Investigating Tissue Heterogeneity using MRI in Prostate Cancer

William Devine

A dissertation submitted in partial fulfilment of
the requirements for the degree of
Doctor of Philosophy.

Centre for Medical Imaging
Division of Medicine
UCL

July 2019

I, William Devine, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Multi-parametric MRI, a promising new technique for grading prostate cancer using MRI, classifies a high number of regions as indeterminate. This is a symptom of the wider problem that clinical usage of MRI in prostate cancer only includes basic techniques and does not directly categorise tissue microstructure.

This work provides insight into the microstructure of the prostate using a combination of new tissue models and acquisition schemes. Each is tested with the aim of producing a method that is better at detecting and grading prostate cancer.

The first section utilises microstructural diffusion models to better quantify tissue heterogeneity in the prostate. The two models investigated provided more information about the heterogeneous nature of the prostate that ADC and showed significant difference between lesions and normal tissue.

The next section looks into combining multi-echo T2 (ME-T2) sequences with quantitative tissue modelling called Luminal Water Imaging (LWI). This work produced an optimal LWI fitting technique and acquisition.

Then the ability of LWI to detect the PI-RADS v2.0 score of regions of interest was examined, showing that it was able to differentiate between scores better than ADC. This work also showed that LWI can differentiate between tumour and normal tissue with an AUC of 0.81 ($p<0.05$) when compared to ADC with an AUC of 0.75 ($p<0.05$) in this dataset.

The next section further improves the acquisitions using larger datasets. It showed that correcting for imperfect pulse refocusing could improve on the performance of LWI in detecting PCa. This work also showed that fewer echoes could be used in the acquisition.

Neural networks were then used to detect and grade prostate cancer using the data points from both multiple b-value diffusion and ME-T2 decay curves. The neural network's ability to distinguish between different PIRADS scores was shown to have an AUC of 0.87 ($p<0.05$) using 32-echo data.

Impact Statement

Prostate cancer caused 11,631 deaths in the UK in 2016. Any technique which could help to diagnose and grade prostate cancer earlier and more accurately would reduce this number. The impact of this work is modest at present but is an invaluable first step. Any medical advance starts with a preliminary result. These results are important because they are designed to prove that a technique is worth investigating further, generating enough interest that the work will be taken up by other centres and make it to clinical practice. By presenting my work in conferences, and publishing a paper on the topic, the field of quantitative T2 imaging in the prostate is becoming more widely recognised as a strong alternative to other MR based models.

The benefits of the insight presented in this thesis are clear. The methodologies for acquiring MRI datasets and modelling the resultant signal decays contained in this work could be used in the future to create larger studies into their efficacy, with the ultimate aim being to non-invasively detect and grade prostate cancer accurately. This will reduce the need for invasive biopsies, which themselves have been shown to miss a significant number of tumours, and reduce the impact that prostate cancer has on the lives of millions of men across the world.

Within academia, my work into the investigation of new MRI biophysical models, the work into making these models as simple as possible whilst still providing better sensitivity and specificity that current techniques, and the optimisation of the acquisitions necessary for these models to function have all had an impact on MR imaging of the prostate. For example, it is clear through many presentations and conversations at events, seminars and conferences that my work into reducing the echo train length needed to quantitatively model the T2 spectra of prostate tissue has shown others that this type of technique is not only clinically feasible but effective in detecting prostate cancer. This work into quantitative T2 modelling was strong enough that the technique is now included in a large grant designed to further investigate its potential.

Future work is already building upon the results contained within this thesis to create a methodology that ultimately aims to replace invasive methods such as biopsy as the go to method for diagnosing and staging prostate cancer. What is needed is a large repeatability study into quantitative T2 modelling which will give an insight as to whether it would work across multiple centres. This would be the next step on the road towards clinical acceptance.

Contents

List of Figures 9

List of Tables 13

1 Introduction 15
- 1.1 Thesis Structure . . . 17
- 1.2 Publications . . . 18
  - 1.2.1 Papers . . . 18
  - 1.2.2 Conference Abstracts . . . 18

2 Background 21
- 2.1 The Prostate . . . 21
  - 2.1.1 Anatomy . . . 21
  - 2.1.2 Prostate Cancer . . . 23
  - 2.1.3 Diagnosis . . . 24
  - 2.1.4 Histology . . . 26
  - 2.1.5 Grading . . . 27
  - 2.1.6 Multi-Parametric MRI . . . 29
- 2.2 Nuclear Magnetic Resonance . . . 31
  - 2.2.1 Coils . . . 33
  - 2.2.2 Measuring T2* & T2 . . . 34
  - 2.2.3 Measuring T1 . . . 35
- 2.3 Pulse Sequences . . . 36
  - 2.3.1 Spin-Echo . . . 36
  - 2.3.2 Pulsed Gradient Spin Echo . . . 40
  - 2.3.3 Multi-Echo Spin Echo . . . 41
  - 2.3.4 Fast Spin Echo . . . 42
- 2.4 Modelling . . . 42
  - 2.4.1 Diffusion modelling . . . 43
    - 2.4.1.1 ADC . . . 43
    - 2.4.1.2 Diffusion Tensor Imaging . . . 46
    - 2.4.1.3 Diffusion Kurtosis Imaging . . . 47
    - 2.4.1.4 Stretched Exponential Model . . . 48
    - 2.4.1.5 Statistical model . . . 49
    - 2.4.1.6 Multi-exponential DWI . . . 51
    - 2.4.1.7 Spherical Mean Technique . . . 52
    - 2.4.1.8 VERDICT . . . 52

2.4.2 T2 modelling . . . 54
2.4.2.1 Traditional T2-weighted Imaging . . . 54
2.4.2.2 Multi-exponential T2 . . . 55
2.4.2.3 T2 Distribution . . . 56
2.4.2.4 mcDESPOT . . . 58
2.5 Image unwarping . . . 59
2.6 Machine learning . . . 60
2.7 Summary . . . 65

3 Diffusion Modelling 67
3.1 Distributed diffusivity model . . . 68
3.1.1 Methods . . . 69
3.1.2 Results and Discussion . . . 72
3.2 Spherical Mean Technique . . . 73
3.2.1 Background . . . 74
3.2.2 Methods . . . 76
3.2.3 Results . . . 77
3.2.4 Discussion . . . 78
3.3 Summary . . . 79

4 T2 Modelling 81
4.1 Model Selection . . . 82
4.1.1 Comparing constrained and unconstrained fitting . . . 86
4.1.1.1 Simulation . . . 87
4.1.1.2 Phantom . . . 89
4.2 Sequence Design . . . 90
4.2.1 Type of sequence . . . 91
4.2.2 Parameters of sequence . . . 91
4.2.2.1 Simulation . . . 92
4.2.2.2 Phantom Study . . . 93
4.3 Preliminary comparison of T2 and diffusion modelling . . . 97
4.3.1 Methods . . . 97
4.3.2 Results . . . 98
4.3.3 Discussion and Conclusion . . . 99
4.4 Differentiating tumour from normal tissue . . . 100
4.5 Summary . . . 102

5 Simplified Luminal Water Imaging for the Detection of Prostate Cancer from Multi-Echo T2 MR Images 103
5.1 Abstract . . . 103
5.2 Introduction . . . 104
5.3 Materials and Methods . . . 107
5.3.1 Simulation . . . 108
5.3.2 Patient Selection . . . 108
5.3.3 MRI Acquisitions . . . 109
5.3.4 Regions of Interest and histologic examination . . . 109
5.3.5 Data processing . . . 110

5.3.6 Statistics . . . 112
5.4 Results . . . 113
5.5 Discussion . . . 116

6 Optimisation for Clinical Use 119
6.1 Sequence optimisation . . . 120
6.2 Imperfect pulse refocusing . . . 122
6.2.1 Background . . . 123
6.2.2 Simulation . . . 125
6.2.3 In-vivo . . . 127
6.2.4 Discussion . . . 128
6.3 Summary . . . 129

7 Machine Learning for Prostate Cancer Detection 135
7.1 Comparison of Machine Learning Techniques . . . 135
7.1.1 Methods . . . 136
7.1.2 Results & Conclusion . . . 137
7.2 Neural networks for scoring of tumours . . . 137
7.2.1 Introduction . . . 138
7.2.1.1 Patient Selection . . . 139
7.2.2 MRI Acquisitions . . . 139
7.2.2.1 Regions of Interest and histologic examination . . . 140
7.2.2.2 Image Transformation . . . 140
7.2.2.3 Neural networks . . . 141
7.2.2.4 Statistics . . . 142
7.2.3 Results . . . 143
7.2.4 Discussion . . . 143
7.3 Summary . . . 147

8 Conclusion and Future Work 149
8.1 Conclusion . . . 149
8.2 Contributions . . . 151
8.3 Future Work . . . 152

A Imaging Biomarker Roadmap 155

List of Figures

2.1 Cross section of human body, including prostate (reproduced with permission from Cancer Research UK). . . . 22
2.2 Cross section of prostate (reproduced with permission from the Canadian Cancer Society[28]). . . . 23
2.3 Example of TRUS guided biopsy with (reproduced with permission from Cancer Research UK). . . . 25
2.4 Ultrasound image of prostate highlighting darker tumour region with an arrow (reproduced from Harvey et al. [68] with permission). . . . 26
2.5 Diagram of Gleason Scoring system (reproduced from Humphrey et al. [71] with permission). . . . 28
2.6 Histology from prostate biopsy, graded Gleason 3+5 adenocarcinoma. The separate areas of Gleason 3, with the lighter colour and better differentiated glands, and Gleason 5, with a darker colour and little glandular space, can be clearly seen (reproduced from Humphrey et al. [71] with permission). . . . 29
2.7 Free induction decay of an excited hydrogen nucleus. . . . 34
2.8 Free induction decay of an excited hydrogen nucleus. . . . 34
2.9 Typical spin-echo sequence. . . . 37
2.10 Gradient application in a typical spin-echo sequence. . . . 38
2.11 K-space diagram for a typical spin-echo sequence. . . . 38
2.12 Typical PGSE sequence. . . . 40
2.13 RF pulses and Signal respectively for a MESE sequence, showing T2 and T2* signal decays. . . . 42
2.14 First two echoes of a typical FSE sequence. . . . 43
2.15 Example models for diffusion data. . . . 50
2.16 VERDICT tissue model (reproduced from Panagiotaki et al. [120] with permission). . . . 53
2.17 T2 spectrum using NNLS fitting (reproduced from Mackay et al. [101] with permission). . . . 56
2.18 Diagram of example neural network with two hidden layers. . . . 61
2.19 Diagram of example neural network with two hidden layers. . . . 62

3.1 A graph showing the Gaussian diffusion distribution along the b=0 axis. The overall signal is the sum of all of the exponentials, as explained in Equation 3.1. . . . 70
3.2 Comparison of mono-exponential and distribution models in a single voxel. . . . 70

3.3 Maps of ADC, Mean Diffusivity (D_m)of Distribution and Standard Deviation (σ) of Distribution for all three patients. The tumour is outlined in red. . . . 71
3.4 Maps of T2w, DWI, ADC, Mean Diffusivity of Distribution (D_m) and Standard Deviation of Distribution (σ) for one patient. The tumour is outlined in red. . . . 71
3.5 Zeppelin response function, with the parallel and perpendicular diffusivities ($\lambda_{\parallel}$ and $\lambda_{\perp}$) labelled. The zeppelin shape is axially symmetric which reduces the number of parameters when compared to a normal tensor. . . . 75
3.6 T2w, ADC, b=90 and b=2000 images for an example PZ tumour. . . . 77
3.7 ADC and FA maps of an example healthy prostate. . . . 78

4.1 Example of the three components of multi-parametric MRI, axial T2, apparent diffusion coefficient (ADC) and dynamic contrast enhanced (DCE) images. . . . 81
4.2 Simple diagrams representing a bi-exponential, NNLS and constrained two-Gaussian fit respectively for a two-compartment tissue. . . . 83
4.3 Images of phantom experiment. . . . 94
4.4 Maps of LWF and axial T2 for two example tumours. The tumours are outlined in red in the first row. . . . 99

5.1 Diagram of prostate histology in A) Benign tissue B) Malignant adenocarcinoma. The four separate micro-environments present in the prostate are shown in the key. Using LWI, the short-T2 component is made up of the stroma and epithelia and the long-T2 component is made up of the lumen (the T2 of the vasculature is not considered to have a significant effect on the LWI model). Using the VERDICT model, the intra-cellular compartment is made up of epithelia, the extracellular-extravascular compartment is made up of stroma and lumen and the vascular compartment is made up of the vasculature. . . . 105
5.2 LWF map, axial T2 image and ADC map for one patient. The region of healthy tissue is highlighted by the arrow on the left of each image, the tumour by the arrow on the right of each image. This figure shows distortions in the PZ in the ADC image, highlighting a disadvantage of DWI over ME-T2 modelling. . . . 116
5.3 T2 distributions for an example pixel in a) normal tissue and b) tumour. These pixels are taken from the subject in Figure 5.2 in the regions highlighted by the arrows. . . . 116
5.4 Bland-Altman and correlation plots for median LWF values of ROIs from two separate readers. . . . 116

6.1 ROC curves for 32-echo (blue), 16-echo(green), 8-echo(red)and 6-echo (orange) datasets. . . . 121
6.2 T2 signal and LWI fitting for pixel of benign tissue. The images on the left are the fitted T2 distribution of the pixel and the image on the right is the T2 signal decay alongside the model fit. . . . 131
6.3 T2 signal and LWI fitting for pixel of cancerous tissue.The images on the left are the fitted T2 distribution of the pixel and the image on the right is the T2 signal decay alongside the model fit. . . . 132

6.4 First eight echoes of the signal decay of the myelin component for three refocusing flip angles. All of the magnitudes are normalized to the value of the first echo for the 180° case (reproduced from Prasloski et al. [124] with permission). 133
6.5 A graph showing the first five pulses of a CPMG pulse train with perfect 180° pulse refocusing, including the phase graph and the resultant signal showing the five corresponding echoes. 133
6.6 A graph showing the first three pulses of a CPMG pulse train with imperfect pulse refocusing, including the phase graph and the resultant signal. 134

7.1 All signal decays for test set plotted against time. Red are the pixels predicted to be PI-RADS 3,4&5 and blue are the pixels predicted to be PI-RADS 1 & 2. 144
7.2 This plot depicts those pixels that were predicted to be normal tissue. The darker lines are pixels that were predicted correctly and the lighter lines those for which the neural network prediction was wrong. 144
7.3 This plot depicts those pixels that were predicted to be tumour. The darker lines are pixels that were predicted correctly and the lighter lines those for which the neural network prediction was wrong. 145
7.4 (a) This is a T2 weighted image of a prostate containing a large tumour in the left PZ. (b) The final neural network, using the 32-echo data alone, has predicted which pixels are normal tissue and which are tumour (tumour = yellow, normal = blue). The algorithm is trained on PZ tumours so although it is predicting the TZ to be mostly tumour this is misleading because the algorithm has not seen any TZ data. The PZ is highlighted in red. 145

A.1 The imaging biomarker roadmap for cancer studies (reproduced from O'Connor et al. [115] under a creative commons license.) 156

List of Tables

2.1 Spin-echo tissue weightings. 37
2.2 Summary of recent works correlating Gleason Grade and ability to detect tumour with ADC in the peripheral zone of the prostate using a 3T scanner. It is noted that some methodologies may differ. 44

3.1 Mean values and the range of values of ADC, D_m and σ for both tumour and normal tissue 71
3.2 AUC values when differentiating between three different PI-RADS v2 groupings. 77
3.3 Sensitivity and specificity values (as percentages) when differentiating between three different PI-RADS v2 groupings. 78
3.4 Sensitivity, specificity and AUC values when differentiating between normal tissue and tumour as defined by histological analysis. 78

4.1 LWF estimates using a two-delta tissue model. These are mean values across varied values of μ_1 and μ_2. 88
4.2 LWF estimates using a two-Gaussian tissue model. These are mean values across varied values of μ_1 and μ_2. 88
4.3 LWF values for each fitting method using 64-echo sequences with various TE values (Ground truth LWF = 0.2). 90
4.4 Mean LWF values for simulation using 32-echo scan. 93
4.5 Mean LWF values for simulation using 64-echo scan. 93
4.6 Results using 32-echo sequence. 95
4.7 Results using 64-echo sequence. 95
4.8 Scan parameters of multi-echo T2 and diffusion sequences. 97
4.9 Correlation and p-values comparing multi-compartment T2 and VERDICT models. 99
4.10 Mean values and range for model parameters. 101

5.1 Mean estimated LWF values for both the constrained and unconstrained models using different ground truth LWF values in simulation. These are the mean values using both the delta and Gaussian ground truth models over a range of μ_1 and μ_2 values. 115
5.2 Three separate analyses using LWF and ADC to predict PI-RADS v2 scores. The p-value comes from the logistic regression model and the other statistics are from an ROC analysis. 115
5.3 ROC analysis of LWF in detecting malignant lesions (Gleason 3+3 and above). 115

5.4 p-values of Kruskal-Wallis tests between ADC and LWF for the ROC analyses performed on each of the four score groupings tested. A p-value of 0.013 means that the null hypothesis, that the ADC and LWF predictions have the same AUC values, can be rejected with 95% confidence. 115
5.5 Correlation between LWI parameters and the intra-cellular fraction (fIC), extracellular-extravascular fraction (fEES) and vascular fraction (fVASC) parameters of the VERDICT diffusion model. 115

6.1 LWF values for simulation using 32 and 64 echo scan data respectively for an actual flip angle of 180°. 125
6.2 LWF values for simulation using 32 and 64 echo scan data respectively for an actual flip angle of 160°. 126
6.3 A comparison of the ability of LWI parameter values produced using the EPG algorithm to differentiate between PI-RADS v2 score groupings.The p-values are from a logistic regression and the AUC, sensitivity and specificity values are from an ROC analysis. 128

7.1 The average results of the cross-validated ROC analyses along with results from other studies. 137
7.2 For each of the 5 datasets, presented in the table are the average results of the cross-validated ROC analyses alongside the number of layers and nodes in the neural network that produced these values. 143

Chapter 1

Introduction

Prostate cancer is becoming more of a problem as life expectancy increases. With the incidence rate in the UK expected to rise by 12% by 2035 [149] it is important that a reliable, non-invasive method is found to detect, diagnose and stage prostate cancer.

Magnetic Resonance Imaging (MRI) provides a high level of detail in comparision to other imaging modalities. If there is a suspicion of prostate cancer in a patient then the imaging technique typically used to assess it is multi-parametric MRI (mp-MRI), a combination of diffusion-weighted, T2-weighted and Dynamic Contrast Enhanced images. mp-MRI provides sensitivity and specificity of between 87-93% and 41-47% respectively on a 1.5 T scanner [6]. However, the relatively low specificity leads to approximately 40% of lesions being classed as indeterminate in that study, whereby a radiologist cannot decide whether they are tumour or normal tissue when the images are inconclusive.

The aim of this work is to improve upon current MRI techniques in order to better diagnose and grade prostate cancer. One of the problems with current MRI techniques is that the information gathered is usually subject to a very simple model. As an example, the Apparent Diffusion Coefficient models the signal decay as a single exponential and hence does not utilise all of the data in the signal. This problem is currently being tackled in the prostate in the form of more intricate diffusion models but none have successfully impacted patient care because ADC and T2-weighted images are still the only ones used in the hospitals that use MR. The difficult trade-off between a model

that can relate directly to tissue microstructure and a dataset that can be acquired in a clinically feasible amount of time has proven difficult to solve.

The methods described in this thesis seek to improve on the sensitivity and specificity of mp-MRI for both the detection and grading of tumours. There are two main directions in which this is addressed in this work, one being the use of novel acquisitions designed to investigate the tissue microstructure and the other being models that can relate the signal to that microstructure without being overly complex.

The 'Imaging Biomarker Roadmap for Cancer Studies' produced by O'Connor et al. [115] lays out a framework for conducting research into imaging biomarkers, shown in Appendix A. It lays out two main translational gaps that need to be bridged to take a technique from an idea to a clinically accepted technique. The first translational gap looks to prove that the technique should be rolled out into multiple research centres and the second looks to use multi-centre trials to push it into routine clinical use.

