kidney.

To assess the accuracy of the network, each of the five volumes per validation subject was segmented by the trained network, in theory the TKV predicted for each volume should have been the same. Figure 3.4a shows the predicted TKV against the manually segmented “true” TKV with each subject plot in a different colour. While there is a spread in the predicted values, there is also a reasonable variation in manual TKV. For three out of the seven repeatability subjects, the predicted TKV has a smaller standard deviation than the manually segmented data, this indicates that the algorithm may actually be more consistent than the humans. To identify if a systematic error is present, a Bland-Altman plot was generated (Figure 3.4b). From this figure it is possible to see that the algorithm is slightly over estimating the TKV by 0.69% ( $1.6 \text{ mm}^3$ ) but there is no correlation between difference in TKV and true TKV over all subjects. There is a more subtle trend between each repeat on the same subject though, the volumes with a smaller true TKV are consistently over estimated more than those volumes with a larger true TKV. This again points to an issue with the manual segmentation rather than the algorithm.

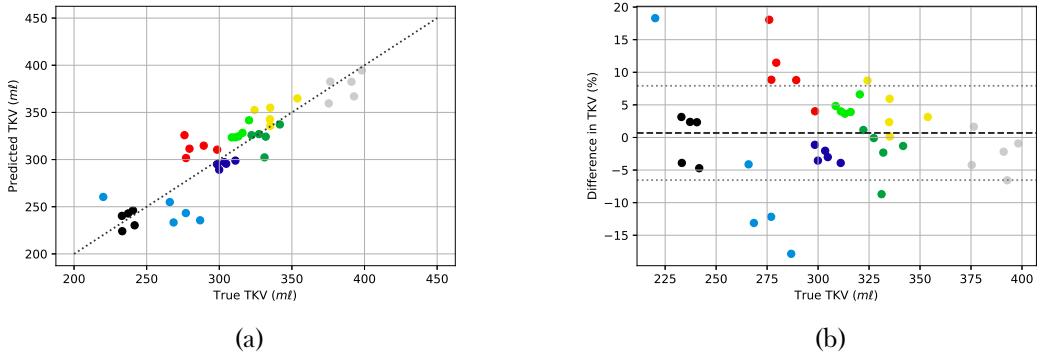


Figure 3.4: (a) The predicted TKV plot against the manually segmented “true” TKV. Each subject is plotted in a different colour. (b) A Bland-Altman plot to identify and systematic error in the networks performance. Each subject is plotted in a different colour.

While assessing the ability of the algorithm to predict TKV is important, it is also necessary to assess the raw segmentation as, for example, the algorithm may be over-estimating the size of central slices but under-estimating the size of edge slices. This type of inaccuracy could be masked in the TKV however will be visible in the dice scores. These are shown in Figure 3.5.

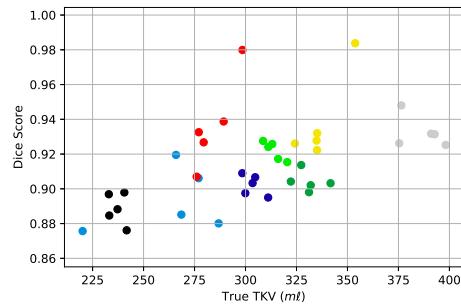


Figure 3.5: The dice scores for all volumes in the validation data. Each subject is plotted in a different colour.

The mean dice score over all forty volumes is  $0.91 \pm 0.02$ . Here we see a slight trend towards more accurate predictions for larger kidneys. The reason for this becomes clear when we look at the ROI the algorithm is outputting, Figure 3.6.

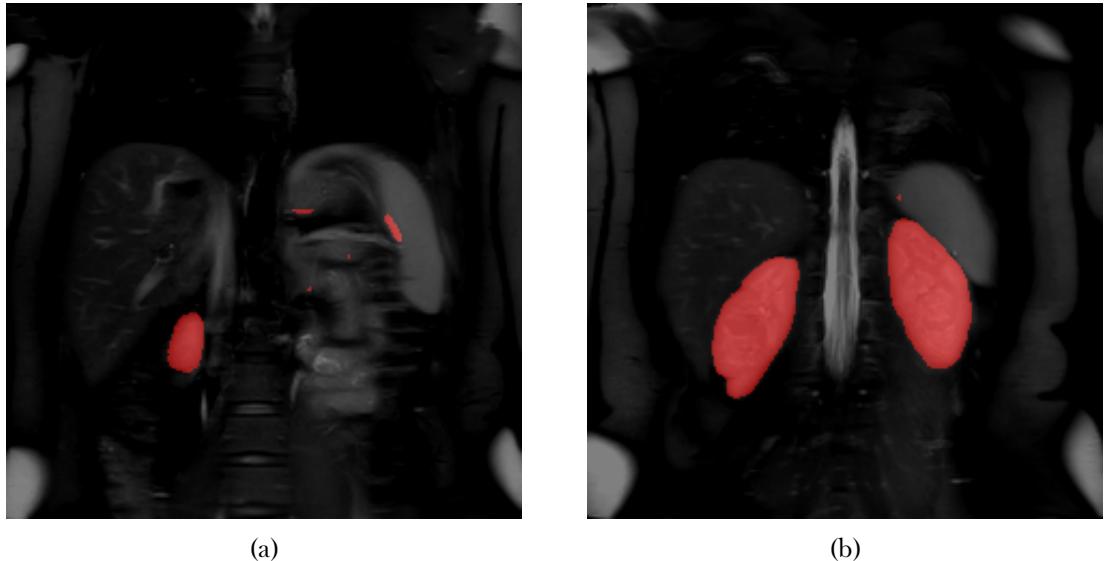


Figure 3.6: (a) A slice from the posterior of the volume. (b) A slice from the centre of the kidneys.

In Figure 3.6a the algorithm is assigning false positives on the right hand side of the image in the area a kidney would be expected further into the body. The amount of false positives decrease as the slices move in an anterior direction as kidney comes into the slice, 3.6b. The algorithm works on each slice individually as a two-dimensional image rather than as a three-dimensional volume. This means that, as the majority of slices in the training data have two kidneys in them, the algorithm is more likely to generate false positives if there is no kidney in the slice. For smaller kidneys, there are more slices with no kidney in them and therefore the overall dice score is lower, hence the trend observed in Figure 3.5.

There are multiple methods of reducing this tendency in the algorithm. The false positives tend to be spatially independent through slices, this means that it would be relatively simple to remove them in post-processing by reconstructing the two-dimensional slices back into a three-dimensional volume and removing masked areas with a small volume or areas that are very thin in the anterior-posterior direction. Another method would be to modify the architecture to a Recursive Neural Network (RNN) with Long Short-Term Memory (LSTM). This would also keep the large advantage of working with two-dimensional images, that the algorithm generalises to  $n$  slices, but means that the algorithm also has some memory of what is in the surrounding slices [72, 73]. Finally the algorithm could be re-written as a three-dimensional FCN, this would give the greatest degree of accuracy between slices however comes at the expense of simple generalisation with regards to number of slices or the slice thickness and would require much more data collection as

amount of training/test data would be reduced by a factor of approximately thirteen.

To establish how the network is performing with the relatively small amount of training data, predictions were made on the training and testing data and the dice score plot, Figure 3.7.

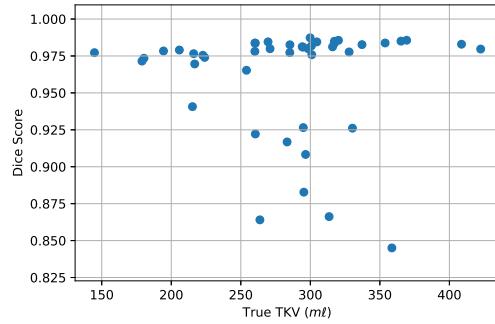


Figure 3.7: The dice score of predictions made on the training data.

The algorithm is performing better on most of the training data than it did on the validation data although with 80% of the volumes segmented more accurately than the validation data. This indicates that a certain degree of over fitting is occurring as the 20% of volumes that are not segmented as well are most likely the testing data. Earlier in the development of this network it was established that augmentation did not improve the performance of the trained network however this should be explored again in light of this result as some basic augmentation would lead to a smaller disparity between the training data and test data and thus allow for better performance when segmenting the validation data.

An indication that some degree of data augmentation would be beneficial is also seen when investigating if there is any difference in performance of the network between healthy and CKD kidneys. The manually segmented mean TKV of the healthy subjects is  $330 \pm 35$  mL and for subjects with CKD is  $268 \pm 32$  mL therefore it would be expected that the algorithm would perform better on the healthy subjects given their larger kidney volume. This is not the case though, the mean dice score of the validation images for healthy kidneys is  $0.89 \pm 0.02$  and for kidneys with CKD this increases to  $0.93 \pm 0.02$ . As there are more healthy subjects in the training data (26 versus 23) it would be expected that the network would perform better for these subjects however the greater degree of variability in geometry and size of the CKD kidneys means they essentially have some degree of augmentation built into them. If this were replicated via data augmentation in the whole

training dataset then an increase in accuracy across the board may be observed.

### 3.4 Conclusions and Future Work

This method has been shown to produce promising results delivering an mean dice score of  $0.91 \pm 0.02$  over eight unseen scans with a mixture of healthy and CKD subjects resulting in a mean TKV difference of 0.69% when compared to the manually segmented TKV. This is especially promising as the algorithm will improve in accuracy as more training data is collected, something which the renal group at SPMIC are actively pursuing by adding this scan to almost every subject that goes in the scanner. Efforts have been made to avoid the quintessential machine learning mistakes such as imbalanced training data making the algorithm too specific and un-generalisable and the network only working for healthy subjects. We have also peaked inside the black box of the algorithm to check that it is behaving in a sensible and expected manner.

There is still work to be done on this segmentation method, as explained above, there are signs that data augmentation may improve both the accuracy and generalisability of the algorithm, this should be implemented and evaluated. By implementing data augmentation, the false positives observed on fringe slices may decrease however, if this is not the case then there are multiple solutions to reduce the prevalence of these errors such as basic binary filtering or modifying the networks architecture. There seems to be a reasonable degree of variability in the manually segmented data, to investigate this, the manual masking process should be repeated by a second observer at assess if this variability in the data is due to acquisition or human interpretation.

Another common segmentation task in renal imaging is to generate ROI for the renal cortex and medulla. There are some automated methods of achieving this once an overall renal mask has been produced [12], however there has been no work on the application of machine learning to this task. During the acquisition of the  $T_2$  weighted data in Section 3.2.1, a sequence designed to optimise the contrast between cortex and medulla was also collected on each subject, an example of which is shown in Figure 3.8. Using this data it may be possible to develop the algorithm further so it can segment each tissue type within the kidneys.



Figure 3.8: An example of the data collected to enable segmentation of the renal cortex and medulla.

Ultimately the goal of this work is to produce an easy to use segmentation tool that can be utilised by clinicians and scientists alike. As such, time should be spent making the software easy to use with a simple front end/Graphical User Interface (GUI).

# **Chapter 4**

## **Other Work**

Below is a brief synopsis of other work undertaken over the past two years.

### **4.1 Layer Based Analysis of the Kidneys**

The vast majority of analysis of renal MRI data is based around defining ROI within the kidneys. While this method has provided excellent results, it is by no means perfect as these ROI need to be manually defined, leading to human bias, or defined by an automated method which, as outlined in Chapter 3, can be difficult to generalise.

Taking inspiration from the analysis pipelines used by neuroimagers [74–76], a method of dividing the kidneys into layers of equal thickness was developed. This method uses a three-dimensional FreeSurfer mesh on the surface of the kidney [77] and levelsets to produce a map of how far each voxel is from the surface. From this map it is possible to place voxels into layers of any thickness. An example of this method being applied to both the brain and an ex-vivo kidney sample can be seen, figures 4.1a and 4.1b respectively.

One of the main challenges in transferring this technique is coping with the reduced FOV that comes with body imaging. Given that the method essentially asks how far is each voxel from the closest vertex on the mesh, if the mesh has a large hole in it where the slices stop covering the kidney, then the quantitative nature of this depth map is compromised.

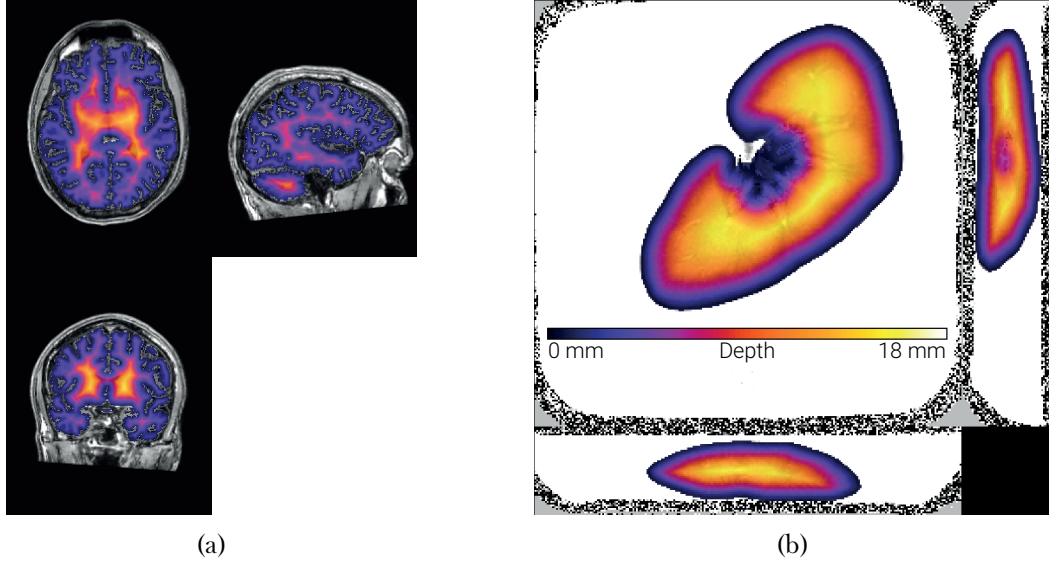


Figure 4.1: (a) A depth mask of the brain. Lighter areas are deeper inside the brain. (b) A depth mask applied to a quantitative  $T_1$  map.