This work is centred around taking novel techniques and performing the initial validation on small sample sizes needed to warrant larger, multi-centre trials, which is the first translational gap in the biomarker roadmap. This consists of technical validation, clinical validation and an initial evaluation of cost effectiveness. Of these, this work seeks to first and foremost provide clinical validation of the method. This is due to the fact that a positive clinical validation better motivates further work into technical validation such as repeatibility.

Clinical validation first needs a suitable dataset and then seeks to use that dataset to show that the method is fit for purpose by first evaluating its sensitivity and specificity for the task at hand. These two first steps are the basis of the majority of this thesis. Hence it is the aim of this work not to take these techniques all the way to the clinic but to provide the initial evidence of sensitivity and specificity for detecting and grading prostate tumours needed to generate the motivation to carry out larger, multi-centre studies.

Our aim is to investigate models for both diffision and T2 weighted images of the prostate that relate directly to its constituent tissues. Investigating acquisitions and models that

look into better modelling microstructure should provide the increased sensitivity and specificity needed if MR methods are to be able to replace the invasive, painful methods of diagnosis that patients are currently subjected to.

Problem statement: Investigation of non-invasive imaging techniques with sufficient sensitivity and specificity in detecting and grading prostate tumours for clinical use. The ways in which this project seeks to achieve this are to:

- Investigate novel diffusion and T2 based biophysical models that solve the unmet issues of current clinical and research techniques
- Create an optimal dataset on which the models can be tested
- Optimise the fitting of the models alongside the acquisitions in order to create the best possible combination of the two
- Provide clinical validation into the sensitivity and specificity of these techniques in a significant number of patients

1.1 Thesis Structure

Chapter 2 provides a detailed summary of the background material on prostate cancer, Nuclear Magnetic Resonance, diffusion and T2 Magnetic Resonance Imaging and diffusion and T2 modelling.

Chapter 3 presents two techniques for modelling diffusion heterogeneity in the prostate and carries out preliminary investigations into their ability to detect prostate cancer.

Chapter 4 chooses optimal acquisitions and models in order to perform Luminal Water Imaging (LWI), a novel technique for calculating the fractional volume of glandular lumen in prostate tissue by quantitatively modelling multiple-echo T2-weighted images.

Chapter 5 then clinically validates this LWI method in a larger number of patients. It looks into how well it can detect and grade prostate cancer and also tries to correlate

the model back to tissue microstructure. This chapter also contains an analysis into the repeatibility of this method.

Chapter 6 uses this clinical dataset to optimise the acquisition further and look into whether imperfect pulse refocusing could be affecting the ability of the LWI model to detect and grade prostate cancer.

Chapter 7 seeks to utilise a machine learning technique on multi-echo T2 data in order to carry out the same classification tasks as LWI without applying a biophysical model to the data.

1.2 Publications

1.2.1 Papers

Simplified Luminal Water Imaging for the Detection of Prostate Cancer from Multi-Echo T2 MR Images

Devine W., Giganti F., Johnston E., Punwani S., Alexander D.C. and Atkinson D. (published online December 2018, awaiting print publication) Journal of Magnetic Resonance Imaging.

1.2.2 Conference Abstracts

Optimisation of luminal water imaging for classification of prostate cancer

Gong F., Devine W., Giganti F., Johnston E., Atkinson D., and Punwani S. (Accepted January 2019) 27th meeting of the International Society for Magnetic Resonance in Medicine, Toronto, Canada (2019).

Detecting prostate cancer in multi-echo T2 images with neural networks

Devine W., Giganti F., Johnston E., Sidhu H.S., Panagiotaki E., Punwani S., Alexander D.C. and Atkinson D. (2018). 26th Post-graduate Symposium of the British Chapter of

the International Society for Magnetic Resonance in Medicine, Imperial College London, London, UK (2018).

Two compartment fitting for Luminal Water Imaging: multi-echo T2 in Prostate Cancer

Devine W., Giganti F., Johnston E., Panagiotaki E., Punwani S., Alexander D.C. and Atkinson D. (2018) Proceedings of the 26th meeting of the International Society for Magnetic Resonance in Medicine, Paris, France (2018).

Multi-echo T2 modelling to predict PIRADS 2.0 score

Devine W., Giganti F., Johnston E., Panagiotaki E., Punwani S., Alexander D.C. and Atkinson D. (2018) Proceedings of the 26th meeting of the International Society for Magnetic Resonance in Medicine, Paris, France (2018).

Linking a multi-compartment T2 model to diffusion microstructure in prostate cancer

Devine W., Johnston E., Bonet-Carne E., Punwani S., Alexander D.C., and Atkinson D. (2017). Proceedings of the 25th meeting of the International Society for Magnetic Resonance in Medicine, Hawaii, USA (2017).

Why is the Peripheral Zone of the Normal Human Prostate High in ADC Value and T2-Weighted Signal Intensity?

Johnston, E., Bailey, C., Bonet-Carne, E., Pye, H; Heavey, S., Patel, D., Sridhar, A., Siow, B., Mertzanidou, T., Devine, W., Kalasthry, J., Clemente, J., Hawkes, D., Whitaker, H., Rodriguez-Justo, M., Shaw, G., Alexander, D., Freeman, A., Bourne, R., Panagiotaki, E., Punwani, S. (2017). Proceedings of the 25th meeting of the International Society for Magnetic Resonance in Medicine, Hawaii, USA (2017).

Microstructural heterogeneity in normal and cancerous prostate using Gaussian diffusivity distributions

Devine W., Johnston E., Bonet-Carne E., Punwani S., Alexander D.C., and Atkinson D. (2016). 22nd Annual Scientific Meeting of the British Chapter of the International Society for Magnetic Resonance in Medicine, Leeds, UK (2016).

Chapter 2

Background

This chapter seeks to introduce the anatomy of the prostate and the pathology of prostate cancer, followed by a brief summary of Nuclear Magnetic Resonance and Magnetic Resonance Imaging. Then current tissue models are described in detail. At the end of the chapter there is a summary on the work done in this field to date and an analysis of where the opportunities lie.

2.1 The Prostate

2.1.1 Anatomy

The prostate is an integral part of the male reproductive system, secreting fluid that protects sperm as they leave the body. It also contains muscular tissue designed to help expel semen during ejaculation. As shown in Figure 2.1, the prostate is a small gland positioned below the bladder and in front of the rectum, with the urethra carrying urine from the bladder out of the body. The average size of a healthy prostate is a little larger than a walnut, weighing between 7-11 grams.

The tissue within the prostate has four major components; blood vessels, glandular space (also known as the lumen), epithelia and stroma. Large glandular spaces, which make up the lowest percentage of the prostate's overall volume, store fluid and are connected to

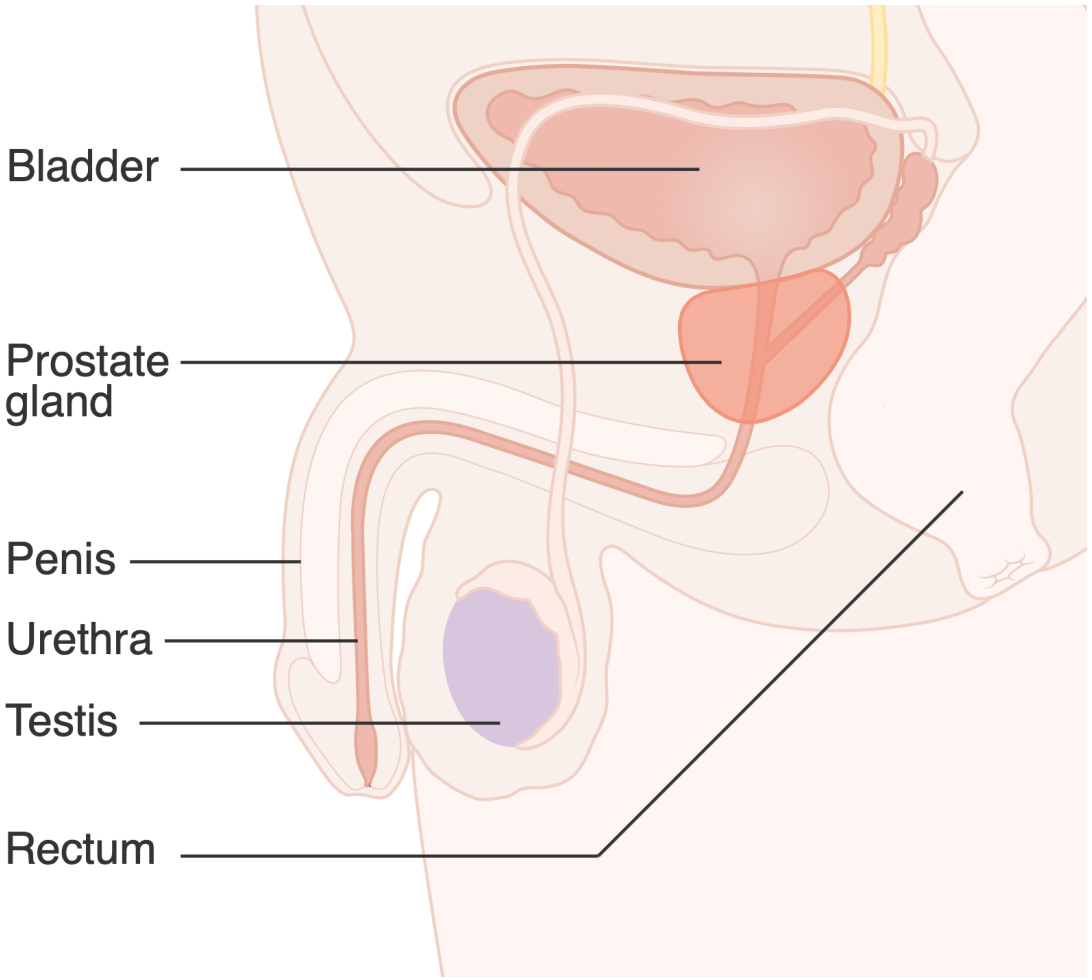


FIGURE 2.1: Cross section of human body, including prostate (reproduced with permission from Cancer Research UK).

the base of the prostate by ducts that carry the fluid to the urethra during ejaculation. These glandular spaces are surrounded by epithelial cells which produce this fluid and secrete it into the glandular space. A matrix of muscular stromal tissue links these glands together, giving structure to the prostate and forcing the fluid out of the glands during ejaculation.

On a larger scale the prostate is made up of a number of separate tissue types, or 'zones'. Each is made up of different proportions of gland, stroma and epithelia and each has its own individual purpose. Figure 2.2 shows the central, transition and peripheral zones.

The central zone makes up roughly 25% [108] of the volume of the healthy prostate. It consists of large, irregularly shaped glands, cuboidal epithelial cells and very compact stromal tissue. The ducts of the central zone meet the urethra at the verumontanum,

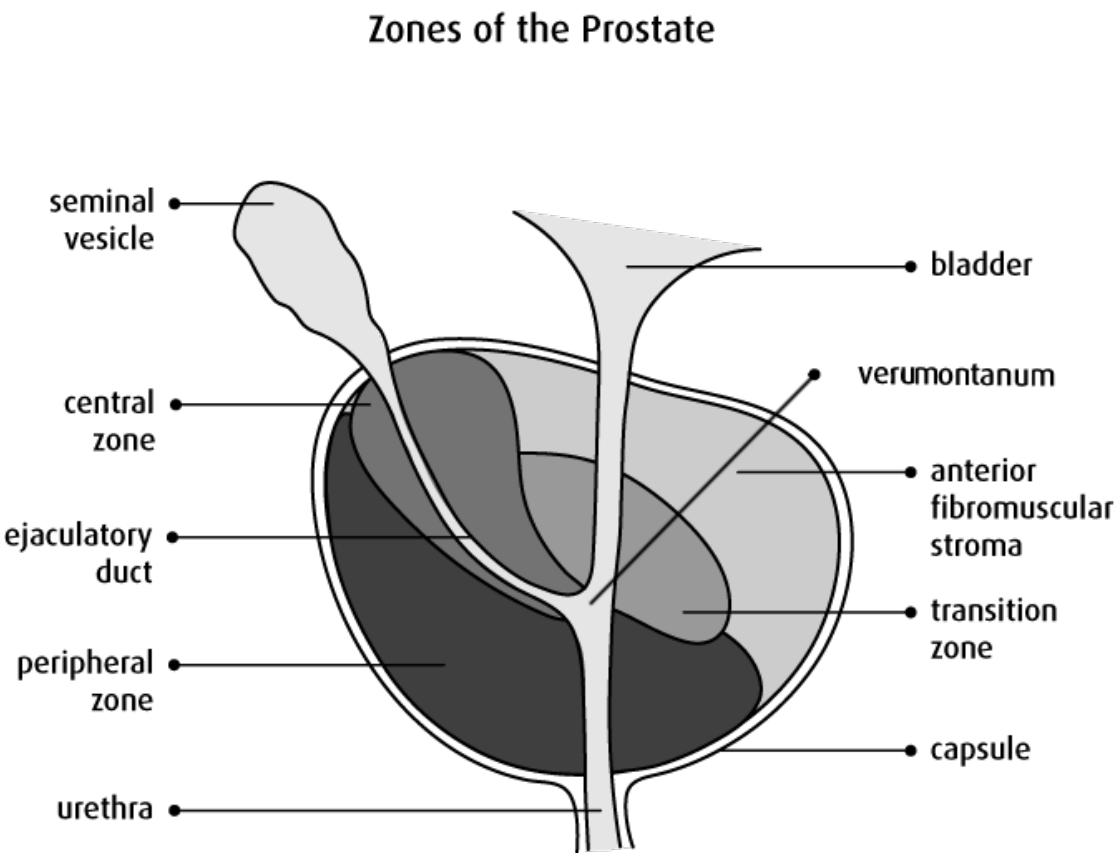


FIGURE 2.2: Cross section of prostate (reproduced with permission from the Canadian Cancer Society[28]).

the junction between the urethra and the ejaculatory ducts coming from the seminal vesicles. Most of the prostatic fluid is generated in the central zone.

The peripheral zone, by comparison, constitutes around 70% [108] of the volume of the prostate. It contains small, simple, regularly arranged glands lined by tall columnar epithelial cells and surrounded by loosely interwoven stroma.

The smallest zone of the prostate is the transition zone, making up approximately 5% [108] of the healthy prostate volume. It is composed of two lobules either side of the urethra. Histologically, it closely resembles the peripheral zone tissue, though the stroma are more compact.

2.1.2 Prostate Cancer

The risk of developing benign prostate hyperplasia (BPH) and prostate cancer increases with age. Therefore, over the past few decades, a greater life expectancy around the

world has meant that men are living to an age where the likelihood of being diagnosed with either of these diseases is significantly increased. BPH is the slow enlargement of the prostate due to a steady cell proliferation and is common in older men, with studies showing that over half of men in their 60s and 80-90% of men over 70 have BPH [128]. BPH usually originates in the transition zone and has been found to have no link to prostate cancer[31].

Prostate cancer, or carcinoma of the prostate, begins when normal semen-secreting epithelial cells mutate into cancer cells, with almost 70% of cases starting in the peripheral zone [109]. With 11,300 deaths in the UK in 2014 being attributed to adenocarcinoma in the prostate [2] it is the second most common cause of cancer related death in men, with its incidence rate predicted to rise by 12% by 2035 [149]. Across the world, in 2018 there were approximately 1.3 million new cases of prostate cancer, making it the second most sommonly occuring cancer in men [25].

2.1.3 Diagnosis

Current clinical practice when diagnosing a patient is the triple assessment. This section describes the most common clinical process, though some centres do utilise MRI prior to biopsy to aid in the detection of tumours. The triple assessment involves performing three tests in order to detect and grade PCa. These three tests are:

1. Digital Rectal Examination (DRE) - A finger is inserted into the rectum to feel for abnormalitles of the prostate. The normal prostate is smooth and symmetrical so irregularities raise the possibility of a tumour.

2. Prostate Specific Antigen (PSA) Test - PSA is a protein formed in the prostate and a low level is always present in healthy men. However a significantly raised level can be an indicator of possible malignancy.

3. Trans-Rectal Ultrasound (TRUS) guided biopsy - As Figure 2.3 shows, this involves the insertion of an ultrasound probe into the rectum until it is close to the prostate. Once the probe has located the target area of the prostate a needle is passed along

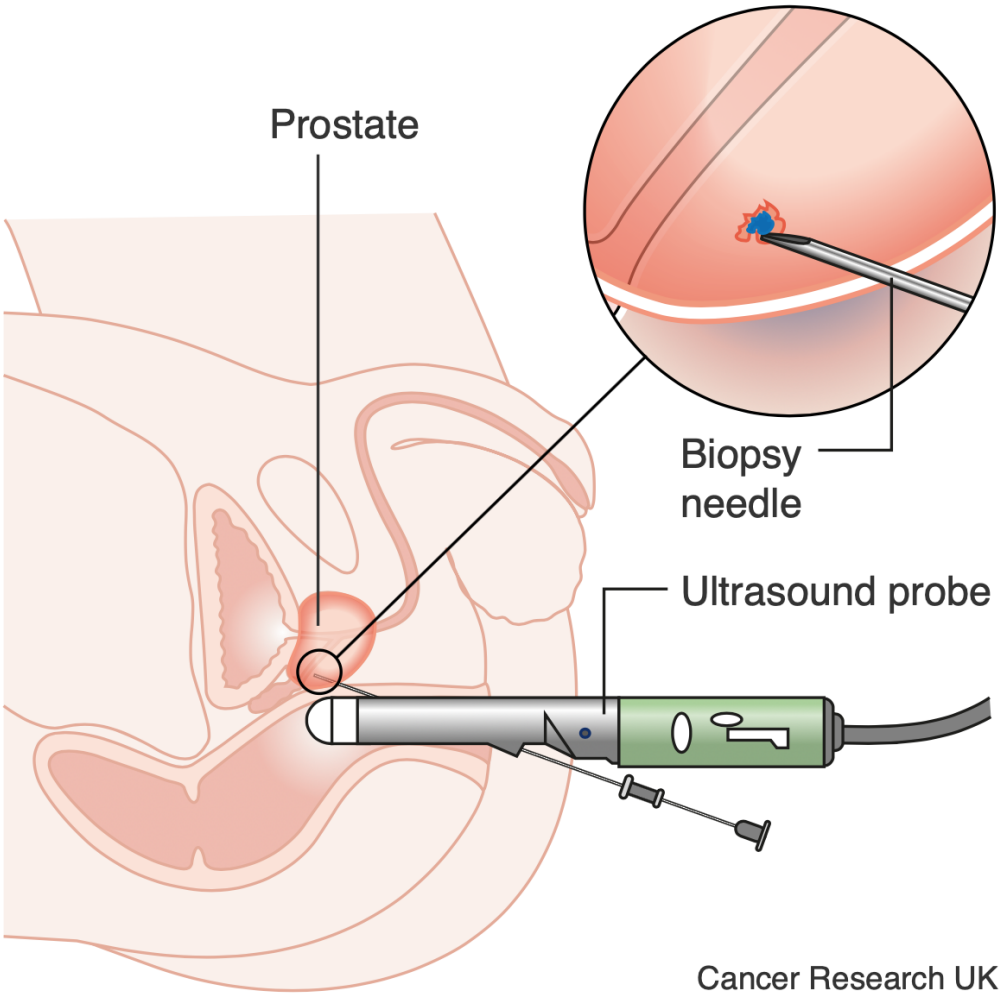


FIGURE 2.3: Example of TRUS guided biopsy with (reproduced with permission from Cancer Research UK).

the probe, through the rectum wall and into the area of interest, removing a small portion of tissue to be examined under a microscope. This process can be repeated for multiple regions of the prostate or a template biopsy of evenly spaced needles can be taken.

Though the PSA test is not highly specific, with only 30% of raised PSA levels turning out to be associated with tumour [5], it has proven to be a useful tool in initially assessing the risk without resorting to invasive measures.