This levelset method was compared to two other methods of producing layers in the kidneys, a two-dimensional and a three-dimensional version of the Twelve Layer Concentric Objects (TLCO) method [78, 79]. Each method was tested with a volume that included the entire kidney, and a cropped volume that only included the central section of the kidney, simulating the reduced FOV that is common in body imaging. The layers generated were applied to a  $T_1$  map with the mean and standard deviation of  $T_1$  in each layer being plotted, Figure 4.2.

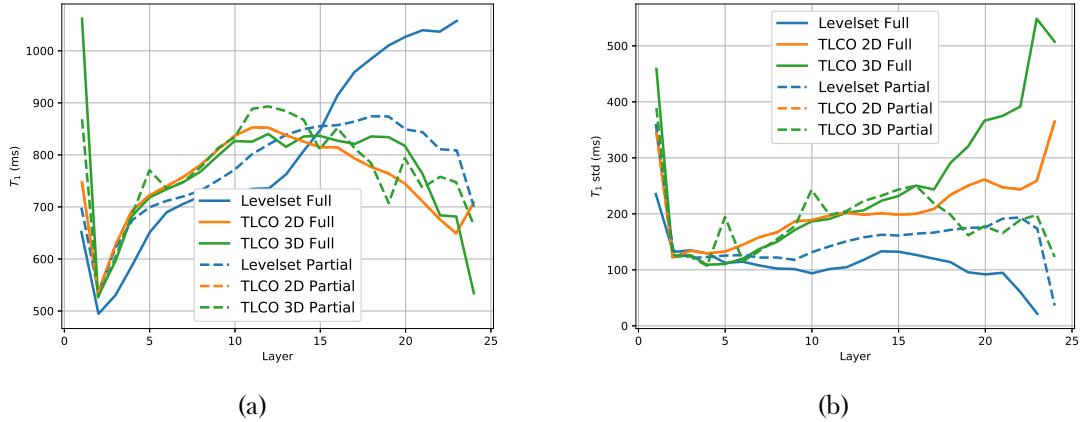


Figure 4.2: (a) The mean  $T_1$  within each layers produced by each of the three methods when processing either the full volume of the kidney or only the central slices. (b) The standard deviation of the  $T_1$  within the layers produced by each method.

In Figure 4.2b we can see that the layers output by the levelset method when applied to a full volume dataset produce the minimum standard deviation, this means the layers are the most anatomically sensible as a smooth transition of  $T_1$  is expected with depth in

the kidney and therefore the variance in each layer should be relatively small. Given this method produces the most realistic layers, other methods will be compared to it.

Neither of the TLCO methods manage to capture the increase in  $T_1$  that can be seen deeper in the kidney. They also have a much larger standard deviation per layer for deeper layers than the levelset method implying that the layers produced are a mixture of cortex and medulla. When comparing the performance of each method with only a partial volume of kidney, the levelset method produces the results that are closer to that of the full volume levelset.

Given this method can be used both in-vivo and on ex-vivo samples, they will make for an interesting additional analysis pipeline for the work in Chapter 1.

## 4.2 Diffusion Tensor Imaging in the Kidneys

As part of the nephrectomy protocol, Chapter 1, we want to be able to assess the microstructure of the kidneys, one avenue to pursue for this is the use of Diffusion Tensor Imaging (DTI). FA has been shown to correlate with GFR [80] and mean tract length is an indicator of ureteropelvic junction obstruction [81] showing that tractography can be used to assess renal structure.

One of the major hurdles to overcome in developing a renal DTI protocol for this study was correction of both EPI readout distortions and eddy current induced distortions. This was achieved using a pipeline based around FSL’s topup [82, 83] and eddy [84] routines. Key acquisition parameters for the sequence are 64 directions arranged over a whole spherical shell to assist with eddy current correction, this whole acquisition is then repeated with the opposite phase-encode direction to enable EPI distortion correction. A b-value of 600 s/mm<sup>2</sup> is used. FA maps are generated using FSL and tractography is processed using an in-house pipeline developed using the Dipy Python library [85].

Using this pipeline, images such as those in figures 4.3 and 4.4 could be produced. This protocol is now ready to be used on patients undergoing a nephrectomy.

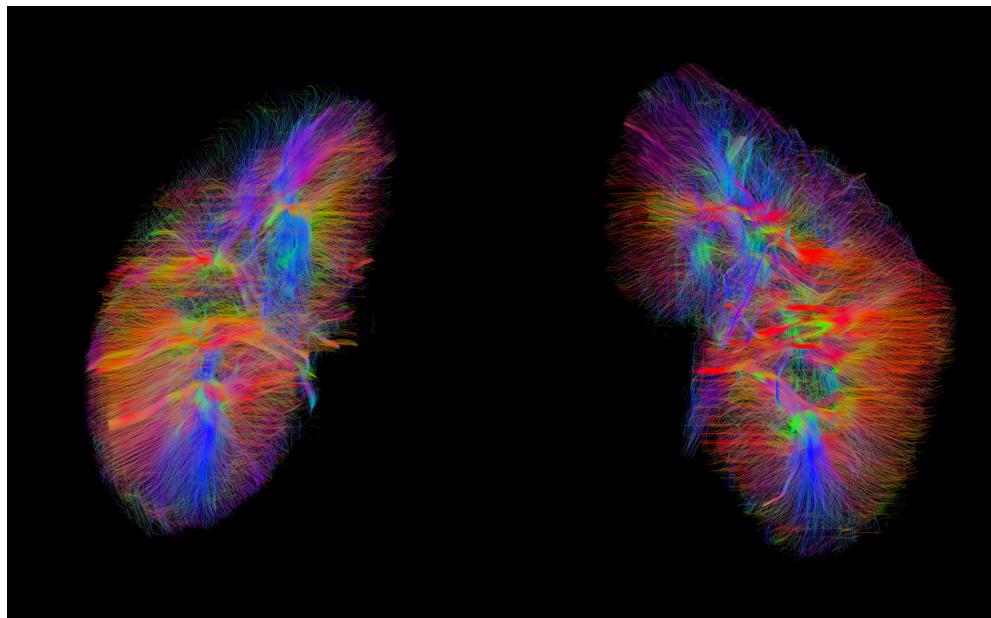


Figure 4.3: Example tractography generated using the above protocol.