Either an abnormal Digital Rectal Examination or a raised PSA level will usually lead to further investigation using the histological diagnosis of biopsied tissue. This involves the removal and visual examination of a small tissue sample under a microscope, which

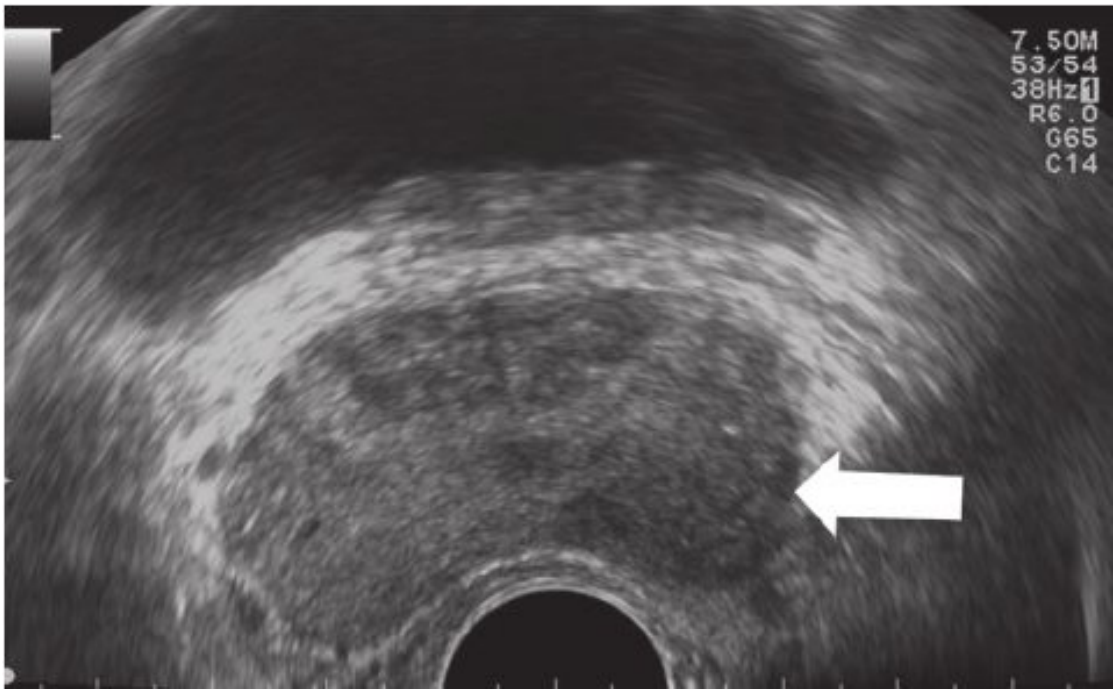


FIGURE 2.4: Ultrasound image of prostate highlighting darker tumour region with an arrow (reproduced from Harvey et al. [68] with permission).

is essential before treatment can begin. This currently takes the form of a Trans-Rectal Ultrasound (TRUS) guided needle biopsy.

Though TRUS biopsy accuracy is limited, ultrasound has still proven to have an advantage over Digital Rectal Examinations[94]. An example of an ultrasound image is shown in Figure 2.4, with the prostate in the centre and a suspected tumour visible as a dark area highlighted by the arrow. This example is specifically chosen to clearly show cancerous tissue whereas many tumours are still not visible on ultrasound.

2.1.4 Histology

Histology is the study of tissue microstructure, performed by examining cells and tissues under a microscope. Normal peripheral zone and central zone tissue have slightly different structures. Both contain an arrangement of similarly sized, evenly spaced lumen surrounded by a thin, even layer of epithelial cells and are held together by a regular stromal matrix. Carcinoma arises mainly from epithelial cells, causing them to rapidly replicate and invade into the surrounding tissue. In doing so, the epithelium causes

irregularities in the shape and arrangement of the glands within the tissue, eventually destroying them and forming one solid mass of tumour cells.

2.1.5 Grading

One of the challenges in prostate cancer is that it is exceedingly difficult to differentiate between those tumours that will lie dormant without affecting lifespan or quality of life and those aggressive forms that require radical intervention. The most common method of grading the carcinoma to gauge its progression and severity is using the Gleason Grading System, first proposed by Donald Gleason in the 1960s [59]. This is a system in which the two largest areas of prostate cancer within the tissue specimen are each given a score preformed visually by experience based on the diagrams in Figure 2.5.

Gleason pattern 3 is made up of well separated glands, though they can vary greatly in their size and shape. As we move up to grade 4 the glands become fused and are not very well formed and by grade 5 there are only sheets of cells with only occasional luminal space. To calculate the overall grade, half of the score is based on the most common tissue architecture and the other half is based on that of the second most common tissue architecture. These two numbers are then combined to produce a total score for the cancer. At present, the lowest grade used when examining needle biopsies is 3+3 due to the fact that grades lower than this are prone to poor reproducibility, poor correlation with prostatectomy and can mislead patients and clinicians into believing that they have an indolent tumour [47]. This information can then be passed on to the care team who will decide the best course of action.

Figure 2.6 shows histology in the form of adenocarcinoma tissue taken from a biopsy. This sample was graded as Gleason 3+5, the Gleason 3 tissue being the lighter tissue containing some well differentiated glandular lumen and the Gleason 5 being the tightly packed region in the centre of the upper edge of the sample.

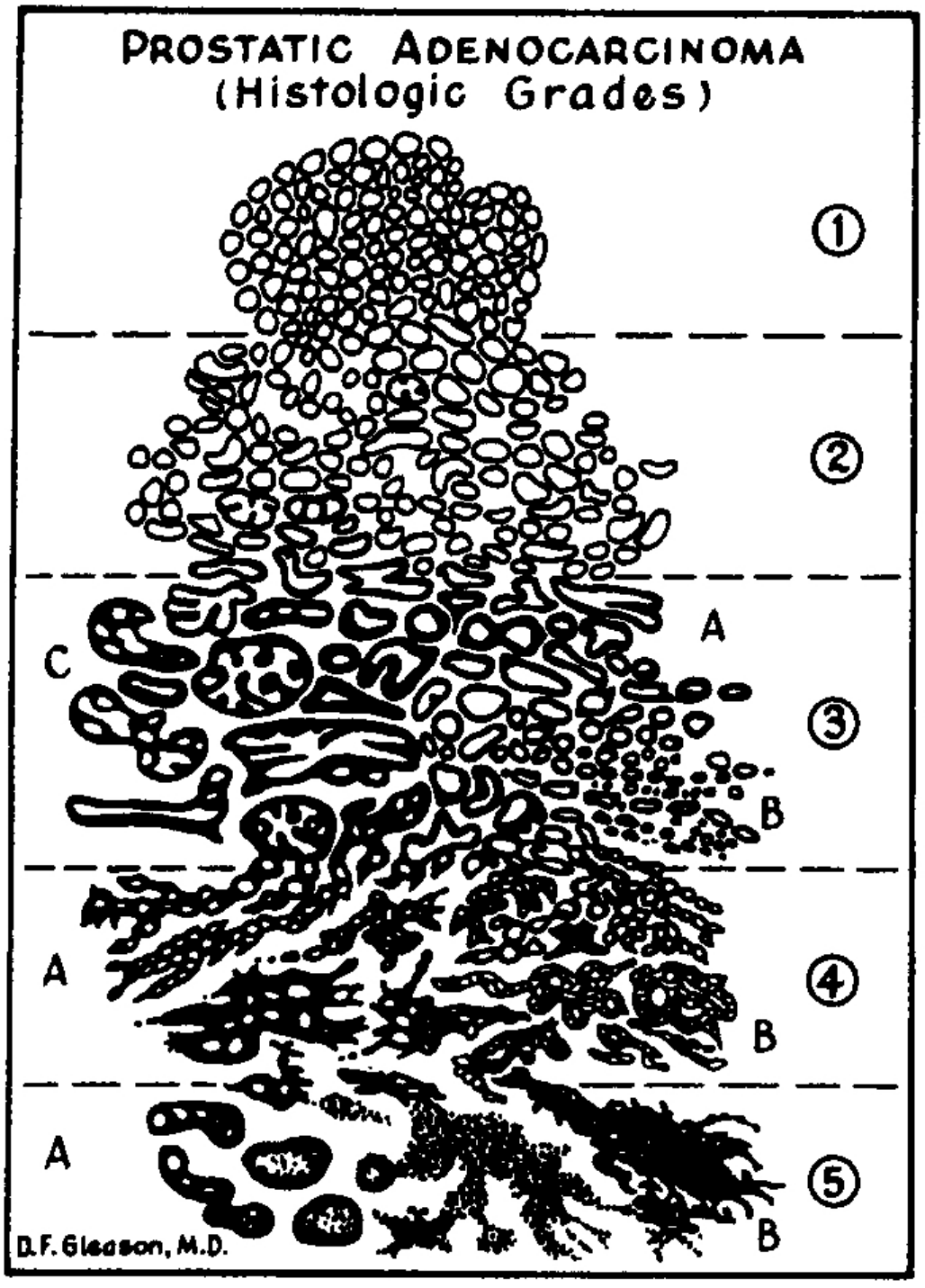


FIGURE 2.5: Diagram of Gleason Scoring system (reproduced from Humphrey et al. [71] with permission).

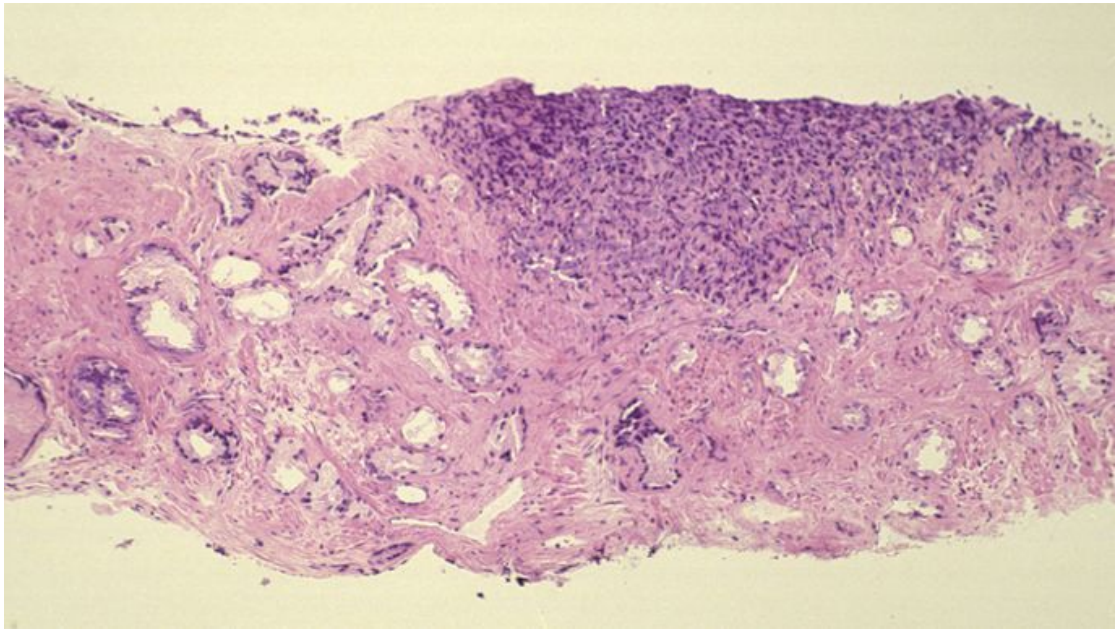


FIGURE 2.6: Histology from prostate biopsy, graded Gleason 3+5 adenocarcinoma. The separate areas of Gleason 3, with the lighter colour and better differentiated glands, and Gleason 5, with a darker colour and little glandular space, can be clearly seen (reproduced from Humphrey et al. [71] with permission).

2.1.6 Multi-Parametric MRI

Biopsy is an invasive (and in many cases uncomfortable) process and is still suspected to miss or undergrade tumours. There has been a lot of investigation into using MRI as a non-invasive method for detecting and grading prostate tumours. The most promising method to date for doing this is Multi-Parametric MRI (mp-MRI).

mp-MRI combines T2-weighted, diffusion-weighted and Dynamic Contrast Enhanced (DCE) imaging, utilising the unique information provided by each to better detect tumours in the prostate. Images of the prostate in each of the three modalities are acquired and are interpreted by a radiologist, often in combination with clinical information such as PSA and DRE results.

Both diffusion and T2 imaging are sensitive to tissue microstructure. It has been previously shown by Gibbs et al. that T2 and the Apparent Diffusion Coefficient (ADC) values (described in Section 2.4.1.1), though both related to histology and Gleason Score, do not correlate with each other [55]. This shows that the two techniques are sensitive to different microstructure. T2 imaging has proven to be sensitive to tumours in the

prostate but is much less specific than ADC. Hence it was hypothesised by Gibbs et. al. that by combining the two, along with a functional method such as DCE MRI, clinicians would be better able to localise prostate cancer.

The radiologists report the likelihood of the prostate containing areas of clinically significant prostate cancer using a method of visually scoring lesions. The simplest of these scoring methods is five-point Likert scale, where the score ranges from highly unlikely (Likert 1) to highly likely (Likert 5) [167].

The Prostate Imaging Reporting and Data System System (PI-RADS) method was first proposed by the European Society of Urogenital Radiology in 2012 [11]. The newest iteration of this scoring system, PI-RADS v2, was proposed in 2015 [167] and new sets of guidelines are periodically released [119]. It was designed as a structured reporting scheme for PCa whereby mp-MRI images are used to calculate a score out of five [151], where each of the score is:

- PI-RADS 1 - clinically significant cancer is highly unlikely to be present
- PI-RADS 2 - clinically significant cancer is unlikely to be present
- PI-RADS 3 - the presence of clinically significant cancer is equivocal
- PI-RADS 4 - clinically significant cancer is likely to be present
- PI-RADS 5 - clinically significant cancer is highly likely to be present

Many sources have found that mp-MRI has much higher sensitivity and only slightly lower specificity than the individual methods themselves, leading to a higher accuracy in detecting tumours in the prostate overall [4, 6, 46, 72, 83, 95, 148, 158, 170, 176]. One study [40] showed that Multi-Parametric MRI is particularly effective in peripheral zone tumours whilst being little better than the separate metrics in other regions. Multi-parametric MRI was also shown to be a more accurate measure of tumour volume [105] and was able to dramatically improve upon TRUS in guiding the biopsies that were performed [6]. Some studies have also suggested that DCE-MRI does not improve the diagnostic accuracy of mp-MRI [142].

However, mp-MRI combined with PI-RADS v2 also suffers from a number of issues that motivate the investigation of other imaging techniques. mp-MRI provides high sensitivity values but low specificity values in the detection of clinically significant prostate cancer [6, 148], which means that a significant number of men are suspected of having PCa when in fact they have none. Another of the problems with this method is that it classifies 35.6-44.7% of lesions as indeterminate on a 1.5T scanner [6]. This means that when classifying lesions a radiologist cannot decide whether a large proportion of them are in fact a lesion or not. Other problems with mp-MRI are the low inter-reader agreement [64, 140], the proposal that a negative mp-MRI does not guarentee the absence of significant tumour [24], and problems due to EPI-based artifacts in the diffusion images [105]. Differences in ROI drawing across multiple centres have also been highlighted as a downside of mp-MRI in a recent work [162]. Some studies have tried to resolve these issues, such as by using Prostate-Specific Antigen (PSA) levels to distinguish between indeterminate lesions [27, 127], but none has done so conculsively. Motivation for using methods that better quantify the tissue microstructure of the prostate comes from the fact that it has been shown that differences in tissue architecture can significantly affect the detection of PCa using mp-MRI [157].

2.2 Nuclear Magnetic Resonance

In quantum mechanics, 'spin' is a form of angular momentum possessed by all subatomic particles. Magnetic Resonance Imaging (MRI) is most commonly focused on hydrogen nuclei due to the fact that they are abundant in the body. The spin of this positively charged subatomic particle produces a small magnetic field. When placed in a much stronger magnetic field this particle begins to align with it. Due to the fundamentals of quantum mechanics, which will not be covered here, the protons are not able to line up exactly with the larger magnetic field and instead have a small offset from the field. Since the magnetic field is still producing a force along its axis, the axis of the proton begins to rotate about the larger axis, a phenomenon called 'precession'.

MRI is based upon the interaction of hydrogen nuclei in a magnetic field. When the hydrogen nuclei within water and fat molecules are subject to a strong magnetic field, the spin axes of the nuclei align in one of two ways, either precessing about the direction of the larger magnetic field or precessing about an axis 180° opposite to the direction of the main field. There is an almost even balance between the nuclei in either alignment but the strength of the magnetic field causes a varying disposition towards aligning in the direction of the magnetic field. This slight mismatch in the number of nuclei aligning in either direction is the source of the signal used in MRI.

When an electromangetic pulse of a specific frequency is applied to the nuclei in the magnetic field a small number will switch alignments. When the electromagnetic pulse stops, over time the nuclei revert to their original alignment but as they do they release the excess energy imparted by the electromagnetic pulse as an emitted signal which can be detected. This reveals information about the underlying tissue in the process. This section will cover the physics of Nuclear Magnetic Resonance (NMR) and how it is used to contrast between tissue types.

Through a mathematical derivation it can be shown that the angular frequency of precession, also known as the Larmor frequency, is represented by the equation $\omega_0 = \gamma B_0$ where γ is the gyromagnetic ratio (an intrinsic characteristic of the nuclei of a particular isotope) and B_0 is the strength of the magnetisation field. From this equation it follows that the larger the magnetic field the faster the nucleus will precess about the B_0 axis.

If we monitor the combined magnetisation of many hydrogen nuclei within this magnetic field in the direction parallel to the main B_0 axis, there is a net positive magnetic field. However in the plane perpendicular to the main field there is zero magnetisation as all of the precessions are out of phase and hence cancel one another out. Subject to only the B_0 field, the bulk magnetisation is exclusively in the B_0 direction.

If a Radio-Frequency (RF) pulse is applied at approximately the frequency at which the nuclei are precessing, ω_0, the bulk effect is that they begin to precess in phase and an overall magnetisation in the orthogonal plane can be detected. Therefore the bulk magnetisation vector, denoted as M_0, starts to move from the B_0 direction towards the orthogonal plane. M_0 can be made to travel through any chosen angle $\alpha = \gamma B_1 t_p$ by

setting the values of B_1 and t_p. Here α is the flip angle, B_1 is the effective magnitude of the RF pulse, and t_p is the duration of the pulse. The rates at which the excited nuclei separately dephase and return to B_0 equilibrium forms the basis of T1 and T2 imaging.

2.2.1 Coils

An MRI scanner is made up of a number of separate coils. The main magnetic field is produced by a large cylinder made up of coils that are wrapped around the bore, producing a strong linear field in the z-direction through the bore of the scanner.

Just inside the bore of the main field are the gradient coils, designed to produce a gradient in the magnetic field so the field strength varies spatially. There are three gradient coils, used to encode the x, y and z spatial axes of the scan.

Figure 2.7 shows the two coil configurations utilised in MRI scanners to attain these gradients. The first is made up of what is called a 'Maxwell Coil' where two coils at a certain distance apart produce a very linear magnetic gradient when currents are passed through them in opposite directions. As the diagram shows, the coils produce a gradient in the z-direction where the field is strongest near the coils and zero in the middle. The second setup in Figure 2.7 is called a 'Golay Coil', whereby the 'saddle-shaped' coils produce a field that varies in the y-direction. When rotated about the z-axis of the scanner 90° this configuration can also produce a gradient in the x-direction. The combination of one Maxwell Coil and two Golay Coils produce gradients in all three spatial dimensions that enable the scanner to localise a signal to a particular point in space.

The final component of the scanner is the RF coil which generates a rotating magnetic field that, when oscillating at a frequency close to that of the precession of the nuclear spins, imparts energy into the hydrogen nuclei that causes them to change their alignment. This coil, when switched off, also acts as a reciever for the signal released by the hydrogen nuclei as they return to their initial energy states. In this way the nuclei are excited by the RF coil and then their subsequent relaxation is measured using the same coil.

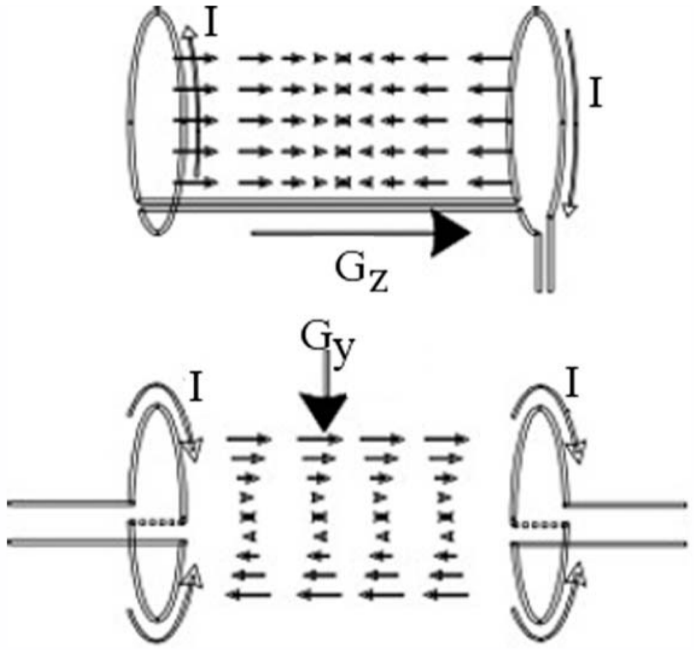


FIGURE 2.7: Free induction decay of an excited hydrogen nucleus.

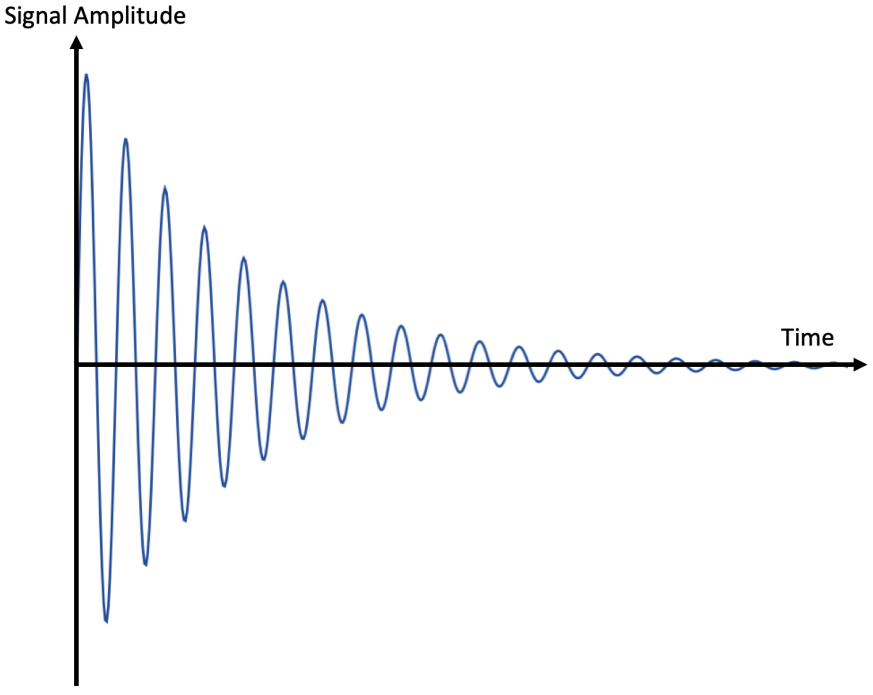


FIGURE 2.8: Free induction decay of an excited hydrogen nucleus.

2.2.2 Measuring T2* & T2

Once the bulk magnetisation vector M_0, which is the vector of net magnetisation of the precessing nuclei, has been rotated into the perpendicular plane by an $\alpha = 90°$ pulse, two relaxation processes happen simultaneously, called T1 and T2 decays. T1

decay describes the return of the bulk magnetisation vector in the z-direction to its initial value and T2 decay refers to the reduction in the component of the magnetisation vector in the perpendicular x-y plane to zero.

Measuring magnetisation in the transverse plane, the precessing nuclei should produce a signal oscillating at the Larmor frequency. However due to relaxation, the amplitude of the signal decreases exponentially with time as the protons dephase. This process is known as Free Induction Decay (FID), illustrated in Figure 2.8. The decay is described by the formula: $S = S_0 \exp(-TE/T2^*)$ where 'T2*' is the decay constant.

There are several mechanisms that cumulatively cause T2* decay: dephasing due to a magnetic gradient, dephasing due to magnetic field inhomogeneities and 'spin-spin relaxation'. Spin-spin relaxation occurs when two protons precessing at the same frequency come into close contact, momentarily producing a local magnetic field which causes both to alter their precessional frequency. As they then move apart, they regain their original frequency but the phase shift they incur during the interaction cannot be undone. The combination of these effects means that T2* decay is very fast, in the order of milliseconds.

There is a distinction between the dephasing due to magnetic gradients and field inhomogeneities and dephasing due to spin-spin relaxation. The former two are predictable and reversible processes whereas spin-spin relaxation is unpredictable and irreversible. The exponential decay resulting from only the spin-spin relaxation is an important characteristic of the tissue and is represented with the decay constant 'T2'. With the right combination of radio frequency pulses these effects can be separated.

2.2.3 Measuring T1

In the longitudinal direction, the magnetisation increases having been excited into the perpendicular direction by the pulse. This longitudinal relaxation is called 'Spin-Lattice relaxation', with the magnetisation in the B_0 direction after a 90° pulse following an exponential, rising from zero after the pulse to a steady state value M_0.

Though T1-weighted imaging is an important field within MRI, as this thesis is based upon diffusion and T2 weighted imaging it will not be fully explained here. T1-weighting is only considered insofar as how it could affect our T2 and diffusion images. As shown in Table 2.1, a long TR value weights images towards T2 rather than T1. As a rule of thumb in T2 imaging, as long as the TR of the sequence is longer than the intrinsic T1 value of the tissue being imaged then the T1 effect will be insubstantial.

2.3 Pulse Sequences

A pulse sequence is a pre-determined series of radio frequency pulses designed specifically to examine a particular phenomenon. They come in many forms and, combined with varying magnetic gradients, can probe many different features of the tissue. The purpose of changing these sequences of pulses will become apparent but in short it allows us to manipulate the image produced, for example enabling us to suppress fat signals or measure how water diffuses in a certain region.

2.3.1 Spin-Echo

In order to separate the T2* decay from the T2 decay it is necessary to select the appropriate pulse sequence. The simplest sequence is Spin-Echo. As shown in Figure 2.9, a spin echo sequence is comprised of a 90° pulse, flipping the magnetisation vector M_0 into the orthogonal plane, followed after a short time by a 180° pulse which refocuses the dephased spins to form a signal, called the spin echo. By choosing the time between the two pulses to be TE/2, we can effectively set the time after the initial pulse that we receive the spin echo as TE.

Initiating a pulse sequence with a 90° pulse, M_0 is tipped fully into the transverse plane. As the nuclei begin to relax and dephase the T2* decay causes a very fast drop-off of signal in the transverse plane. If at that point a 180° pulse is administered the axes of the nuclei flip to face the opposite direction. In doing this, the direction in which the protons precess is also flipped, meaning that the nuclei begin to rephase. The dephasing due to gradients and inhomogeneities can be reversed, as the conditions for any given

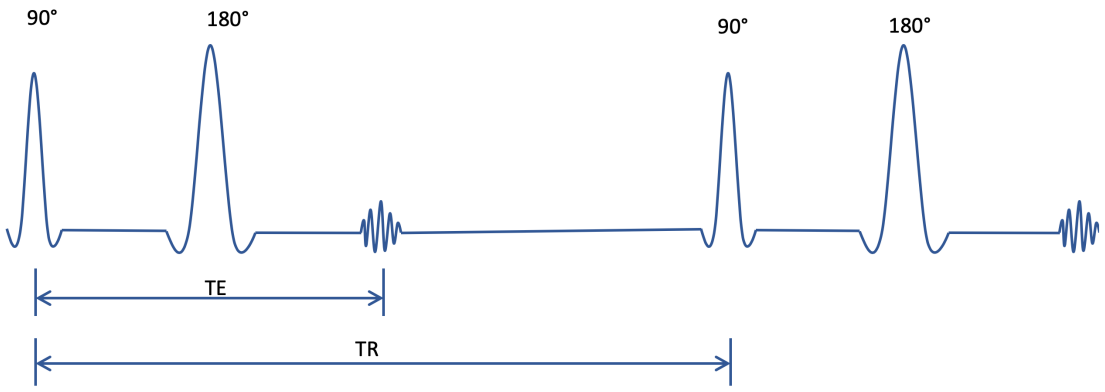


FIGURE 2.9: Typical spin-echo sequence.

	Short TR	Long TR
Short TE	T1-weighted	Proton Density weighted
Long TE	-	T2-weighted

TABLE 2.1: Spin-echo tissue weightings.

nuclei remain the same, whereas the nuclei dephased by spin-spin relaxation cannot be rephased because the localised interacting magnetic fields cannot be recreated. Motion and the diffusion can also cause reductions in the signal due to the fact that those nuclei that were excited by the initial 90° pulse will have moved relative to the gradient fields and so will not be excited properly by the subsequent 180° pulse.

Following a 180° pulse most but not all of the nuclei will rephase and after a short time the signal in the transverse magnetic field will spike. The reduction in amplitude between the original signal and the echo is only due to spin-spin relaxation so the decay between echoes is T2 and not T2* decay.

By changing the echo time, TE, and the repetition time, TR, which is the time between successive 90° pulses, the signal can be either T1-weighted or T2-weighted. Table 2.1 shows the different weightings achieved by varying TE and TR. Proton density weighting is when the image intensity is related to the density of hydrogen nuclei within the tissue.

To generate an image, magnetic gradients need to be used in the three orthogonal directions. A gradient G_{ss} encodes in the slice select direction whilst gradients G_x and G_y encode in the other two orthogonal directions.

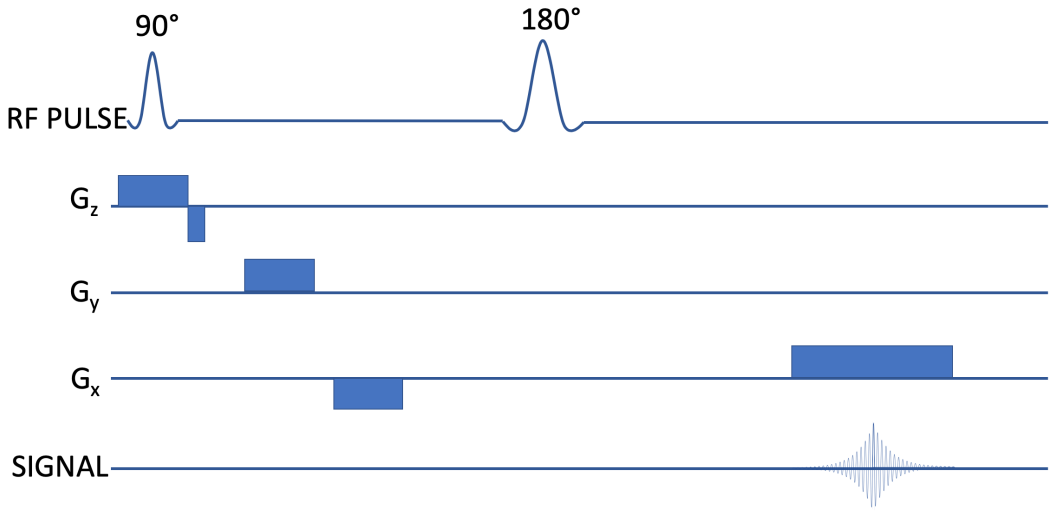


FIGURE 2.10: Gradient application in a typical spin-echo sequence.

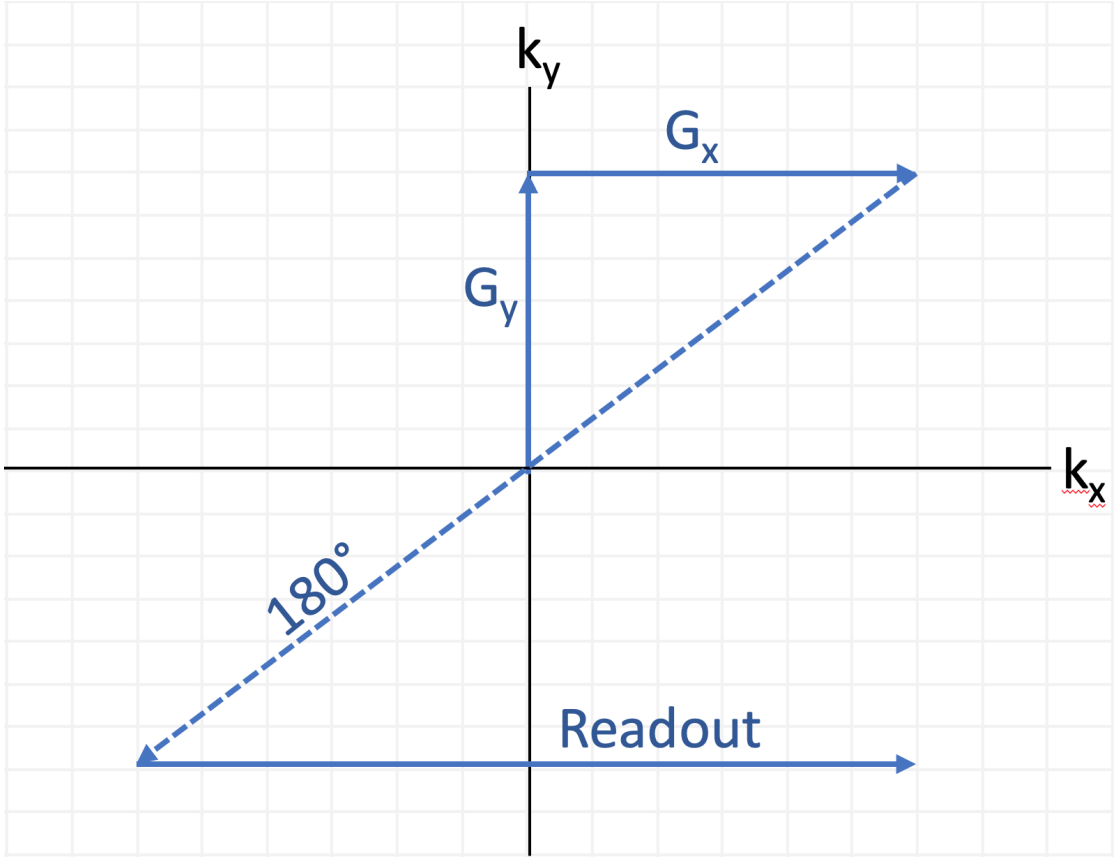


FIGURE 2.11: K-space diagram for a typical spin-echo sequence.

In the main magnetic field B_0 the precession of a spin has an angular frequency $\omega = \gamma B_0$. However if a magnetic field gradient $\mathbf{G}$ is applied in the direction $\mathbf{r}$ (both vectors) then this becomes:

$$\omega = \gamma B_0 + \gamma \mathbf{G}\mathbf{r} \tag{2.1}$$

If this is defined in a rotating frame of reference, where the original precession $\omega = \gamma B_0$ is our reference, it becomes $\omega = \gamma \mathbf{G}\mathbf{r}$. In order to work out the phase ϕ of a spin at time t, the frequency is integrated as:

$$\phi(\mathbf{r}, t) = \int_0^t \gamma \mathbf{G}(\tau)\mathbf{r}\ d\tau \tag{2.2}$$

And if the gradient $\mathbf{G}$ is constant through time then:

$$\phi(\mathbf{r}, t) = \gamma \mathbf{G}\mathbf{r}t = k_{\mathbf{r}}\mathbf{r} \tag{2.3}$$

where the spatial frequency of the image under a gradient $\mathbf{G}$ in the direction $\mathbf{r}$, $k_{\mathbf{r}}$, is defined as $k_{\mathbf{r}} = \gamma \mathbf{G}t$.

For different gradient directions $\mathbf{r}$ and gradient magnitudes $\mathbf{G}$ the value of the spatial frequency changes. If a map of spatial frequency values is constructed by acquiring multiple signals, the map can be used to create an image. The spatial frequency matrix, or k-space, is our method for representing the spatial frequency map of the image we wish to create. K-space is the Fourier transform of the image and vice versa.

Figure 2.10 shows how these gradients are used in combination to perform a spin echo sequence and Figure 2.11 shows how these gradients affect the position of the spins in k=space. As the 90° pulse is applied, G_{ss} is applied in the slice select direction to select the slice of the image. After this G_y, a gradient in the y direction, is applied, moving the position of the spins in k-space from the origin up the k_y axis. Then the gradient G_x is applied, moving the spins out parallel to the k_x axis. The 180° pulse then inverts the position of the spins through the origin to the other side of the k-space. Once

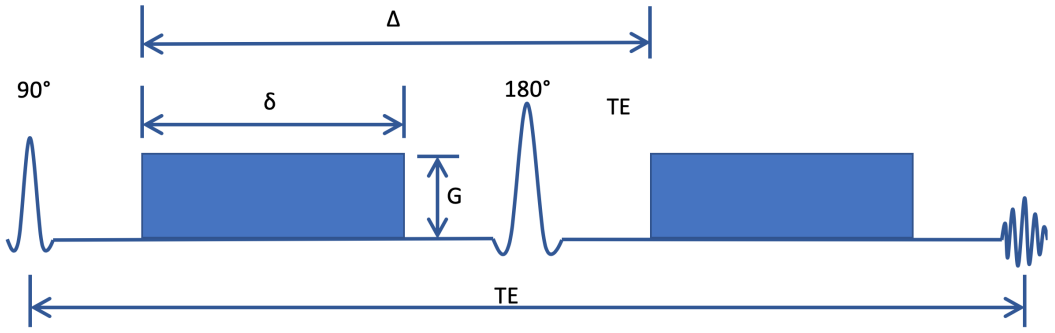


FIGURE 2.12: Typical PGSE sequence.

the position in k-space has moved to the correct position, the readout gradient is then applied and at each individual point along the readout line the signal is acquired. This whole process is then repeated for another line of k-space until all of the lines in the matrix are filled. Then in order to get an image the k-space matrix undergoes a Fourier transform.

2.3.2 Pulsed Gradient Spin Echo

Altering the pulse sequence and gradient parameters can make an MRI scan sensitive to the diffusion of water within a subject. Diffusion imaging can provide unique information on the underlying tissue structure.

A common sequence for sensitising the MR signal to diffusion is the Pulsed Gradient Spin Echo (PGSE) sequence, represented in Figure 2.12. This sequence is also known as the Single Diffusion Encoding (SDE) sequence. After the initial 90° pulse a strong directional gradient is applied in a specified direction, characterised by a gradient strength G, duration δ and gradient spacing Δ.

After time δ the gradient is switched off and a 180° pulse is applied. The gradient is then switched back on after time Δ from the application of the first gradient. The first gradient causes the precessional frequencies to vary with position along the applied gradient and the second gradient, after a 180° pulse, causes the phases of the protons to refocus and produce a measurable signal.

If the water molecules are static, the first gradient will alter their precessional frequency, causing them to experience a phase change relative to those at other positions along the gradient. The 180° pulse flips the spins in the transverse plane, meaning that the second gradient pulse, with the same magnitude and duration, will rephase the spins perfectly, producing a larger signal. However if the molecules are moving in the direction of the diffusion gradient, the dephasing first gradient and the refocusing second gradient experienced by the molecules will not be of the same magnitude as they are no longer in the same position along the field. The spins will remain somewhat out of phase, meaning that they will produce a diminished signal. In this way we can obtain a signal that is closely related to the mean level of random diffusional motion within a pixel.

2.3.3 Multi-Echo Spin Echo

As we have seen in the previous section a signal weighted to T2 decay can be created using a spin echo sequence. From this signal the T2 relaxation rate of the tissue can be estimated, which gives a useful insight into the composition of the underlying tissue. However, acquiring just one T2-weighted image means that the tissue in a voxel must be estimated to have a single T2-relaxation rate, which is not always the case in the underlying tissue. In some circumstances it may be desirable to collect multiple echoes at different echo times to better map out the exponential decay of the signal in the orthogonal plane.

A sequence used for collecting multiple echoes of the same decay is Multi-Echo Spin Echo (MESE), shown in Figure 2.13. Here a single 90° pulse is followed by a train of 180° pulses, each creating a single echo that diminishes in size as the spins undergo T2 decay. The signal intensities at these echo times then trace out a decay curve for T2 relaxation, reducing the effects of noise and providing more information about the nature of the decay.