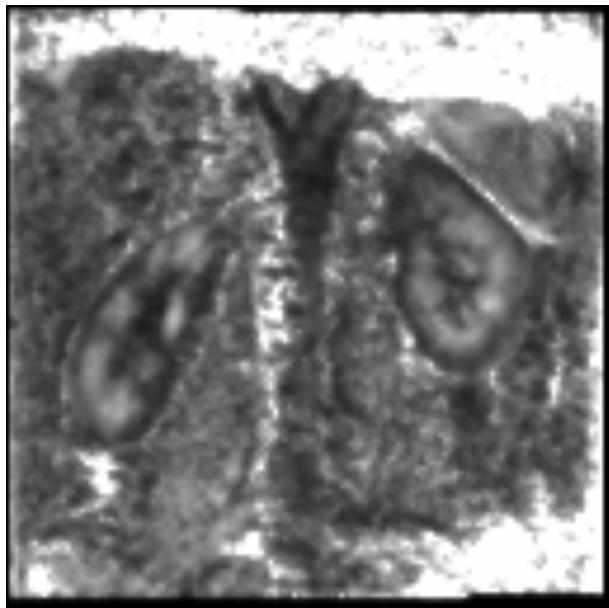


Figure 4.4: An example FA map generated using the protocol above.

### 4.3 Designing Novel MR Phantoms using Additive Manufacturing

Phantoms used in the field of MRI are generally limited to blocks of doped agar [86] or arrays of liquids of known properties as used in Section 1.2.3. While these types of phantom are useful, they have remained largely unchanged for many years. With the advent of modern manufacturing techniques such as Additive Manufacturing (AM) far more complex phantoms could be designed and built relatively inexpensively.

We plan to begin by validating the manufacturing process using a phantom similar to one designed by Berry et al [87], Figure 4.5. These phantoms restrict the direction in which the fluid within them can diffuse leading to a higher FA for the phantoms with smaller tubules. A technical drawing of one of these proposed phantoms can be seen in appendix ???. These small structures would be very difficult and costly to produce before recent developments in AM.

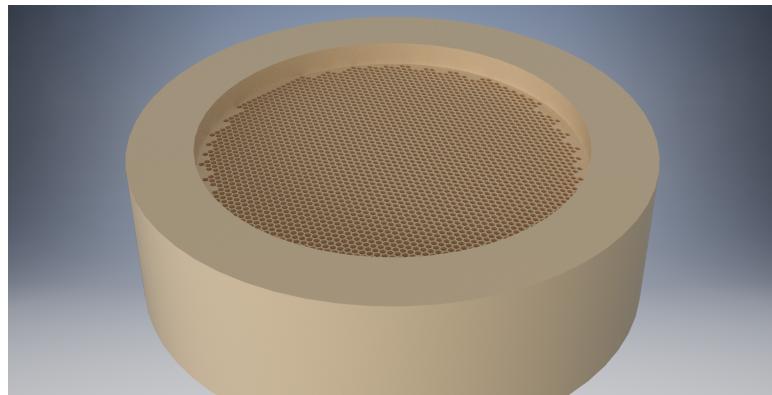


Figure 4.5: A render of one of the proposed phantoms, in this case each small hexagon is  $50\ \mu\text{m}$  in diameter with a  $10\ \mu\text{m}$  wall separating them.

Extending this concept, we intend to print a phantom to simulate fibre crossing, appendix ???. This phantom contains small tubules, half of which always remain in the same direction, the other half of which gradually rotate though forty five degrees. This phantom therefore simulates multiple angles of fibres crossing found in-vivo and should enable the verification of multiple DTI tractography reconstruction algorithms, something that has only been performed using somewhat crude phantoms until now [88].

A final proposal is to use custom printed phantoms to investigate the accuracy of flow measurements in the kidneys. Similar work has been carried out in other areas of the body

[89, 90] however, no such investigation has been done on renal vessels. We propose to begin with a basic flow phantom, Figure 4.6a and appendix ??, and verify that the MRI measurements correlate well with those predicted using Computational Fluid Dynamics (CFD), Figure 4.6b, before moving onto more complicated geometry of the renal blood vessels comparing in-vivo, ex-vivo and in-silico results.

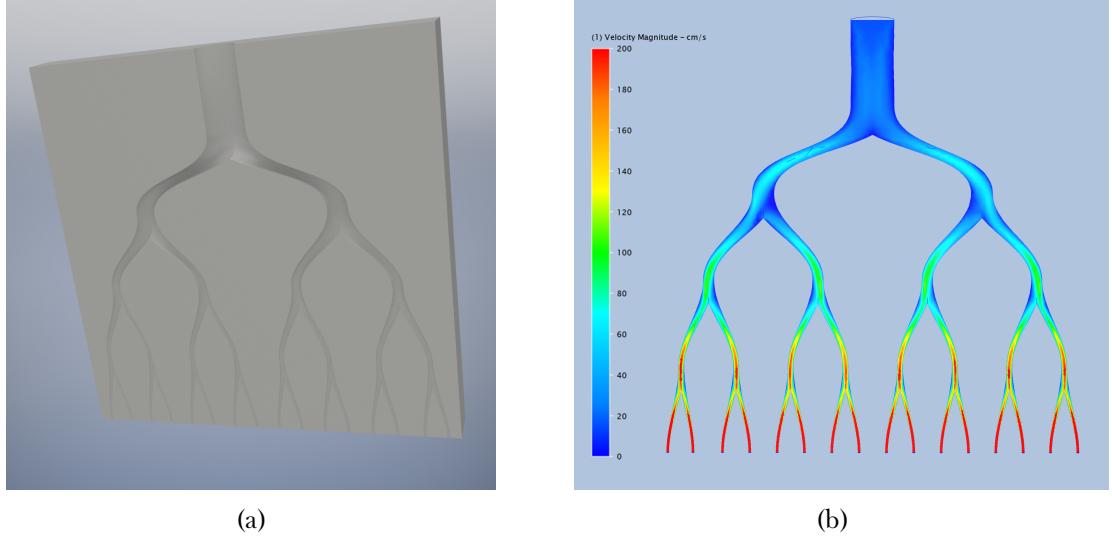


Figure 4.6: (a) The proposed basic flow phantom with branching structure. (b) CFD simulations of the flow through the proposed phantom.

# Acknowledgements

We thank Hanzhang Lu for sharing the TRUST methodology.

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

We are grateful for access to the University of Nottingham's Augusta high performance computing service.