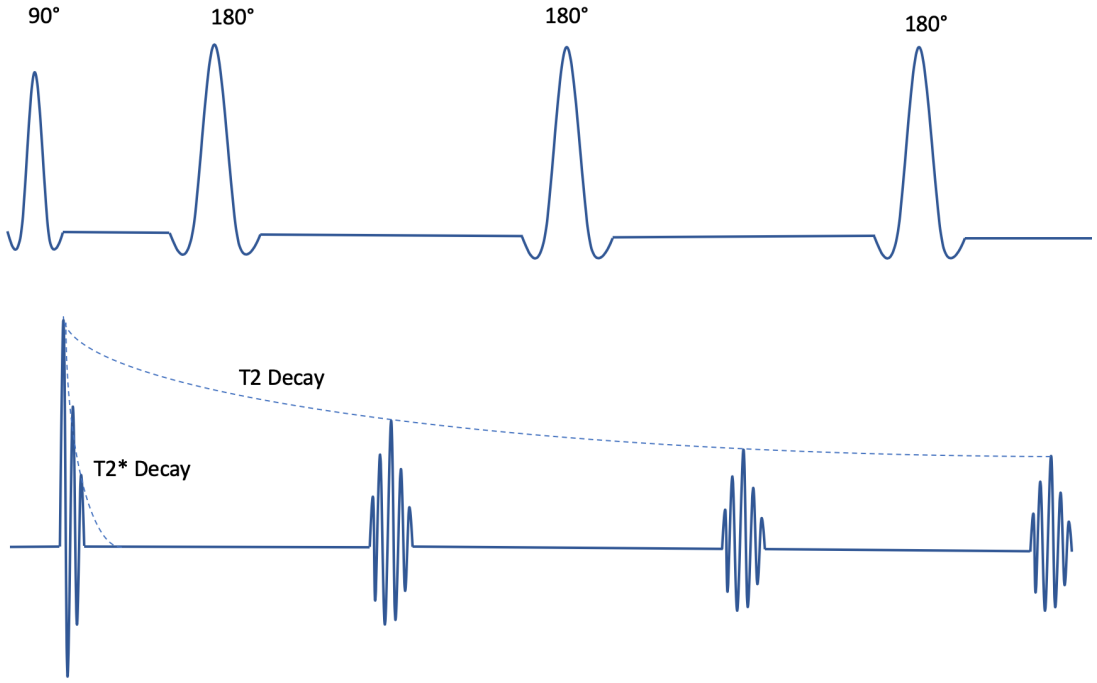


FIGURE 2.13: RF pulses and Signal respectively for a MESE sequence, showing T2 and T2* signal decays.

2.3.4 Fast Spin Echo

An alternative method for collecting multiple T2-weighted echoes is to use the Fast Spin Echo (FSE) sequence, also known as Turbo Spin Echo. Represented in Figure 2.14, the FSE sequence closely resembles a MESE sequence. The difference between the two is that different phase encoding gradients are used between each 180° pulse, meaning that multiple rows of k-space are acquired after each excitation. In this way, images are produced for each effective TE separately rather than collecting all echoes for a single line in k-space as MESE does. The term effective TE reflects the fact that an FSE image is a blend of multiple different TE values which can be heavily weighted towards one value, though not perfectly.

2.4 Modelling

Using pulse sequences we can discover new and unique information about tissue microstructure. Nevertheless, to better extract this information from the measurements

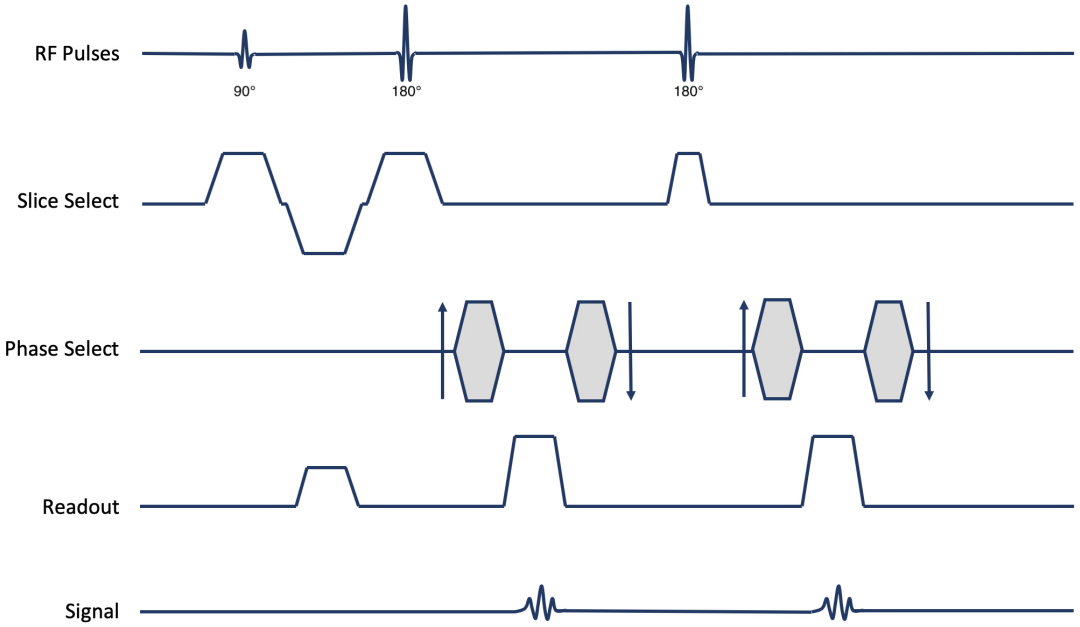


FIGURE 2.14: First two echoes of a typical FSE sequence.

we employ models which relate various features (for example, water diffusivity in the tissue or T2 properties) to the acquired data given the acquisition parameters.

This section will contain separate analyses of the more commonly used modelling techniques for both Diffusion and T2 imaging followed by a concluding remark on the differences between them and where new opportunities lie.

2.4.1 Diffusion modelling

Diffusion MRI employs diffusion-weighted sequences to sensitise the signal to the displacement of the water molecules in the tissue. As the displacement of water molecules is influenced by cellular architecture, a model which links tissue features (for instance, cell size and density) to the diffusion measurements can be used to obtain maps of microstructural properties.

2.4.1.1 ADC

The most basic model for diffusion imaging is the Apparent Diffusion Coefficient (ADC) model. It begins with the assumption that the underlying tissue is homogeneous, which

			Grading			Detection
Author	Year	Scanner	Population	r	P	P
Jyoti [79]	2018	3.0T	33	-0.365	0.014	-
Barbieri [10]	2016	3.0T	84	-0.47	<0.01	<0.01
Bittencourt [14]	2012	1.5T	24	-0.63	<0.01	-
Hambrock [67]	2011	3.0T	51	-0.497	<0.0001	<0.001
Itou [73]	2011	1.5T	60	-0.437	<0.0005	-
Rosenkrantz [130]	2011	1.5T	45	-	-	<0.0001
Turkbey [159]	2011	3.0T	48	-0.69	0.003	-
Langer [85]	2010	1.5T	24	-	-	<0.001
Tamada [154]	2008	1.5T	215	-0.497	<0.0001	<0.0001
Yoshimitsu [175]	2008	1.5T	37	-0.18	0.014	-

TABLE 2.2: Summary of recent works correlating Gleason Grade and ability to detect tumour with ADC in the peripheral zone of the prostate using a 3T scanner. It is noted that some methodologies may differ.

makes modelling much simpler. Using the PGSE sequence [1], along with the single, isotropic compartment and free diffusion assumptions, a physical derivation leads us to the relationship:

$$S = S_0 \; e^{-b \text{ x } ADC} \tag{2.4}$$

where S is the measured signal intensity using a diffusion weighting b to measure the diffusivity ADC. S_0 is the signal when there is no diffusion weighting applied ($b = 0$).

The value b in Equation 2.4 is the diffusion weighting of the pulse sequence, named the 'b-value', where $b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3)$. The b-value for a PGSE sequence is dependent on the gyromagnetic ratio γ, the gradient strength G, the duration of the gradient pulse δ, and the time between the two pulses Δ, which can be understood better by examining Figure 2.12.

The ADC can be estimated, by rearranging Equation 2.4, from the signal intensities acquired with at least two diffusion weightings (b = 0 and a non-zero b value usually around b = $1000 s/mm^2$). This model is termed a mono-exponential model because it fits a single exponential decay to the diffusion signal intensity values, a detail which will be important in understanding the move to more complicated models later in the chapter.

Past studies have shown that ADC is negatively correlated to Gleason Score. Table 2.2 is a summary of recent works in this area. The third column shows the correlation between ADC and Gleason Score found in these studies, with their corresponding p-values revealing that all studies found a significant relationship between the two ($P<0.05$). In all of these studies a statistically significant difference was found between the ADC values for normal tissue and for tumour, as suggested by the very p-values in the column of Table 2.2.

A falling ADC with increased Gleason Score means that the overall diffusivity of the underlying tissue is reduced as the cancer proliferates. This is of interest as it gives some insight into the microstructural changes occurring within the cancerous tissue. The tissue component with the highest ADC, the lumen, has been found in histology to be decreasing in fractional volume with increasing Gleason Grade whereas the components with low ADC, epithelia and stroma, are increasing in fractional volume [71]. However the exact contribution of each of these effects individually remains unmeasurable.

A recent study also looked into using quantitative values of ADC alongside mp-MRI images in order to reduce the number of patients who have unnecessary biopsy [123]. It found that the minimum ADC value of an ROI could rule out the need for biopsy in 33% of those men who would have required biopsy based on just mp-MRI.

However, there are limitations to ADC. From Table 2.2 it is clear that the correlation between ADC and tumour grade is variable and not strong enough to be certain of a tumour's grade purely based on ADC. Furthermore, ADC does not take into account vascularity within the tissue. The blood flowing through blood vessels affects the ADC of the tissue because if it is not accounted for in the model it can be mistaken for diffusion within the tissue. Methods for taking this into account in the model are described in the next sections.

If the aim of MRI in the prostate is to remove the need for biopsy, a correlation with tumour grade is a strong indicator of whether an MR technique will be effective or not. From these studies it can be concluded that using present MR techniques ADC alone cannot grade prostate tumours well enough to replace biopsy.

2.4.1.2 Diffusion Tensor Imaging

Diffusion Tensor Imaging (DTI) is a technique designed to take the anisotropic diffusion properties of certain tissues into account. ADC assumes isotropic diffusion and averages out the directional variation in diffusion-weighted images. DTI removes this assumption and accounts for anisotropy in the tissue. In order to measure the DTI parameters, field gradients in each of six orthogonal directions are applied in order to detect orientational differences in the amount of diffusion in a voxel. The diffusion tensor model is represented as:

$$S = S_0 e^{-b\mathbf{g}^\intercal \mathbf{D}\ \mathbf{g}}$$

where the diffusion tensor:

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix}$$

and the gradient direction is $\mathbf{g}$. Though this tensor has nine separate elements, due to its symmetrical nature only six are unique. The diffusion tensor $\mathbf{D}$ is estimated by acquiring six directions or more and a $b = 0$ image. Once the tensor has been estimated for a given voxel the diffusivity λ_1 can be estimated along the main axis of the tensor, and the diffusivities λ_2 and λ_3 along the two orthogonal axes, by calculating the eigenvalues of $\mathbf{D}$. From these the more relevant values of Mean Diffusivity (MD) and Fractional Anisotropy (FA) can be calculated, where:

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \tag{2.5}$$

and

$$FA = \sqrt{\frac{1}{2}} \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}. \tag{2.6}$$

This technique has shown mixed results in the brain, encountering problems with crossing fibers and multiple tissue compartments. In the prostate it has generally been found that whilst the MD from DTI can differentiate prostate cancer from benign tissue the FA cannot[18, 21, 33, 53, 84, 104, 161, 172]. Bourne et al. use diffusion microscopy to image individual cell compartments[18, 21, 22]. They assign a high FA value to stromal tissue, a lower FA to epithelia and no FA to lumen. They go on to hypothesise that the generally low values of FA, and its inability to differentiate malignant and benign tissues, are due to the heterogeneity in both orientation and tissue composition in the tissue within a traditional DTI voxel.

2.4.1.3 Diffusion Kurtosis Imaging

Assuming that most diffusion decay curves do not perfectly trace out an exponential, a method to quantify the deviation from the perfect curve could be useful in investigating the tissue microstructure. A number of methods have been previously suggested to describe the deviation of the diffusion data from the ideal homogeneous, single compartment model, including the Kurtosis model, the Stretched Exponential model and the Statistical model.

The Diffusion Kurtosis Imaging (DKI) model, initially proposed by Jensen et al. [75, 77], is based on a Taylor series expansion of the signal intensity. The Kurtosis parameter K is related to the fourth moment of the displacement of the spin during the sequence, a mathematical description that will not be covered here. It quantifies the deviation from the ideal exponential decay, where $K = 0$ is the monoexponential fit and increased heterogeneity causes a rise in K.

The key relationship in DKI is:

$$S = S_0 \exp(-bD + \frac{1}{6}b^2D^2K) \tag{2.7}$$

In the brain DKI proves to be sensitive to microstructural change [34, 48, 75–77, 91, 93, 99, 113, 171] as well as providing a better fit to diffusion data than ADC [98] and being more robust than other diffusion models [50, 145]. DKI in the prostate has been

able to differentiate cancerous and benign tissue [74, 107, 120, 125, 132, 153, 155] and has good sensitivity and specificity to tumours[132]. DKI also has a higher information content than monoexponential, stretched exponential and biexponential models [23, 74], provides a better fit to diffusion data than ADC, Stretched Exponential or Biexponential models [74, 120] and is repeatable and robust to noise [74].

However there remain a number of problems with DKI. The main issue is its relation to histology where, whilst some relationships between DKI, cellularity and cellular heterogeneity has been postulated [90], the model does not consider each tissue component separately and so can not separate the effects of different microstructures [56]. Furthermore its superiority to other diffusion methods has been refuted in a number of other studies [125, 129, 156]. DKI needs a specific range of b-values, with the maximum lying between 1000-3000 s/mm^2 [131], meaning that scan times can be longer than traditional diffusion acquisitions [125].

2.4.1.4 Stretched Exponential Model

The Stretched Exponential Model [13] is similar to DKI in that it tries to quantify the variation from the monoexponential with a single variable. The basis of this method is the equation:

$$S = S_0 \exp\{-(bD)^{\alpha}\} \tag{2.8}$$

where α is the stretching factor that alters the shape of the traditional monoexponential fit.

Some papers have found a link between tissue heterogeneity and the stretching factor α [12, 13, 39, 66]. However, when compared to other diffusion models, the Stretched Exponential has been found to be less repeatable and not fit the data as well as DKI and two-compartment models [23, 74, 112]. It also contains less information than DKI [23] and is less robust to noise than DKI [93]. Furthermore, as with DKI, the fact that it only accounts for a single tissue compartment means that it is limited in its relation to histology [56].

2.4.1.5 Statistical model

Diffusion MRI is important in the detection and characterization of prostate cancer due to its unparalleled sensitivity to tumour microstructure. Many of the measures currently used to estimate diffusion within a voxel do not consider the heterogeneous nature of biological tissue [85], which affects the quality of the data fitting. Introducing statistical modelling [173] helps to relate diffusion images to the underlying tissue microstructure [118].

The Statistical Model used by Yablonskiy et al. [173] fits a Gaussian distribution to the diffusivity values. This means that the heterogeneity of the diffusivity values within a voxel can also be quantified using a variance parameter. The equation governing the model is:

$$S = S_0 \int_0^\infty P(D) exp(-bD) dD \tag{2.9}$$

Assuming a Gaussian distribution with mean μ and variance σ^2:

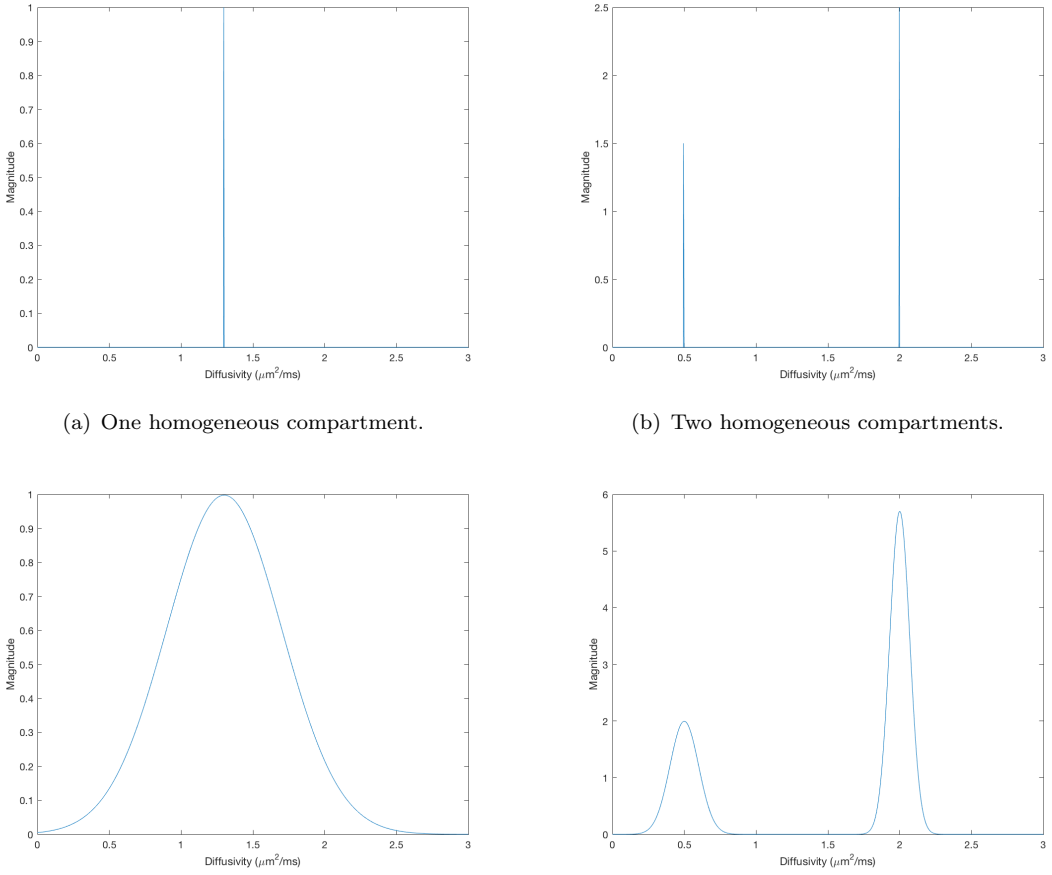
$$P(D) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-(D-\mu)^2/2\sigma^2) \tag{2.10}$$

The Statistical Model and DKI have an identical mathematical form but rely on different underlying assumptions. The Kurtosis parameter is in fact a specific property of the distribution, related to P(D) by:

$$K = \frac{\sigma^2}{\mu^2} \tag{2.11}$$

Examples of the different diffusivity models are shown in Figure 2.15. For instance, 2.15(d) shows two Gaussian curves representing two compartments with mean diffusivity values at 0.5 and 2 $\mu m^2/ms$ respectively. The compartment with the greater mean diffusivity also has the higher peak, suggesting that more of the signal comes from a tissue with a higher diffusivity.

The Statistical Model has proven sensitive to microstructure [93], is robust to noise, can differentiate tumour from normal tissue and importantly can theoretically be better related to histology than either DKI or the Stretched Exponential [146]. However this particular form of the Statistical Model still assumes a single tissue compartment.



(a) One homogeneous compartment.

(b) Two homogeneous compartments.

(c) One heterogeneous compartment.

(d) Two heterogeneous compartments.

FIGURE 2.15: Example models for diffusion data.

When modelling compartments using distributions, an important factor is which distribution to use. There is a wide range of possibilities, each with their own benefits and downsides. Whilst a good fit to the data is important the distribution must not be overly complex, which could open up the possibility of overfitting. Furthermore it would be very computationally advantageous to use a simple model. The two most common choices in the literature are the Gaussian distribution and the Gamma distribution. Whilst some works have found Gamma distributions to better fit the data [118, 139], Shinmoto et al. found that Gamma distributions only have an advantage within prostate cancer and fit PZ and BPH tissue no better than a Gaussian. The benefit of the Gaussian is that it is mathematically simpler to implement and the two parameters have a more direct tissue relevance and are independent of each other, unlike when the mean and variance of the Gamma distribution are calculated from α and β.