# Chapter 5

## References

1. Friedli, I. *et al.* New Magnetic Resonance Imaging Index for Renal Fibrosis Assessment: A Comparison between Diffusion-Weighted Imaging and T1 Mapping with Histological Validation. *Scientific Reports* **6**. ISSN: 2045-2322. pmid: 27439482 (21st July 2016).
2. Risdon, R. A., Sloper, J. C. & De Wardener, H. E. Relationship Between Renal Function and Histological Changes Found In Renal-Biopsy Specimens From Patients With Persistent Glomerular Nephritis. *The Lancet. Originally Published as Volume 2, Issue 7564* **292**, 363–366. ISSN: 0140-6736 (17th Aug. 1968).
3. Uribe, C. F. *et al.* In Vivo 3 T and Ex Vivo 7 T Diffusion Tensor Imaging of Prostate Cancer: Correlation with Histology. *Magnetic Resonance Imaging* **33**, 577–583. ISSN: 0730-725X, 1873-5894. pmid: 25721995 (1st June 2015).
4. Sengupta, S. *et al.* High Resolution Anatomical and Quantitative MRI of the Entire Human Occipital Lobe Ex Vivo at 9.4T. *NeuroImage*. ISSN: 1053-8119 (20th Mar. 2017).
5. Wolf, M. *et al.* Magnetic Resonance Imaging T1- and T2-Mapping to Assess Renal Structure and Function: A Systematic Review and Statement Paper. *Nephrology Dialysis Transplantation* **33**, ii41–ii50. ISSN: 0931-0509 (suppl\_2 1st Sept. 2018).
6. Franke, M. *et al.* Magnetic Resonance T2 Mapping and Diffusion-Weighted Imaging for Early Detection of Cystogenesis and Response to Therapy in a Mouse Model of Polycystic Kidney Disease. *Kidney International* **92**, 1544–1554. ISSN: 0085-2538 (1st Dec. 2017).
7. Li, X., Bolan, P. J., Ugurbil, K. & Metzger, G. J. Measuring Renal Tissue Relaxation Times at 7 T. *NMR in Biomedicine* **28**, 63–69. ISSN: 1099-1492 (2015).

8. Zhang, J. L. *et al.* Reproducibility of R2\* and R2 Measurements in Human Kidneys in *Proc. Intl. Soc. Mag. Reson. Med.* **19** ISMRM. **19** (2011), 2954.
9. Szumowski, J. *et al.* Signal Polarity Restoration in a 3D Inversion Recovery Sequence Used with Delayed Gadolinium-Enhanced Magnetic Resonance Imaging of Cartilage (dGEMRIC). *Journal of Magnetic Resonance Imaging* **36**, 1248–1255. ISSN: 1522-2586 (1st Nov. 2012).
10. Branch, M., Coleman, T. & Li, Y. A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems. *SIAM Journal on Scientific Computing* **21**, 1–28. ISSN: 1064-8275 (1st Jan. 1999).
11. System Phantom for Diagnostic Quantitative MRI <https://www.qalibre-md.com/system-phantom/> (2019).
12. Cox, E. F. *et al.* Multiparametric Renal Magnetic Resonance Imaging: Validation, Interventions, and Alterations in Chronic Kidney Disease. *Frontiers in Physiology* **8**. ISSN: 1664-042X (2017).
13. Dawe, R. J., Bennett, D. A., Schneider, J. A., Vasireddi, S. K. & Arfanakis, K. Post-mortem MRI of Human Brain Hemispheres: T2 Relaxation Times during Formaldehyde Fixation. *Magnetic Resonance in Medicine* **61**, 810–818. ISSN: 1522-2594.
14. Van der Kolk, A. G. *et al.* Imaging the Intracranial Atherosclerotic Vessel Wall Using 7T MRI: Initial Comparison with Histopathology. *American Journal of Neuroradiology*. ISSN: 0195-6108, 1936-959X. pmid: 25477359 (4th Dec. 2014).
15. Stevens, L. A., Coresh, J., Greene, T. & Levey, A. S. Assessing Kidney Function — Measured and Estimated Glomerular Filtration Rate. *New England Journal of Medicine* **354**, 2473–2483. ISSN: 0028-4793. pmid: 16760447 (8th June 2006).
16. Farrugia, A. Albumin Usage in Clinical Medicine: Tradition or Therapeutic? *Transfusion Medicine Reviews* **24**, 53–63. ISSN: 0887-7963, 1532-9496. pmid: 19962575 (1st Jan. 2010).
17. Pruijm, M., Milani, B. & Burnier, M. Blood Oxygenation Level-Dependent MRI to Assess Renal Oxygenation in Renal Diseases: Progresses and Challenges. *Frontiers in Physiology* **7**. ISSN: 1664-042X. pmid: 28105019 (5th Jan. 2017).
18. Coresh, J., Astor, B. C., Greene, T., Eknoyan, G. & Levey, A. S. Prevalence of Chronic Kidney Disease and Decreased Kidney Function in the Adult US Population: Third National Health and Nutrition Examination Survey. *American Journal of Kidney Diseases*

- eases: *The Official Journal of the National Kidney Foundation* **41**, 1–12. issn: 1523-6838. pmid: 12500213 (Jan. 2003).
19. De Lusignan, S. *et al.* Identifying Patients with Chronic Kidney Disease from General Practice Computer Records. *Family Practice* **22**, 234–241. issn: 0263-2136 (1st June 2005).
  20. Drey, N., Roderick, P., Mullee, M. & Rogerson, M. A Population-Based Study of the Incidence and Outcomes of Diagnosed Chronic Kidney Disease. *American Journal of Kidney Diseases* **42**, 677–684. issn: 0272-6386 (1st Oct. 2003).
  21. Amato, D. *et al.* Prevalence of Chronic Kidney Disease in an Urban Mexican Population. *Kidney International* **68**, S11–S17. issn: 0085-2538 (1st Aug. 2005).
  22. Chadban, S. J. *et al.* Prevalence of Kidney Damage in Australian Adults: The Aus-Diab Kidney Study. *Journal of the American Society of Nephrology* **14**, S131–S138. issn: 1046-6673, 1533-8450. pmid: 12819318 (suppl 2 7th Jan. 2003).
  23. Jungers, P. *et al.* Late Referral to Maintenance Dialysis: Detrimental Consequences. *Nephrology Dialysis Transplantation* **8**, 1089–1098. issn: 0931-0509 (1st Jan. 1993).
  24. Sesso, R. & Belasco, A. G. Late Diagnosis of Chronic Renal Failure and Mortality on Maintenance Dialysis. *Nephrology Dialysis Transplantation* **11**, 2417–2420. issn: 0931-0509 (1st Dec. 1996).
  25. Klebe, B. *et al.* The Cost of Implementing UK Guidelines for the Management of Chronic Kidney Disease. *Nephrology Dialysis Transplantation* **22**, 2504–2512. issn: 0931-0509 (1st Sept. 2007).
  26. Precious, K. *NHS England » Programme Budgeting* <https://www.england.nhs.uk/resources/resources-for-ccgs/prog-budgeting/> (2017).
  27. Niendorf, T. *et al.* How Bold Is Blood Oxygenation Level-Dependent (BOLD) Magnetic Resonance Imaging of the Kidney? Opportunities, Challenges and Future Directions. *Acta Physiologica* **213**, 19–38. issn: 1748-1716.
  28. Chong, S. P., Merkle, C. W., Leahy, C. & Srinivasan, V. J. Cerebral Metabolic Rate of Oxygen (CMRO<sub>2</sub>) Assessed by Combined Doppler and Spectroscopic OCT. *Bio-medical Optics Express* **6**, 3941–3951. issn: 2156-7085. pmid: 26504644 (11th Sept. 2015).
  29. Jordan, J. E., Pelc, N. J. & Enzmann, D. R. Velocity and Flow Quantitation in the Superior Sagittal Sinus with Ungated and Cine (Gated) Phase-Contrast MR Imaging. *Journal of Magnetic Resonance Imaging* **4**, 25–28. issn: 1522-2586 (1st Jan. 1994).