2.4.1.6 Multi-exponential DWI

It has been postulated that diffusivity is measurably different in each of the three major tissue compartments within the prostate [19, 20, 32, 85, 114, 122], with epithelial cells being the most restricted, followed by stromal cells and then glands, and blood vessels having the highest diffusivity. In terms of how restricted the water molecules are physically, these observations make sense. The epithelial and stromal cells are much smaller in scale to the lumen and blood vessels so the molecules within them are much more likely to encounter a boundary. Multiple studies designed to investigate this have indeed found that the diffusion curve is non-monoexponential, suggesting that there is more than one tissue compartment in the voxel [21, 147, 160].

The simplest multi-exponential model, a bi-exponential, takes the form:

$$\frac{S}{S_0} = \alpha e^{-bD_{fast}} + (1-\alpha)e^{-bD_{slow}}$$

where α represents the fraction of the tissue made up of the fast diffusing component, D_{fast} is the mean diffusion of the fast diffusion compartment and D_{slow} is that of the slow diffusion compartment. However this can easily be extended to three or more exponentials by adding more parameters to the equation.

More of the water molecules within stromal and epithelial cells will be bound to other molecules, meaning that their decay rate should be faster than those freely moving within the lumen. However the multi-compartment approach must be approached with caution as its results can be misleading. It has been shown that two compartments can even be found within a single cell [143] and that there are many different confounding factors, such as cell geometry and orientation, that can lead to such an effect [35, 112].

The fact that tissue microenvironments within the prostate have different MR properties motivates the use of models with increased numbers of parameters to try to capture some of the extra detail. As described earlier, prostate cancer destroys the lumen in the tissue. Hence we would expect the amount of signal received from the luminal compartment to reduce with the grade of the tumour.

2.4.1.7 Spherical Mean Technique

Both Bourne et al.[18, 21] and Xu et al. [172] found in ex-vivo prostates that areas of high stromal content have a larger fractional anisotropy. Few studies find diffusional anisotropy in normal prostate tissue. This can be attributed to the fact that the large voxel size of traditional DTI sequences allows tissue and directional orientation to mask this effect[18]. Bourne et al. used diffusion microscopy with extremely high resolution to solve this problem but this can only be done ex-vivo. A technique for measuring anisotropy on a microscopic scale whilst limited to traditional pixel sizes is needed.

Intriguing papers by Kaden et al. [80, 81] propose a novel model, the Spherical Mean Technique (SMT), to achieve this in the brain by separating the effects of individual axon anisotropy from the orientational dispersion of these axons. They achieve this by modelling the signal as the convolution of the individual axons with their dispersion.

The approach of removing the effects of orientation dispersion holds promise for the prostate. As discussed, the stromal and epithelial tissue are themselves anisotropic in nature but this effect is diminished on the macroscopic scale by orientational heterogeneity. Hence the scale of this effect must be investigated if an accurate model of the prostate is to be created. This is an area that will be investigated in this thesis.

2.4.1.8 VERDICT

The VERDICT diffusion model (which stands for Vascular, Extracellular, and Restricted Diffusion for Cytometry in Tumours), proposed by Panagiotaki et al.[120, 121], is a multi-compartment model for tumours that can be modified for use in the prostate, each compartment of which is itself represented by diffusion tensors. These tensors, which are chosen for the goodness of fit they achieve to the data, can reveal something about the underlying tissue microstructure.

Figure 2.16 shows how the VERDICT model is related to tissue microstructure. The epithelia and the stromal nuclei, which are both highly restricted, are represented by a 'sphere' with a solid boundary at distance R, diffusivity within the boundary of d_{ic}

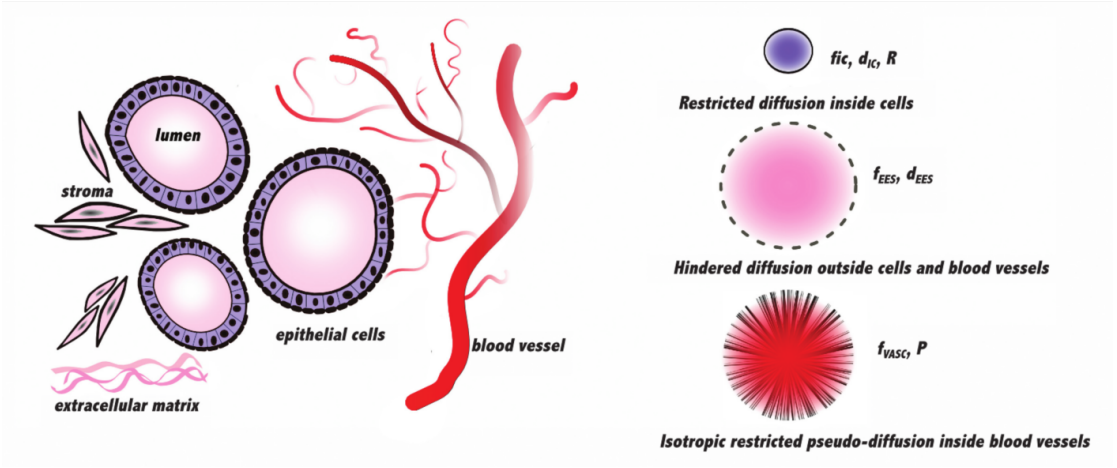


FIGURE 2.16: VERDICT tissue model (reproduced from Panagiotaki et al. [120] with permission).

and a volume fraction within the tissue of f_{ic}. Similarly the lumen are represented by a diffusivity d_{EES} and volume fraction f_{EES} but no solid boundary, as within realistic diffusion times few water molecules will encounter a boundary there. The vascular compartment is represented by a collection of isotropically distributed 'sticks' with a pseudo-diffusion coefficient P and volume fraction f_{vasc}.

The VERDICT model has proven to be able to discern cancer from benign tissue whilst being more specific to biological microstructural features than either ADC or DKI[121]. It is hypothesised that this link to microstructure could eventually eliminate the need for DCE MRI[120]. An optimised version of the technique, recently presented by Bonet-Carne et al.[16], created maps that were more robust to noise than the original VERDICT.

Machine learning techniques have recently been used to classify PCa using the VERDICT parameters, producing an AUC of 86.7% [36]. A recent conference publication also showed that the f_{IC} parameter better discriminates between Gleason 3+4 and Gleason 3+3 than ADC and that the image quality of VERDICT and ADC are comparable[26]. Preliminary studies have also been able to successfully link the VERDICT model parameters to histology [17].

2.4.2 T2 modelling

T2 relaxometry calculates the actual tissue T2 value from data acquired with a multiple echo sequence. Simple models can estimate an average T2 value for the tissue within a whole voxel whereas more complicated models attempt to discern different tissue compartments and their respective T2 values, at the expense of scan time and SNR.

2.4.2.1 Traditional T2-weighted Imaging

T2-weighted images acquired using a single Spin Echo sequence provide good soft tissue contrast and high image resolution. The relationship between the signal intensity and the underlying T2 relaxation rate for a perfectly homogeneous tissue is:

$$S = S_0 \exp(-TE/T2) \tag{2.12}$$

TE is the time between the initial 90° pulse and the echo, S_0 is the signal intensity at time zero and $T2$ is the intrinsic transverse relaxation rate of the tissue. Equation 2.12 shows that a high signal intensity is equivalent to a long tissue T2 value.

A Single Spin Echo sequence is not enough if we are to estimate the value of T2 using Equation 2.12. Due to the fact that the signal intensity cannot be measured at time $TE = 0$, S_0 must be estimated from datapoints at different TEs. This is where multiple echo sequences become necessary. This model is called a mono-exponential model because a single exponential decay is fit to the signal intensities. To understand how T2 relaxation rates can inform us about tissue types and anatomy, it is important to highlight that different soft tissues have different T2 values.

It has been shown that the T2 measurement is related to the abundance of free water [110], cell density [52, 86], luminal, epithelial and stromal volumes [85] and levels of citrate in the prostate [96, 97], which is itself related to luminal partial volume [38, 58, 144]. Whilst current T2-weighted images can localize tumours with relatively good sensitivity and specificity [63, 65, 72, 138, 166] there is varying literature relating to a

correlation to Gleason Score [45, 82, 144]. Many lesions are scored as indeterminate, meaning that more information might help in making a decision.

2.4.2.2 Multi-exponential T2

Much of the work done on multi-compartment T2 models is applied to the brain, so much of the literature in the following section, though not directly related to the prostate, gives a good representation of the possibilities of these techniques in probing microstructure. Multi-exponential fits are the simplest multi-compartment models. Schad et al.[137] discovered in the brain that, whilst normal white and grey matter generally displayed a monoexponential behaviour, cancerous tissue was better represented by a bi-exponential, and went on to show that this difference meant that the model was able to discriminate well the tumour region. Armspach et al. then [9] did a similar study and found that a bi-exponential fit was superior to the monoexponential in Multiple Sclerosis lesions and that the long T2 component was significantly different in MS than in normal white matter.

In the prostate it was found that healthy prostate tissue is multiexponential in nature [82], with the postulation that these components correspond to the epithelial, luminal and stromal regions. Furthermore the biexponential model was favoured over the monoexponential and triexponential in 86% [152] and 96% [57] of cases respectively, in turn proposing but not proving that one of these exponentials comes from the lumen and that the other originates in both the stroma and epithelia combined. Relating epithelial and stromal cells with fast T2 decay and the luminal spaces with slow T2 decay [82, 152] agrees with information about the tissue structure within the prostate, where far more of the water molecules within stromal and epithelial cells will be bound to other molecules. This binding of molecules leads to more hydrogen nuclei being in close proximity to one another, meaning that spin-spin interactions are more common and so the decay rate should be faster than those freely moving within the lumen. Another theorised explanation for the slower decay in the lumen is that there are likely to be more localised static field inhomogeneities in tissue, where there are a lot of different molecule types and tissue microenvironments, than the lumen.

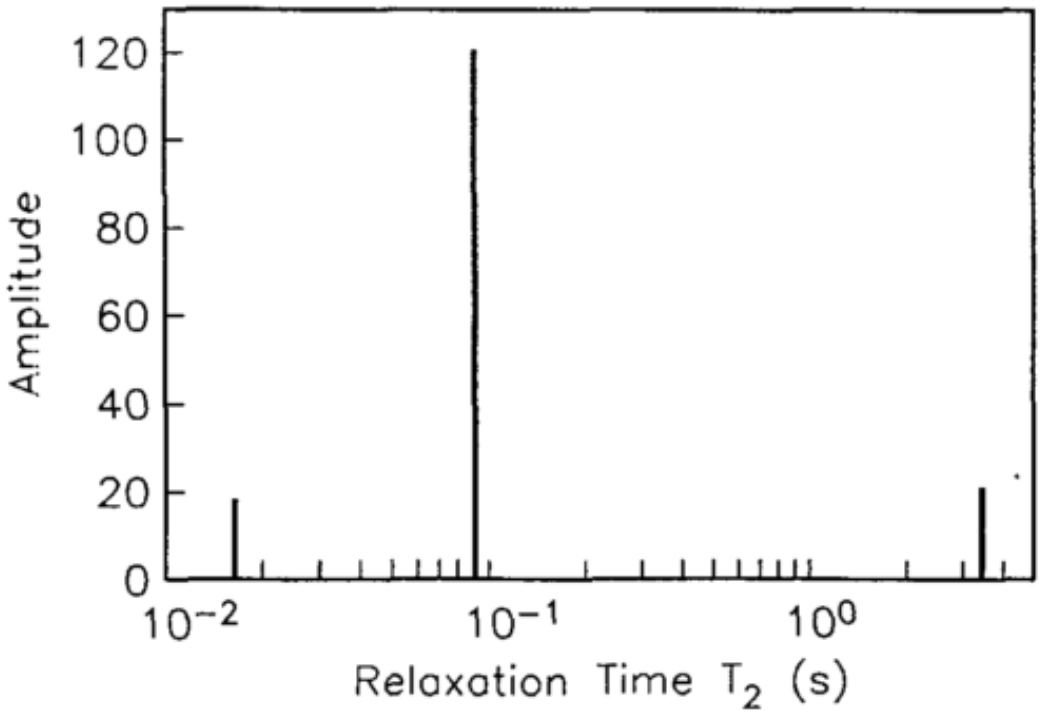


FIGURE 2.17: T2 spectrum using NNLS fitting (reproduced from Mackay et al. [101] with permission).

2.4.2.3 T2 Distribution

Distribution models, wherein the T2 spectrum is modelled as a distribution rather than a number of single values (multi-exponential models), are extensively researched in the brain. Three particularly detailed review papers [7, 89, 100] explain the concept of Myelin Water Imaging (MWI), a distribution model designed to calculate the fraction of the tissue volume in the brain taken up by myelin.

Early pioneers of this multi-distribution modelling were MacKay et al. [101], using an unconstrained non-negative least squares (NNLS) fitting algorithm to produce a discrete spectrum of values like those of Figure 2.17. In the figure we can see three large peaks, showing that in this case the model detects three tissue compartments with different magnitudes and T2 values.

The NNLS method seeks to minimise the weighting vector s in the cost function in Equation 2.13. N is the number of echoes collected in the acquisition, M is the selected

number of parameters to be fit by the algorithm and y is the vector containing the signal intensities and $A_{ij} = e^{-TE_i/T2_j}$.

$$\sum_{i=1}^{N} |\sum_{j=1}^{M} A_{ij} s_j - y_i|^2 \tag{2.13}$$

The unconstrained nature of this method means that it is subject to no limiting assumptions on the constituents of the underlying tissue. However the main problem with MacKay's original method was that due to the nature of the optimisation the peaks often became very narrow, which meant that they, like multi-exponential models, were subject to high noise sensitivity. This also explains why in Figure 2.17 the peaks are so narrow. In order to solve this Graham et al. [62] introduced a regularisation term to help smooth the peaks, making them wider and as a result less sensitive to noise. Most of the subsequent MWI studies employed some variant on this regularisation term as it makes the method much more stable and reliable.

Many papers have found the unconstrained distribution model to provide a better fit to the T2 decay curve than the monoexponential model [7, 8, 62, 89, 165, 168, 169]. By increasing the number of degrees of freedom in a fitting procedure it will naturally have a closer fit but in this case the new parameters are worthwhile as they have been shown to offer new insight into the tissue microstructure. Studies have confirmed the link between the Myelin Water Fraction (MWF), a parameter estimated using the MWI fitting, and the myelin fractional volume attained histologically [7, 88, 89, 100, 111, 165]. This validation shows the value of multi-compartment modelling over traditional monoexponential methods.

When using the unconstrained model the number of variables is large, in some studies numbering in the hundreds of variables. However most standard pulse sequences cannot provide this many datapoints. The large disparity between the number of datapoints and the number of variables means that the problem is underdetermined. In other words, there are a number of solutions to the same problem, leading to less reliable results.

In order to resolve the underdetermined problem in the unconstrained model, papers by Stanisz et al.[150] and Raj et al.[126] use the idea of constraining the distribution to

be a combination of parameterised distributions. Previous studies had shown that the T2 spectrum formed into two peaks of roughly Gaussian shape so in order to reduce the number of degrees of freedom in the optimisation the distribution was constrained to be a combination of two Gaussian distributions. This method showed similar results to the unconstrained version whilst being able to function with fewer datapoints and producing less noise in the final parameter maps.

2.4.2.4 mcDESPOT

A limitation of all modelling methods reliant on multi-echo data is that the acquisition is frequently so long that only a single slice sequence is clinically feasible, heavily limiting its viability. The current gold-standard MESE acquisition is the Carr-Purcell-Meiboom-Gill (CPMG) sequence [168]. Alternative strategies of acquiring T2 relaxometry data including FSE, are quicker but still relatively time consuming [44, 116].

mcDESPOT (Multi-Compartment Driven Equilibrium Single Pulse Observation of T1/T2) is a method designed to reduce the scan time necessary to model multiple compartments in both T1 and T2 relaxation. It uses the SPGR [37] and SSPF [29] pulse sequences and is designed to keep the TR of the sequence constant whilst varying the flip angle α. It has been found that these steady state sequences, in conjunction with the DESPOT1 and DESPOT2 models [42, 43], are able to calculate accurate and repeatable T1 and T2 values using only two datapoints per voxel (as described in Deoni et al. [41]), enabling a much faster and higher resolution scan.

The DESPOT1 and DESPOT2 models only accounted for one tissue compartment, which was addressed by formulating the mcDESPOT model [41]. This combines the original one compartment models with the assumption of two exchanging water compartments. This gives mcDESPOT the ability to calculate the MWF in brain tissue in a fraction of the time of normal sequences. Furthermore, the inclusion of the exchange between water compartments has not been addressed by any other two compartment model described thus far, giving this model an extra degree of information. mcDESPOT has proven itself able to calculate the T1 and T2 of both the fast and slow components,

along with the exchange rate between the compartments and the MWF, in a shorter time and with a higher resolution than is possible with other sequences.

The downsides of this method are that it is very susceptible to field inhomogeneities and errors in the flip angle. Deoni et al.'s work [41] then goes on to describe that although distribution models could theoretically be used with mcDESPOT, to remove the necessity to specify the number of compartments, the MWF is still considered to be the most clinically useful parameter and hence these increases in complexity may not be worth the additional information. Furthermore, the mcDESPOT model takes into account water exchange between different cell compartments, which is its main strength. However in the prostate we can assume negligible exchange between the lumen and the rest of the tissue due to it's large size relative to the surrounding cells [56]. Hence the main advantage of the mcDESPOT method seems mostly redundant in the prostate.

2.5 Image unwarping

In Chapter 7, in order to perform voxel-level classification using machine learning, based on a combination of T2 and diffusion data, the images must be correctly aligned. There are many methods for registering medical images [103, 117, 163]. Many of these methodologies perform well when tested on simulated signals and phantom tissue models but their validation in-vivo has been challenging.

In order to align diffusion and T2 images, in this work the diffusion images were unwarped using in-house MATLAB code developed by a software engineer within the research group. The main source of misalignment between DWI and T2-weighted images is the inhomogeneous B0 field caused by inhomogeneities in the magnet inside the scanner, magnetic materials within the field and technical problems. These inhomogeneities cause echo planar imaging (EPI) based images to become distorted when compared to T2-weighted images as the reconstruction process is based on the assumption of a homogeneous field and the low bandwidth of EPI makes it more susceptible to field variations.

In order to undistort the diffusion images a B0 map was acquired as part of the scanner protocol. Using the bandwidth/pixel ratio calculated by the scanner the number of pixels by which the image was shifted due to these inhomogeneities could be calculated.

After the distortion correction the next step is to re-slice the DWI in order that it is in the same plane and has the same resolution as the T2-weighted images. This was done using 3D interpolation whereby the interpolating kernel takes into account both the slice thickness and in-plane voxel size and ensures anti-aliaising, again using in-house software.

2.6 Machine learning

Machine learning is a data analysis method that has the ability to automatically learn and improve from experience in order to predict outcomes without being explicitly programmed to do so.

Typically machine learning is split up into four broad categories; supervised, unsupervised, semi-supervised and reinforcement learning [61]. Supervised learning uses labelled data to predict the output variables and is usually used for regression and classification. Unsupervised learning is trained on unlabelled data, relying on pattern recognition in the input data to label datapoints. Unsupervised learning is commonly used for the clustering of datapoints. Semi-supervised learning is a combination of supervised and unsupervised learning whereby some of the data is labelled and some is not, usually used in situations where labels are costly to produce. Reinforcement learning maximises rewards and minimises risks using observations it iteratively gathers from its environment.

All of the data used in this thesis is fully labelled so a supervised method would be the most appropriate. In this work neural networks were chosen for all classification tasks due to the fact that they can be used for supervised learning and because of their versatility and their ability to infer non-linear relationships.

An example of a neural network is shown in Figure 2.18. Neural networks consist of an initial layer of input nodes, an end layer of output nodes and a number of 'hidden'

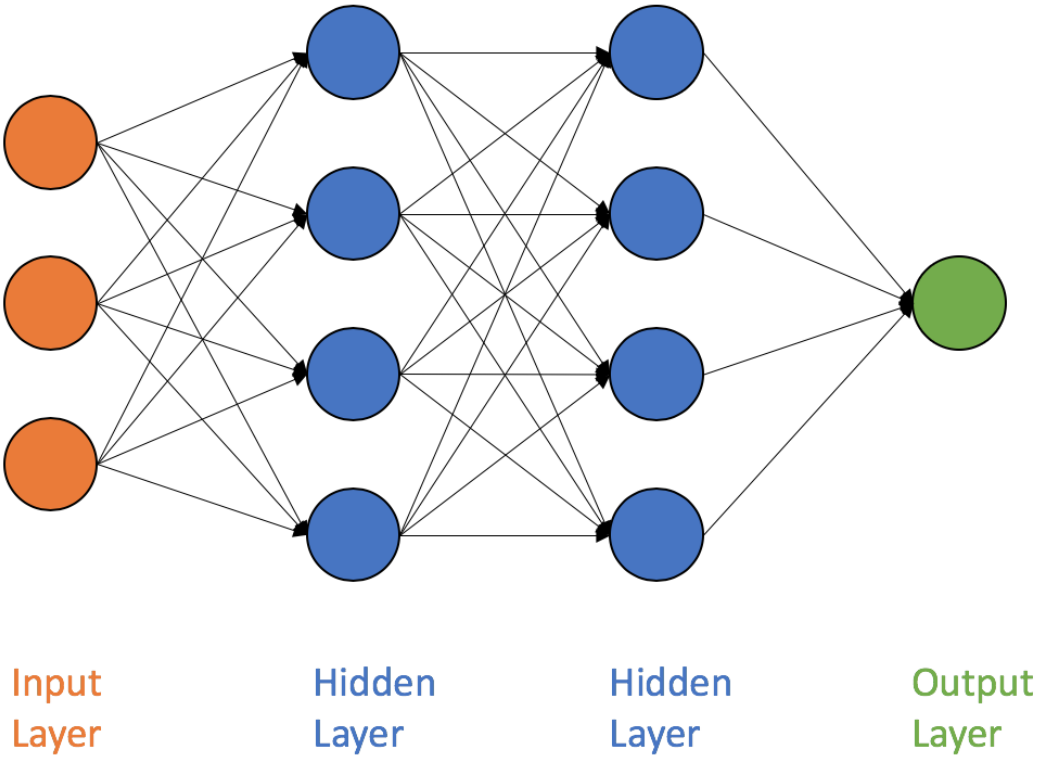


FIGURE 2.18: Diagram of example neural network with two hidden layers.

layers between them. The number of layers between the input and output layers and the number of nodes within each of the hidden layers can be adjusted for the problem at hand. All of the connections between the nodes have weights so that the input to a node is the output of a node in the last layer multiplied by the weight of the connection between them. Each node sums up all of its inputs and a bias and then uses an 'activation function' to convert its inputs into an output.

There is a wide range of activation functions available, three of which are shown in Figure 2.19. Activation functions read in the sum of the weighted inputs as:

$$output = \phi\left(\sum_{n=1}^{N}(\theta_n x_n) + \theta_0\right) \tag{2.14}$$

where N is the number of nodes in the previous layer, ϕ is the activation function, θ_n is the weight of the connection from the n^{th} node in the previous layer and x_n is the output from the n^{th} node in the previous layer.

There are pros and cons to each choice of activation function. The linear activation

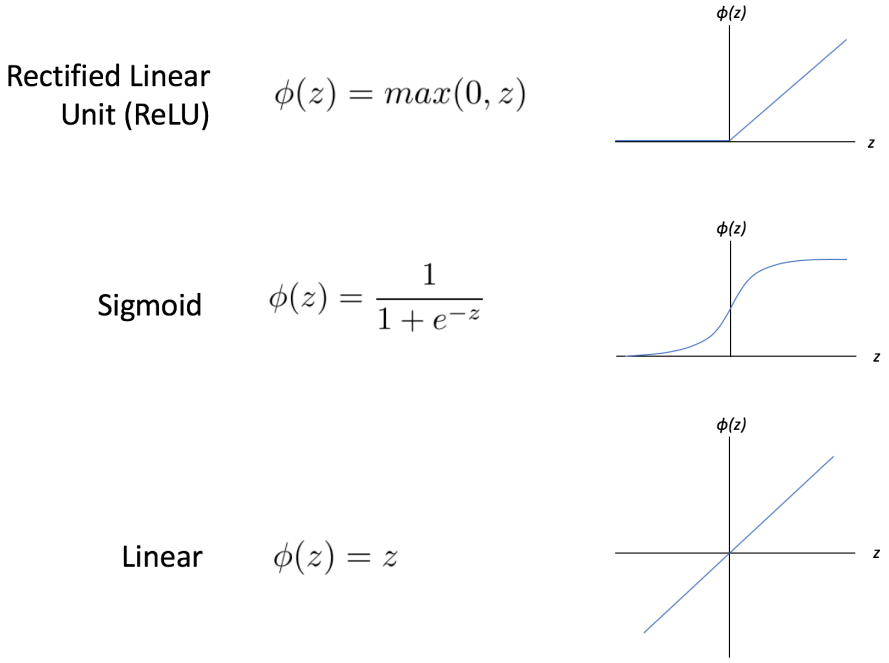


FIGURE 2.19: Diagram of example neural network with two hidden layers.

function is very simple but layers cannot be stacked as any combination of linear functions is still linear and hence can be replaced by a single layer. The sigmoid function is a strong classifier due to its steep gradients in the central region and allows the stacking of layers due to its non-linear nature but the small gradients as z moves further from the origin means that the network can sometimes be very slow to learn. The ReLU function is non-linear, enabling multiple layers, and does not suffer from small gradients at the extremities. Another benefit is that, because the ReLU outputs zero for values of z less than zero, it has fewer neurons activated at any one time, meaning that it is more efficient at representing the data and is better at disentangling variations in the data [60]. However one problem with ReLU is that if a node goes into the zero-gradient region of the activation function it will stop responding to changes in the input, making a large part of the network passive.

Once the size of the neural network has been defined, before any optimisation occurs it has to be initialised. An obvious method of initialisation would be to set all of the weights to zero but this does not work as it means that each set of weights mirror each other and cause the network to become highly redundant. A much more effective initialisation is to set each weight to be a random value within a set range, usually small

values close to zero.

Once the weights have been initialised, a cost function over which the neural network is minimised must be defined. A common cost function when a ReLU or sigmoid activation function is used is cross entropy, given by:

$$J(\Theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{k=1}^{K} y_k^{(i)} log(\phi_\Theta(x^{(i)}))_k + (1-y_k^{(i)}) log(1-(\phi_\Theta(x^{(i)}))_k)\right] + \frac{\lambda}{2m}\sum_{l=1}^{L-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}(\Theta_{ji}^{(l)})^2 \tag{2.15}$$

where $J(\Theta)$ is the cost function of the vector of weights Θ, m is the number of datapoints, K is the number of output nodes, $y_k^{(i)}$ is the actual value of output node k for datapoint i, $\phi_\Theta(x^{(i)})_k$ is the prediction of the model for output node k with weights Θ and input variables $x^{(i)}$, and λ is the weight decay parameter which controls how large the effect of the regularising term is.

Once a cost function has been defined, the weights can be optimised for a training set of input and output values. A common optimisation algorithm is gradient descent:

$$\Theta_{ij}^{(l)} \Leftarrow \Theta_{ij}^{(l)} - \alpha\frac{\partial}{\partial\Theta_{ij}^{(l)}}J(\Theta) \tag{2.16}$$

whereby the weight vector Θ is updated as the previous value minus the gradient of the cost function $J(\Theta)$ multiplied by the learning rate α, where the weights are indexed for each layer of weights l between nodes i and j of the two connected layers.

In order to calculate the values of the output nodes a process called forward propagation is used to move forwards through each layer, calculating the output of each layer based upon the inputs from the last. However, as the gradient descent method needs to know the gradient of the cost function for each weight, a process called back propagation is used to quickly and efficiently calculate these. Through a long derivation [92] it can be shown that for each layer of weights l between nodes i and j of the two connected layers:

$$\frac{\partial}{\partial\Theta_{ij}^{(l)}}J(\Theta) = a_j^{(l)}\delta_i^{(l+1)} \tag{2.17}$$

where $a_j^{(l)}$ is the output value of the j^{th} node in the the layer l and $\delta_i^{(l+1)}$ is the 'error' of node i in layer $l+1$. $\delta_i^{(l)}$ is defined as:

$$\delta_i^{(l)} = (\Theta_i^{(l)})^T \delta^{(l+1)} .* \phi'(z_i^{(l)}) \tag{2.18}$$

where ϕ' is the derivative of the activation function and $z_i^{(l)}$ is the vector of input values for node i in layer l. For a sigmoid activation function it can be shown that:

$$\phi'(z^{(l)}) = a^{(l)} .* (1 - a^{(l)}) \tag{2.19}$$

For the output layer l_o the error is defined as:

$$\delta^{(l_o)} = a^{(l_o)} - y \tag{2.20}$$

where y is the vector of output values and $a^{(l_o)}$ is the vector of estimated output values.

In order to calculate the updated values for each of the weights the differential of the cost function must be found with respect to each of the weights. Equations 2.17-2.20 show that this is most efficiently done by calculating the error of the output layer through forward propagation and then stepping backwards layer by layer to calculate the differential of the cost function for each weight using back propagation.

Machine learning has been used in the prostate already, from segmenting digital histology [102] to detecting prostate cancer using radiomics [15, 164] to combining textural analysis of diffusion and T2 images [49]. Few studies have been made on using machine learning techniques on voxel-wise diffusion or T2-weighted data from the scanner to detect and grade tumours. One work showed that using convolutional neural networks on mp-MRI images to produce lesion probability maps could have similar sensitivity and better specificity than a 6-core prostate biopsy in localising prostate lesions and then goes on to show that these methods also show promise in being able to grade these lesions accurately [174].

2.7 Summary

This chapter has given a summary of current clinical methods and cutting edge diffusion and T2 models. The current method of performing digital rectal examinations, TRUS guided biopsies and histological analysis is invasive and inaccurate. Cutting edge MR sequences and models may be able to improve upon this but are not yet as sensitive and specific to the grades of PCa as histological analysis. However, the utility of a model that could accurately model tissue microstructure is clear.

There is an opportunity to use distribution models in the prostate. None of the ADC, DTI, DKI, Stretched Exponential, multi-exponential or VERDICT models directly quantify a variance for the diffusivity values. Statistical models, especially when multiple compartment, have been shown to relate to individual tissue components but need to be developed further in the prostate. If the relative heterogeneity of the prostate could be quantified, this information could be used to try to detect and grade tumours in the prostate.

Another opportunity lies in better quantifying the diffusion anisotropy of prostate tissue. There are diffusion acquisitions and models in the brain, such as SMT, specifically designed to cope with regions where axons are crossed, an area that traditional diffusion tensors fail to model accurately. There is little to no anisotropy shown by diffusion tensors in the prostate, even though epithelia and stroma are characised as long and thin cells of radius $\sim 10 \mu m$ [120] which would be expected to have anisotropic diffusion characteristics. As glandular lumen could be expected to have more isotropic diffusion due to their structure, a method that could remove the confouding orientation of cells in the prostate could indirectly measure the relative fractions of the tissue made up of cells. As we know PCa is a proliferation of cells that destroys the lumen, it follows that this would then be a method that was sensitive to prostate cancer.

The study of T2 modelling is much less widespread than diffusion modelling, especially in the prostate. Traditional T2-weighted images are well established but make limited use of the information available. Introducing multiple compartments using NNLS fitting improves the model's relation to histology but at the expense of some robustness to

noise. Distribution models have the benefit of being more robust whilst retaining their relationship with histology but as yet are unproven in the prostate. Their application in the brain to calculate the MWF, which similarly contains two distinct T2 components, means that their application to the prostate could be of interest.

Furthermore, machine learning is a strong and versatile technique that is used in a wide range of classification tasks. There is little prior work utilising machine learning methods in the prostate and none combine them with ME-T2 datasets. This novel method for detecting prostate cancer has the strength of not needing a model, meaning that none of the tissue assumptions applied in other circumstances need to be made. It is an area that definitely deserves further research.

Overall, the opportunities highlighted in this section all push the boundaries of current MR techniques by combining richer datasets with novel tissue models. Current methods work well but suffer from not being specific enough to PCa. By more faithfully modelling the prostate with MR, all of these studies aim to provide more information with which to make a better clinical assessment, reducing the need for invasive techniques and hopefully preventing deaths from prostate cancer.

Chapter 3

Diffusion Modelling

Most studies into diffusion modelling use the traditional ADC model, a homogeneous, isotropic single compartment model. As used in multi-parametric MRI, this technique has proven effective in helping in the detection and grading of prostate cancer. However, if some of the disadvantages of these methods are to be addressed then these simplistic models could be replaced by ones that more accurately represent the microstructure of the prostate.

This chapter separately investigates two diffusion MRI models for better quantifying tissue microstructure in the prostate, describing the process of selecting the most appropriate models and then going on to test them in simulated datasets and in-vivo. One model uses a distribution of diffusivity values for each pixel to discover how heterogeneous the underlying tissue is. The second model looks to remove the effects of orientational heterogeneity in order to obtain more accurate diffusional anisotropy values.

These models are related in that they both try to obtain parameters related to microstructure from data at normal clinical resolution, using models on a single voxel instead of trying to increase resolution. They were chosen because they address two of the flawed assumptions made when modelling the Apparent Diffusion Coefficient, one that prostate tissue is made up of a homogeneous tissue that has a single diffusivity value and one that that tissue is diffusion isotropic. Furthermore, both models have

shown promise in other regions of the body and both introduce a small number of extra parameters in order to better model the tissue without requiring a large amount of data to perform accurately.

3.1 Distributed diffusivity model

Diffusion MRI is an established method for modelling microstructure in the prostate. Many intricate models have been developed that try to link diffusion MRI to the histology of the prostate but ADC is still the most popular and is the only one used in the clinic. This is because all of the extra detail that more complex models provide are counteracted by the fact that they require more data to produce robust and repeatible results. Therefore, the idea proposed in this section stems from the fact that a model that finds a middle ground between the simple and repeatible ADC and the intricate and microstructurally-informed models, such as VERDICT, could be more effective than either.

The proposed model is a distribution of diffusivities modelled as a single-Gaussian which produces both a mean and a standard deviation value for diffusivity rather than just a mean value, adapted from past work in the brain [126]. This only adds one extra parameter to the ADC model so it should not require the large amounts of data that some diffusion models need to work effectively. The theory behind this model is that within prostate tumours there is a proportion of the signal originating from the water in the glandular lumen. However, in a prostate tumour the luminal high-diffusivity signal component will be lower as the malignant cells grow into these lumen. It is proposed that within PCa a larger proportion of the overall diffusion-weighted signal will originate from only one compartment as, when tissue becomes more cancerous, the signal will be dominated by one effective compartment that is a combination of both epithelial and stromal cells. Hence the overall tissue heterogeneity, modelled by the standard deviation of the distribution of diffusivities of the pixel, will be reduced. This model does assume that there is only one tissue type within the prostate but takes into account the overall intra-voxel tissue heterogeneity with the variance variable.

A study was designed to investigate the hypothesis that a three-parameter Gaussian distributed diffusivity model would be able to successfully discern between cancerous and normal tissue and provide a better fit to diffusion data than ADC.

3.1.1 Methods

Four men awaiting biopsy for suspected prostate cancer were recalled for diffusion imaging within two weeks of their multiparametric prostate MRI (mp-MRI) on a Philips Achieva 3T MRI scanner using a 32-channel cardiac coil. The diffusion weighted MRI was performed in the axial plane using a pulse gradient spin echo sequence, acquired using single-shot echo planar imaging. Three orthogonal directions were used for each of the 6 b values (0, 90, 500, 1500, 2000 and 3000 $s\ mm^2$). The other parameters of the DWI were: TR/TE = 2000-3707/50–80 ms; voxel size = 1.3x1.3x5mm^3; slice thickness = 5mm; number of slices = 14; field of view (FOV) = 220 x 220 mm^2; matrix size = 176 x 176; scan duration = 12m 57s.

A radiologist manually contoured a circular region of interest (ROI) on the ADC map at the suspected epicentre of the tumour and the same ROI was copied onto a normal region of the prostate (as confirmed by previous mp-MRI) in the same zone on the same slice. All lesions were histologically confirmed as prostate cancer following targeted biopsy.

Two separate models were fit to the signal intensity data. The apparent diffusion coefficient (ADC) of the tissue was calculated in order to compare the new method to existing models. Then the distributed diffusivity model was calculated from the signal intensities using the equation:

$$S = S_0 \int_0^\infty P(D) e^{-bD} dD \tag{3.1}$$

Using the Gaussian distribution to model the diffusion spectrum as:

$$P(D) = (2\sigma^2\pi)^{-1/2} e^{-(D-D_m)^2/2\sigma^2} \tag{3.2}$$

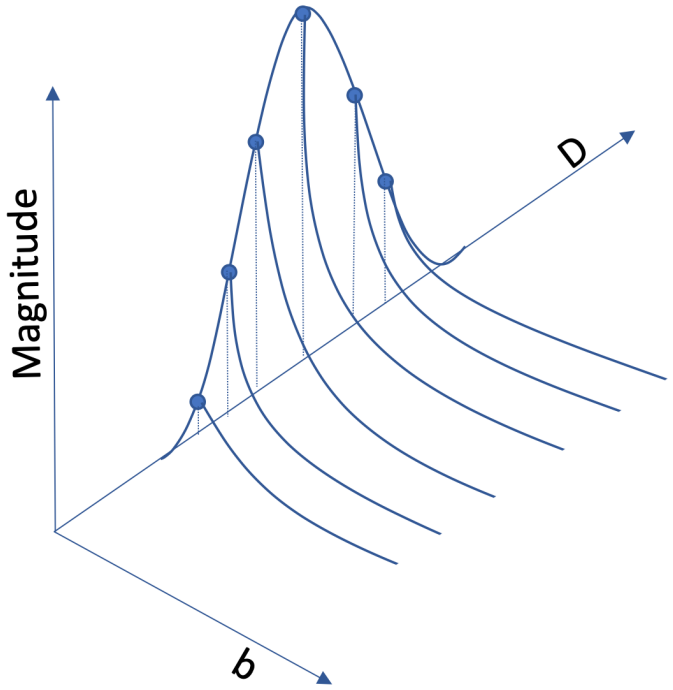
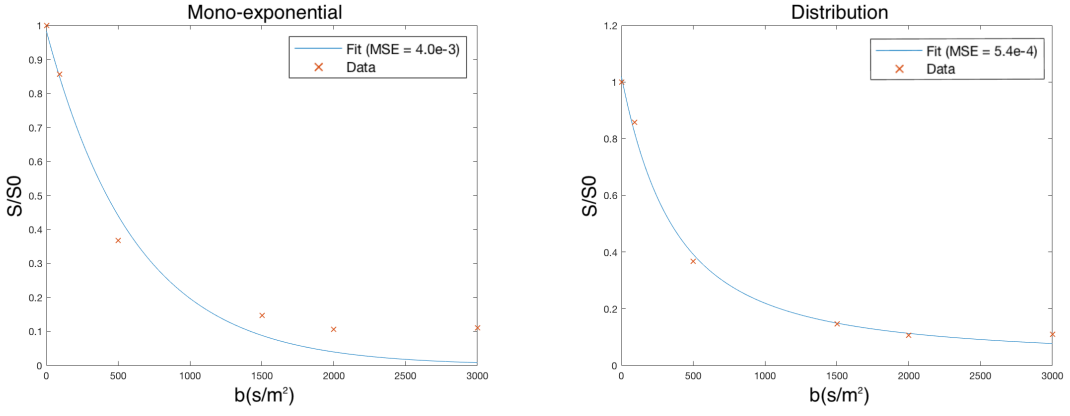


FIGURE 3.1: A graph showing the Gaussian diffusion distribution along the b=0 axis. The overall signal is the sum of all of the exponentials, as explained in Equation 3.1.

This distribition was estimated for each voxel, producing values for mean diffusivity (D_m) and standard deviation (σ) using a constrained nonlinear least squares optimization in a custom-written Matlab tool. It is expected that as the volume of luminal space within a pixel changes, the variance parameter of this model will change.



(a) (b)

FIGURE 3.2: Comparison of mono-exponential and distribution models in a single voxel.