30. Miao, G. *et al.* Reference Value of Presenile Human Hematocrit and Geographical Factors. *Journal of Clinical Laboratory Analysis* **16**, 26–29. ISSN: 1098-2825 (1st Jan. 2002).
31. Gardener, A. G., Francis, S. T., Prior, M., Peters, A. & Gowland, P. A. Dependence of Blood R2 Relaxivity on CPMG Echo-Spacing at 2.35 and 7 T. *Magnetic Resonance in Medicine* **64**, 967–974. ISSN: 1522-2594 (1st Oct. 2010).
32. Shimada, K., Nagasaka, T., Shidahara, M., Machida, Y. & Tamura, H. In Vivo Measurement of Longitudinal Relaxation Time of Human Blood by Inversion-Recovery Fast Gradient-Echo MR Imaging at 3T. *Magnetic resonance in medical sciences: MRMS: an official journal of Japan Society of Magnetic Resonance in Medicine* **11**, 265–271. ISSN: 1880-2206. pmid: 23269013 (2012).
33. Nagdyman, N. *et al.* Comparison between Cerebral Tissue Oxygenation Index Measured by Near-Infrared Spectroscopy and Venous Jugular Bulb Saturation in Children. *Intensive Care Medicine* **31**, 846–850. ISSN: 0342-4642, 1432-1238 (1st June 2005).
34. Lu, H. & Ge, Y. Quantitative Evaluation of Oxygenation in Venous Vessels Using T2-Relaxation-Under-Spin-Tagging MRI. *Magnetic Resonance in Medicine* **60**, 357–363. ISSN: 1522-2594 (1st Aug. 2008).
35. Xu, F., Uh, J., Liu, P. & Lu, H. On Improving the Speed and Reliability of T2-Relaxation-Under-Spin-Tagging (TRUST) MRI. *Magnetic Resonance in Medicine* **68**, 198–204. ISSN: 0740-3194. pmid: 22127845 (July 2012).
36. Liu, P., Xu, F. & Lu, H. Test–Retest Reproducibility of a Rapid Method to Measure Brain Oxygen Metabolism. *Magnetic Resonance in Medicine* **69**, 675–681. ISSN: 1522-2594 (1st Mar. 2013).
37. Liu, P. *et al.* Multi-Site Evaluations of a T2-Relaxation-Under-Spin-Tagging (TRUST) MRI Technique to Measure Brain Oxygenation. *Magnetic resonance in medicine* **75**, 680–687. ISSN: 0740-3194. pmid: 25845468 (Feb. 2016).
38. Jain, V., Langham, M. C. & Wehrli, F. W. MRI Estimation of Global Brain Oxygen Consumption Rate. *Journal of Cerebral Blood Flow & Metabolism* **30**, 1598–1607. ISSN: 0271-678X (1st Sept. 2010).
39. Jain, V. *et al.* Cerebral Oxygen Metabolism in Neonates with Congenital Heart Disease Quantified by MRI and Optics. *Journal of Cerebral Blood Flow & Metabolism* **34**, 380–388. ISSN: 0271-678X (1st Mar. 2014).

40. Driver, I. D. *et al.* Global Intravascular and Local Hyperoxia Contrast Phase-Based Blood Oxygenation Measurements. *NeuroImage* **101**, 458–465. issn: 1053-8119 (1st Nov. 2014).
41. Lee, H., Langham, M. C., Rodriguez-Soto, A. E. & Wehrli, F. W. Multiplexed MRI Methods for Rapid Estimation of Global Cerebral Metabolic Rate of Oxygen Consumption. *NeuroImage* **149**, 393–403. issn: 1053-8119 (1st Apr. 2017).
42. Wright, G. A., Hu, B. S. & Macovski, A. Estimating Oxygen Saturation of Blood in Vivo with MR Imaging at 1.5 T. *Journal of Magnetic Resonance Imaging* **1**, 275–283. issn: 1522-2586 (1st May 1991).
43. Haacke, E. M. *et al.* In Vivo Measurement of Blood Oxygen Saturation Using Magnetic Resonance Imaging: A Direct Validation of the Blood Oxygen Level-Dependent Concept in Functional Brain Imaging. *Human Brain Mapping* **5**, 341–346. issn: 1097-0193 (1st Jan. 1997).
44. Spees, W. M., Yablonskiy, D. A., Oswood, M. C. & Ackerman, J. J. Water Proton MR Properties of Human Blood at 1.5 Tesla: Magnetic Susceptibility, T1, T2, T\*2, and Non-Lorentzian Signal Behavior. *Magnetic Resonance in Medicine* **45**, 533–542. issn: 1522-2594 (1st Apr. 2001).
45. Jain, V., Abdulmalik, O., Propert, K. J. & Wehrli, F. W. Investigating the Magnetic Susceptibility Properties of Fresh Human Blood for Noninvasive Oxygen Saturation Quantification. *Magnetic Resonance in Medicine* **68**, 863–867. issn: 1522-2594 (1st Sept. 2012).
46. Jenkinson, M. Fast, Automated, N-Dimensional Phase-Unwrapping Algorithm. *Magnetic Resonance in Medicine* **49**, 193–197. issn: 1522-2594 (1st Jan. 2003).
47. Hendrikse, J., Lu, H., van der Grond, J., van Zijl, P. C. & Golay, X. Measurements of Cerebral Perfusion and Arterial Hemodynamics during Visual Stimulation Using TURBO-TILT. *Magnetic Resonance in Medicine* **50**, 429–433. issn: 1522-2594 (1st Aug. 2003).
48. Golay, X., Petersen, E. T. & Hui, F. Pulsed Star Labeling of Arterial Regions (PULSAR): A Robust Regional Perfusion Technique for High Field Imaging. *Magnetic Resonance in Medicine* **53**, 15–21. issn: 1522-2594 (1st Jan. 2005).
49. Golay, X., Stuber, M., Pruessmann, K. P., Meier, D. & Boesiger, P. Transfer Insensitive Labeling Technique (TILT): Application to Multislice Functional Perfusion Im-

- ging. *Journal of Magnetic Resonance Imaging* **9**, 454–461. issn: 1522-2586 (1st Mar. 1999).
50. Lu, H., Clingman, C., Golay, X. & van Zijl, P. C. Determining the Longitudinal Relaxation Time (T1) of Blood at 3.0 Tesla. *Magnetic Resonance in Medicine* **52**, 679–682. issn: 1522-2594 (1st Sept. 2004).
  51. Silvennoinen, M., Clingman, C., Golay, X., Kauppinen, R. & van Zijl, P. Comparison of the Dependence of Blood R2 and R2\* on Oxygen Saturation at 1.5 and 4.7 Tesla. *Magnetic Resonance in Medicine* **49**, 47–60. issn: 1522-2594 (1st Jan. 2003).
  52. Liu, P. *et al.* T1 and T2 Values of Human Neonatal Blood at 3 Tesla: Dependence on Hematocrit, Oxygenation, and Temperature. *Magnetic Resonance in Medicine* **75**, 1730–1735. issn: 1522-2594. pmid: 25981985 (Apr. 2016).
  53. Lu, H. *et al.* Calibration and Validation of TRUST MRI for the Estimation of Cerebral Blood Oxygenation. *Magnetic Resonance in Medicine* **67**, 42–49. issn: 1522-2594 (1st Jan. 2012).
  54. Martirosian, P., Klose, U., Mader, I. & Schick, F. FAIR True-FISP Perfusion Imaging of the Kidneys. *Magnetic Resonance in Medicine* **51**, 353–361. issn: 1522-2594 (1st Feb. 2004).
  55. Liu, P., Cox, E. & Daniel, A. *Pro TRUST* version V2018.5. 23rd Aug. 2011.
  56. O'Connor, J. P. *et al.* Comparison of Normal Tissue R1 and R\*2 Modulation by Oxygen and Carbogen. *Magnetic Resonance in Medicine* **61**, 75–83. issn: 1522-2594 (1st Jan. 2009).
  57. Donati, O. F., Nanz, D., Serra, A. L. & Boss, A. Quantitative BOLD Response of the Renal Medulla to Hyperoxic Challenge at 1.5 T and 3.0 T. *NMR in Biomedicine* **25**, 1133–1138. issn: 1099-1492 (1st Oct. 2012).
  58. Pruijm, M. *et al.* Effect of Sodium Loading/Depletion on Renal Oxygenation in Young Normotensive and Hypertensive Men. *Hypertension* **55**, 1116–1122. issn: 0194-911X, 1524-4563. pmid: 20308608 (1st May 2010).
  59. Tumkur, S. M., Vu, A. T., Li, L. P., Pierchala, L. & Prasad, P. V. Evaluation of Intra-Renal Oxygenation during Water Diuresis: A Time-Resolved Study Using BOLD MRI. *Kidney International* **70**, 139–143. issn: 0085-2538 (1st July 2006).

60. Prasad, P. V. & Epstein, F. H. Changes in Renal Medullary pO<sub>2</sub> during Water Diuresis as Evaluated by Blood Oxygenation Level-Dependent Magnetic Resonance Imaging: Effects of Aging and Cyclooxygenase Inhibition. *Kidney International* **55**, 294–298. issn: 0085-2538 (1st Jan. 1999).
61. Nielsen, K., Rehling, M. & Henriksen, J. H. Renal Vein Oxygen Saturation in Renal Artery Stenosis. *Clinical Physiology* **12**, 179–184. issn: 1365-2281 (1st Mar. 1992).
62. Cohen, E. I., Kelly, S. A., Edye, M., Mitty, H. A. & Bromberg, J. S. MRI Estimation of Total Renal Volume Demonstrates Significant Association with Healthy Donor Weight. *European Journal of Radiology. Rheumatoid Disease and Imaging* **71**, 283–287. issn: 0720-048X (1st Aug. 2009).
63. Van den Dool, S. W., Wasser, M. N., de Fijter, J. W., Hoekstra, J. & van der Geest, R. J. Functional Renal Volume: Quantitative Analysis at Gadolinium-Enhanced MR Angiography—Feasibility Study in Healthy Potential Kidney Donors. *Radiology* **236**, 189–195. issn: 0033-8419 (1st July 2005).
64. Chapman, A. B. *et al.* Kidney Volume and Functional Outcomes in Autosomal Dominant Polycystic Kidney Disease. *Clinical Journal of the American Society of Nephrology* **7**, 479–486. issn: 1555-9041, 1555-905X. pmid: 22344503 (1st Mar. 2012).
65. Tangri, N. *et al.* Total Kidney Volume as a Biomarker of Disease Progression in Autosomal Dominant Polycystic Kidney Disease. *Canadian Journal of Kidney Health and Disease* **4**, 2054358117693855. issn: 2054-3581 (1st Jan. 2017).
66. Gong, I. H. *et al.* Relationship Among Total Kidney Volume, Renal Function and Age. *The Journal of Urology* **187**, 344–349. issn: 0022-5347 (Jan. 2012).
67. Zöllner, F. G. *et al.* Assessment of Kidney Volumes From MRI: Acquisition and Segmentation Techniques. *American Journal of Roentgenology* **199**, 1060–1069. issn: 0361-808X (1st Nov. 2012).
68. Seuss, H. *et al.* Development and Evaluation of a Semi-Automated Segmentation Tool and a Modified Ellipsoid Formula for Volumetric Analysis of the Kidney in Non-Contrast T2-Weighted MR Images. *Journal of Digital Imaging* **30**, 244–254. issn: 1618-727X (1st Apr. 2017).
69. Wachinger, C., Reuter, M. & Klein, T. DeepNAT: Deep Convolutional Neural Network for Segmenting Neuroanatomy. *NeuroImage. Segmenting the Brain* **170**, 434–445. issn: 1053-8119 (15th Apr. 2018).

70. Lu, F., Wu, F., Hu, P., Peng, Z. & Kong, D. Automatic 3D Liver Location and Segmentation via Convolutional Neural Network and Graph Cut. *International Journal of Computer Assisted Radiology and Surgery* **12**, 171–182. issn: 1861-6429 (1st Feb. 2017).
71. Mahapatra, D. & Buhmann, J. M. Visual Saliency-Based Active Learning for Prostate Magnetic Resonance Imaging Segmentation. *Journal of Medical Imaging* **3**, 014003. issn: 2329-4302, 2329-4310 (Feb. 2016).
72. Chen, J., Yang, L., Zhang, Y., Alber, M. & Chen, D. Z. in *Advances in Neural Information Processing Systems* 29 (eds Lee, D. D., Sugiyama, M., Luxburg, U. V., Guyon, I. & Garnett, R.) 3036–3044 (Curran Associates, Inc., 2016).
73. Stollenga, M. F., Byeon, W., Liwicki, M. & Schmidhuber, J. in *Advances in Neural Information Processing Systems* 28 (eds Cortes, C., Lawrence, N. D., Lee, D. D., Sugiyama, M. & Garnett, R.) 2998–3006 (Curran Associates, Inc., 2015).
74. Self, M. W., van Kerkoerle, T., Goebel, R. & Roelfsema, P. R. Benchmarking Laminar fMRI: Neuronal Spiking and Synaptic Activity during Top-down and Bottom-up Processing in the Different Layers of Cortex. *NeuroImage* **197**, 806–817. issn: 1053-8119 (15th Aug. 2019).
75. Muckli, L. et al. Contextual Feedback to Superficial Layers of V1. *Current Biology* **25**, 2690–2695. issn: 0960-9822 (19th Oct. 2015).
76. Waehnert, M. D. et al. Anatomically Motivated Modeling of Cortical Laminae. *NeuroImage. In-Vivo Brodmann Mapping of the Human Brain* **93**, 210–220. issn: 1053-8119 (1st June 2014).
77. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage* **9**, 179–194. issn: 1053-8119 (1st Feb. 1999).
78. Piskunowicz, M. et al. A New Technique with High Reproducibility to Estimate Renal Oxygenation Using BOLD-MRI in Chronic Kidney Disease. *Magnetic Resonance Imaging* **33**, 253–261. issn: 0730-725X (1st Apr. 2015).
79. Milani, B. et al. Reduction of Cortical Oxygenation in Chronic Kidney Disease: Evidence Obtained with a New Analysis Method of Blood Oxygenation Level-Dependent Magnetic Resonance Imaging. *Nephrology Dialysis Transplantation* **32**, 2097–2105. issn: 0931-0509 (1st Dec. 2017).

80. Liu, Z. *et al.* Chronic Kidney Disease: Pathological and Functional Assessment with Diffusion Tensor Imaging at 3T MR. *European Radiology* **25**, 652–660. issn: 0938-7994, 1432-1084 (1st Mar. 2015).
81. Delgado, J. *et al.* Pilot Study on Renal Magnetic Resonance Diffusion Tensor Imaging: Are Quantitative Diffusion Tensor Imaging Values Useful in the Evaluation of Children with Ureteropelvic Junction Obstruction? *Pediatric Radiology* **49**, 175–186. issn: 1432-1998 (1st Feb. 2019).
82. Andersson, J. L. R., Skare, S. & Ashburner, J. How to Correct Susceptibility Distortions in Spin-Echo Echo-Planar Images: Application to Diffusion Tensor Imaging. *NeuroImage* **20**, 870–888. issn: 1053-8119 (1st Oct. 2003).
83. Smith, S. M. *et al.* Advances in Functional and Structural MR Image Analysis and Implementation as FSL. *NeuroImage. Mathematics in Brain Imaging* **23**, S208–S219. issn: 1053-8119 (1st Jan. 2004).
84. Andersson, J. L. R. & Sotiroopoulos, S. N. An Integrated Approach to Correction for Off-Resonance Effects and Subject Movement in Diffusion MR Imaging. *NeuroImage* **125**, 1063–1078. issn: 1053-8119 (15th Jan. 2016).
85. Garyfallidis, E. *et al.* Dipy, a Library for the Analysis of Diffusion MRI Data. *Frontiers in Neuroinformatics* **8**. issn: 1662-5196 (2014).
86. Hattori, K. *et al.* Development of MRI Phantom Equivalent to Human Tissues for 3.0-T MRI. *Medical Physics* **40**. issn: 2473-4209 (1st Mar. 2013).
87. Berry, D. B. *et al.* A 3D Tissue-Printing Approach for Validation of Diffusion Tensor Imaging in Skeletal Muscle. *Tissue Engineering Part A* **23**, 980–988. issn: 1937-3341. pmid: 28338417 (17th Feb. 2017).
88. Perrin, M. *et al.* Validation of Q-Ball Imaging with a Diffusion Fibre-Crossing Phantom on a Clinical Scanner. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**, 881–891. issn: 0962-8436, 1471-2970. pmid: 16087433 (29th May 2005).
89. Canstein, C. *et al.* 3D MR Flow Analysis in Realistic Rapid-Prototyping Model Systems of the Thoracic Aorta: Comparison with in Vivo Data and Computational Fluid Dynamics in Identical Vessel Geometries. *Magnetic Resonance in Medicine* **59**, 535–546. issn: 1522-2594 (1st Mar. 2008).

90. Anderson, J. R. *et al.* *Validation of Computational Fluid Dynamics Methods with Anatomically Exact, 3D Printed MRI Phantoms and 4D pcMRI* in 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Aug. 2014), 6699–6701.

## 5.1 List of Acronyms Used

**ADC** Apparent Diffusion Coefficient

**AKI** Acute Kidney Injury

**AM** Additive Manufacturing

**ASL** Arterial Spin Labelling

**BOLD** Blood Oxygen Level Dependent

**CBF** Cerebral Blood Flow

**CFD** Computational Fluid Dynamics

**CKD** Chronic Kidney Disease

**CMRO<sub>2</sub>** Cerebral Metabolic Rate of Oxygen

**DTI** Diffusion Tensor Imaging

**DWI** Diffusion Weighted Imaging

**EPI** Echo Planar Imaging

**eTE** Effective Echo Time

**FA** Flip Angle

**FAIR** Flow-sensitive Alternating Inversion Recovery

**FCN** Fully Convolutional Network

**FFE** Fast Field Echo

**FOV** Field Of View

**FSL** fMRI Software Library

**GFR** Glomerular Filtration Rate

**GPU** Graphical Processing Unit

**GraSE** Gradient Spin Echo

**GUI** Graphical User Interface

**H and E** Haematoxylin and Eosin

**HASTE** Half-Fourier Single-shot Turbo spin Echo

**ISMRM** International Society of Magnetic Resonance in Medicine

**LSTM** Long Short-Term Memory

**ME-TSE** Multi-Echo Turbo Spin Echo

**MRI** Magnetic Resonance Imaging

**NBF** Neutral Buffered Formalin

**PBS** Phosphate-buffered Saline

**PC** Phase Contrast

**PLD** Post Label Delay

**PRELUDE** Phase Region Expanding Labeller for Unwrapping Discrete Estimates

**RBF** Renal Blood Flow

**RNN** Recursive Neural Network

**ReLU** Rectified Linear Unit

**RF** Radio Frequency

**RMRO<sub>2</sub>** Renal Metabolic Rate of Oxygen

**ROI** Region Of Interest

**SE** Spin Echo

**SENSE** SENSetivity Encoding

**SNR** Signal to Noise Ratio

**SPMIC** Sir Peter Mansfield Imaging Centre

**TE** Echo Time

**TFE** Turbo Field Echo

**TFEPI** Turbo Field Echo Planar Imaging

**TI** Inversion Time

**TILT** Transfer Insensitive Labelling Technique

**TKV** Total Kidney Volume

**TLCO** Twelve Layer Concentric Objects

**TR** Repetition Time

**True-FISP** True Fast Imaging with Steady Precession

**TRUST**  $T_2$  Relaxation Under Spin Tagging

**TSE** Turbo Spin Echo

**vNavs** Volume Navigators

**WET** Water suppression Enhanced through  $T_1$  effects