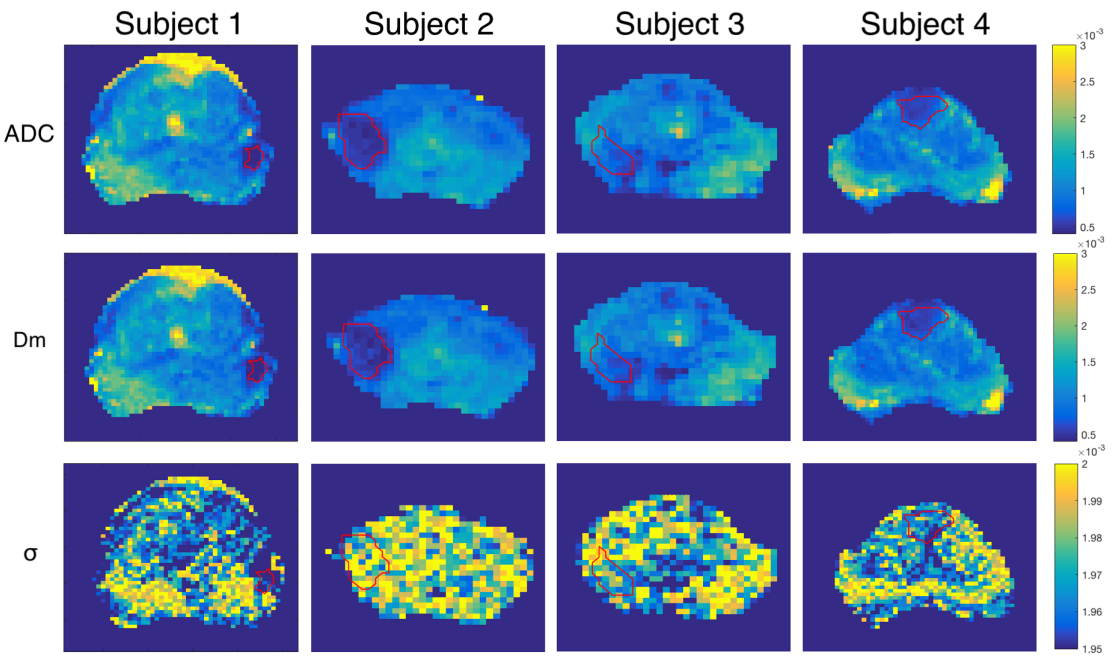


FIGURE 3.3: Maps of ADC, Mean Diffusivity (D_m)of Distribution and Standard Deviation (σ) of Distribution for all three patients. The tumour is outlined in red.

	ADC (x10^{-4})		D_m(x10^{-4})		σ(x10^{-4})	
	Mean	Range	Mean	Range	Mean	Range
Tumour	5.98	5.11-7.72	8.89	7.26-10.71	5.04	2.02-6.95
Normal	8.70	7.47-9.38	19.15	16.21-21.76	9.26	7.28-10.94

TABLE 3.1: Mean values and the range of values of ADC, D_m and σ for both tumour and normal tissue

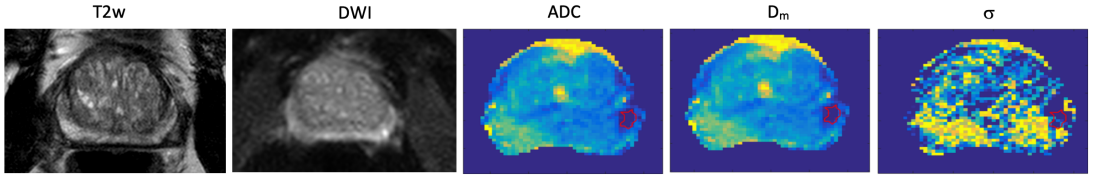


FIGURE 3.4: Maps of T2w, DWI, ADC, Mean Diffusivity of Distribution (D_m) and Standard Deviation of Distribution (σ) for one patient. The tumour is outlined in red.

Figure 3.1 is a visual representation of Equation 3.1, showing how the diffusion distribution $P(D)$ and exponential diffusion signal decay e^{-bD} combine to form the overall signal intensities. Maps of ADC, mean diffusivity (D_m) and standard deviation (σ) were produced for selected slices of interest within each acquisition. In order to analyse the ability of this distributed diffusivity model to differentiate between regions of normal tissue and tumour, a Wilcoxon rank sum test was carried out, where a value of $p < 0.05$ was taken to be significant.

3.1.2 Results and Discussion

All lesions were confirmed as Gleason 3+4. The fact that all of the lesions had the same Gleason score could introduce bias, but, as a preliminary investigation on a small number of patients, the results of this study gives insight into how the model might perform in larger studies. The ROIs were of size 46, 235,128 and 75 pixels respectively for each of the four subjects. Table 3.1 shows the mean values and the range of values of ADC, D_m and σ for both tumour and normal tissue. Figure 3.4 shows images of the same prostate in five different maps for comparison.

Maps of D_m, σ and ADC were obtained, with most of the voxels within the prostate obtaining a fit that was within the bounds set for the optimisation. Figure 3.2 shows the fits of both models for a typical voxel in an area of normal prostate tissue along with the mean-squared error, confirming that the distributed diffusivity model produces a closer fit to the data. Wilcoxon rank sum tests showed p-values of p=0.00031 for D_m, p=0.00062 for σ and p=0.00025 for ADC, which are below p=0.0167 (the Bonferroni-corrected p-value for a 95% significance), suggesting that all are able to distinguish between tumour and normal tissue in the patients studied. Not only do these results show that D_m and σ can be used to distinguish tumour from normal tissue, the distributed diffusivity model provides a better fit to the data than the mono-exponential, as shown by the Mean Square Error values in Figure 3.2. Though this is expected when the number of fitting variables is increased, the fact that the error is reduced by a factor of 10 is significant.

The parameters of the model are shown in Figure 3.3. It can be seen that in the PZ both D_m and ADC are higher, showing that water molecules there are on average less restricted, which ties in with the fact that the normal PZ is known to have a higher luminal partial volume than the rest of the prostate. It is also interesting that σ is raised in the PZ, tying in with the fact that the smaller luminal volumes in the TZ mean that the tissue is more homogeneous there. One limitation of the current method is that the maps for σ in Figure 3.3 look noisy, which is likely to be down to the diffusion acquisition. To rectify this issue further work should aim to optimise the dataset in order to maximise the performance of the model.

Though this study only contains four subjects, it has shown that in these subjects that, based on the p-values of the statistical test carried out, the diffusion distribution model is at least as effective as ADC in its ability to distinguish normal tissue from Gleason 3+4. The small number of subjects and the fact that each tumour has the same Gleason grade means that this experiment is limited. In order to fully investigate this technique a larger study validated by histology must be carried out on tumours with a wider range of Gleason grades. As a further line of research the correlation between these parameters and histology could be assessed in order to probe the relationship of the model to underlying tissue microstructure. For example, the value of σ appears to fall in areas of tumour in Figure 3.3 when compared with normal tissue areas in the same zone of the prostate, which could correspond with the loss of luminal and ductal space [20]. Though anatomically these current findings seem correct, the relationship between this σ value and the microstructural heterogeneity of the tissue should be investigated further. The model could also be extended based on radiological characteristics, for example to account for multiple tissue microenvironments.

3.2 Spherical Mean Technique

Parallel to the experiments looking into a single Gaussian diffusivity model, another diffusion model was investigated. This was motivated by the fact that current methods in the prostate are not able to accurately measure diffusion anisotropy. Diffusion anisotropy is a tissue characteristic that has proven useful in examining tissue in the brain but is underutilised in the prostate. Bourne et al. [20] used high resolution diffusion images (40x40x40μm^3)to show that diffusion anisotropy is present in the prostate.

The prostate contains stromal cells that are thought to be highly diffusion anisotropic but previous work [20] has shown that as the volume of a voxel increases the mean FA for that voxel falls. This shows that high resolution images can detect this anisotropy but in larger voxels the effect of orientational heterogeneity reduces the overall voxel anisotropy.

The Spherical Mean Technique (SMT), as described in Chapter 2, is used in regions of the brain where fibres cross. It was designed to find the diffusion anisotropy of the underlying fibres when the individual fibres are oriented in many directions. This problem has similarities to the prostate in that on a very small scale anisotropy is present but within a larger diffusion MRI voxel this effect is lost. Hence this model is appropriate to the prostate as it could remove the effect of the orientational heterogeneity of the stromal tissue. The difference between normal tissue and tumour should then be more significant as the overall voxel FA is expected to increase with a decrease in the fractional volume of glandular lumen.

3.2.1 Background

The SMT estimates parallel and perpendicular diffusion coefficients whilst removing the effects of orientation. The normalised diffusion signal e_b for gradient g in direction ω is:

$$e_b(g) = \int_{S^2} h_b(g, \omega) p(\omega) d\omega \tag{3.3}$$

where S^2 is a unit sphere, $h_b(g, \omega)$ is the average impulse response function for a microscopic section of tissue where orientation is not a confounding factor and $p(\omega)$ is the orientation distribution of the cells within a voxel. The spherical mean $\bar{e}_b$ of the diffusion signal over all gradient directions is:

$$\bar{e}_b = \frac{1}{4\pi} \int_{S^2} e_b(g) dg. \tag{3.4}$$

Through substitution of these two equations it can be shown that the spherical mean of the impulse response function $\bar{h} = \bar{e}$. This means that the arithmetic average of the signals in all directions is equal to the spherical mean of the impulse response function. Hence by varying the parameters of the response function the best fit to the average of the signals of all gradient directions can be estimated.

In order to be able to constrain this problem a model for the respone function must be chosen. In this work and in previous works the axially symmetric "zeppelin" shape is chosen, where the response function is described by two parameters, the parallel

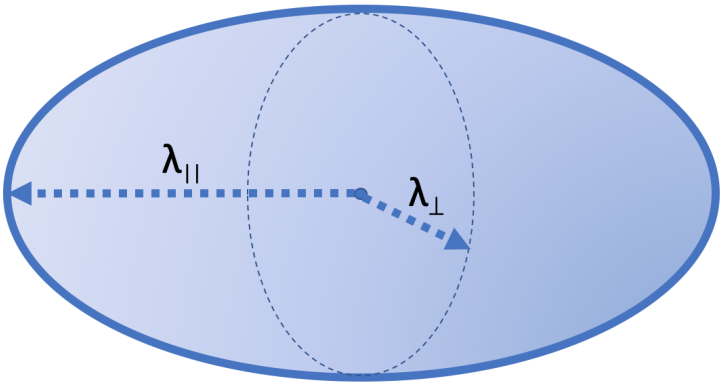


FIGURE 3.5: Zeppelin response function, with the parallel and perpendicular diffusivities ($\lambda_{\parallel}$ and $\lambda_{\perp}$) labelled. The zeppelin shape is axially symmetric which reduces the number of parameters when compared to a normal tensor.

diffusivity ($\lambda_{\parallel}$) and the perpendicular diffusivity ($\lambda_{\perp}$). A representation of the zeppelin response function is shown in Figure 3.5.

In mathematical form the zeppelin function is:

$$h_b(g,\omega) = exp(-b\langle g,\omega\rangle^2\lambda_{\parallel})\ exp(-b(1-\langle g,\omega\rangle^2)\lambda_{\perp}) \tag{3.5}$$

where $\langle g,\omega\rangle$ is the difference between the gradient direction and orientation direction. From this the spherical mean is calculated as [81]:

$$\bar{h}_b(\lambda_{\parallel},\lambda_{\perp}) = \frac{\sqrt{\pi}}{2}e^{-b\lambda_{\perp}}\frac{erf\left(\sqrt{b(\lambda_{\parallel}-\lambda_{\perp})}\right)}{\sqrt{b(\lambda_{\parallel}-\lambda_{\perp})}} = \bar{e}_b(\lambda_{\parallel},\lambda_{\perp}) \tag{3.6}$$

In order to find the values of $\lambda_{\parallel}$ and $\lambda_{\perp}$ that best fit the data an optimisation called the interior-point method was used, which is designed to solve non-linear convex problems.

In previous works, a parameter called the Fractional Anisotropy (FA) was of interest as it combines $\lambda_{\parallel}$ and $\lambda_{\perp}$ into one measure of anisotropy. Fractional anisotropy for a 3-dimensional tensor is defined as:

$$FA = \sqrt{\frac{1}{2}}\frac{\sqrt{(\lambda_1-\lambda_2)^2+(\lambda_2-\lambda_3)^2+(\lambda_3-\lambda_1)^2}}{\sqrt{\lambda_1^2+\lambda_2^2+\lambda_3^2}} \tag{3.7}$$

Which, using the zeppelin model, simplifies to:

$$FA = \frac{\lambda_{\parallel} - \lambda_{\perp}}{\sqrt{\lambda_{\parallel}^2 + 2\lambda_{\parallel}^2}} \tag{3.8}$$

An experiment was designed to investigate whether the SMT could provide anisitropy values that better resembled the tissue microstructure and whether this could then be used to detect and grade regions of interest. The ability of the $\lambda_{\parallel}$, $\lambda_{\perp}$ and FA parameters to categorise ROIs into both groupings of PI-RADS v2 scores and histologically informed normal tissue and tumour were compared to that of ADC.

3.2.2 Methods

Retrospective data, which was acquired for VERDICT modelling in a larger prospective study [78], was used for 65 patients aged 49-79 . The acquisition is a PGSE single shot EPI on a Philips Achieva 3T MRI scanner. DWI was acquired with Single Diffusion Encoding (SDE) single shot EPI sequences over 6 b-values (0, 90, 500, 1500, 2000, 3000 s/mm^2) in 3 orthogonal directions. TR/TE = 2000-3707/50-80 ms; Field of View = 220x220 mm; Voxel size = 1.3x1.3x5mm; Scan duration = 12:57. ROIs were drawn by a board-certified radiologist in regions of both healthy tissue and suspected lesion using mp-MRI images taken as part of the acquisition. In all, 98 ROIs were contoured for this study. After the ROIs were contoured, a radiologist then assigned each a PI-RADS v2 score based on the standard mp-MRI images acquired as part of the study. The SMT maps were processed as described in Section 3.2.1.

The ability of the median SMT parameter values of ROIs to differentiate between PI-RADS v2 score groupings was determined using a Receiver Operating Characteristic (ROC) analysis combined with 5-fold cross validation. Three comparisons were made in this way. PI-RADS 1,2 vs PI-RADS 3,4,5 aims to separate lesions needing further action from those that do not. PI-RADS 1,2 vs PI-RADS 3, and PI-RADS 3 vs PI-RADS 4,5, both investigate the SMT's ability to discern between negative (1,2), indeterminate (3)

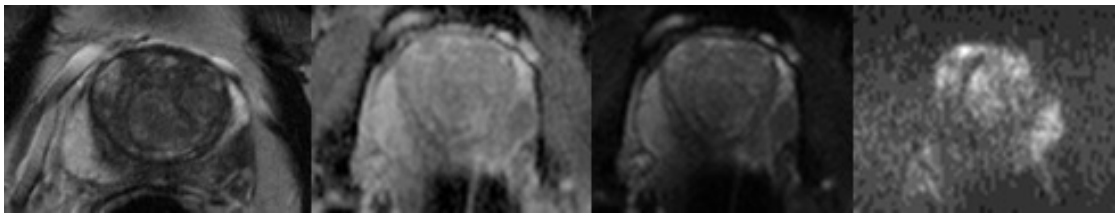


FIGURE 3.6: T2w, ADC, b=90 and b=2000 images for an example PZ tumour.

PI-RADS grouping	$\lambda_{\parallel}$	$\lambda_{\perp}$	FA	ADC
1,2v3	0.7126	0.6932	0.7049	0.7114
3v4,5	0.7417	0.4600	0.6267	0.5283
1,2v3,4,5	0.7404	0.7284	0.7468	0.7296

TABLE 3.2: AUC values when differentiating between three different PI-RADS v2 groupings.

and positive (4,5) disease. The mean values for sensitivity, specificity and area-under-curve (AUC) across the five-folds were computed. Sensitivity and specificity values were calculated from the ROC analysis using an operating point with the shortest distance to the point of perfect discrimination.

The ability of the SMT to discern histologically confirmed tumour from normal tissue was also assessed in a subset of 19 of these patients for which biopsy data was available. 30 ROIs were drawn in either normal regions or tumour using prior knowledge of the location of biopsy results (imaging was pre-biopsy). Values for sensitivity, specificity and area-under-curve (AUC) were also calculated.

3.2.3 Results

For the diffusion ROIs there were 30 PI-RADS 1, 32 PI-RADS 2, 18 PI-RADS 3, 7 PI-RADS 4 and 11 PI-RADS 5. Of the histologically examined ROIs 16 were found to be benign, 3 were Gleason 3+3, 6 were 3+4, 3 were 4+3, 2 were 4+4, 1 was 4+5.

Figure 3.6 contains images of an example prostate containing a tumour in the left PZ.

Figure 3.7 shows an example in the prostate of the FA map calculated using the SMT alongside the ADC map.

Table 3.2 shows the AUC values of each parameter when differentiating between the three different sets of PI-RADS v2 groupings. Table 3.3 shows the sensitivity and specificity

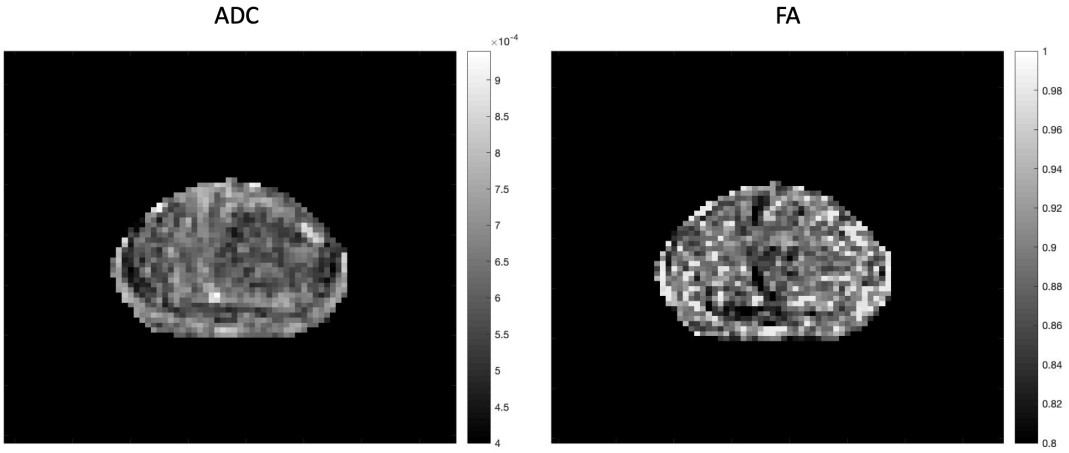


FIGURE 3.7: ADC and FA maps of an example healthy prostate.

	$\lambda_{\parallel}$		$\lambda_{\perp}$		FA		ADC	
PI-RADS	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.
1,2v3	86.0	73.8	82.0	68.6	90.5	65.7	81.7	72.5
3v4,5	90.0	75.0	66.7	63.7	96.0	59.7	70.0	62.0
1,2v3,4,5	88.5	67.4	89.2	63.9	83.4	68.8	86.7	66.5

TABLE 3.3: Sensitivity and specificity values (as percentages) when differentiating between three different PI-RADS v2 groupings.

	$\lambda_{\parallel}$	$\lambda_{\perp}$	FA	ADC
AUC	0.611	0.732	0.556	0.674
Sensitivity(%)	66.7	75.0	58.3	75.0
Specificity(%)	61.1	72.2	61.1	61.1

TABLE 3.4: Sensitivity, specificity and AUC values when differentiating between normal tissue and tumour as defined by histological analysis.

values for the same test. Table 3.4 shows the AUC, sensitivity and specifivity values for each SMT parameter when differentiating histologically confirmed tumour from normal tissue.

3.2.4 Discussion

The first point to note from the results is that the fractional anisotropy values calculated through the SMT are substantially higher than those seen in previous studies using DTI [18, 20, 54]. The effects of orientation distribution mean that a lower value for FA is expected from DTI but, as there is no ground truth as to what the diffusion anisotropy should be, the accuracy of these values cannot be quantified. However, the ability of

the SMT estimates to differentiate PI-RADS v2 scores is shown to be similar to that of ADC in this study, even when differentiating between indeterminate lesions and positive lesions (PI-RADS 3 v PI-RADS 4,5). This shows that the tensor produced by the SMT is providing more insight into changes in the tissue than ADC.

The histology results in 3.2 show that when simply trying to detect prostate cancer the $\lambda_\perp$ looks to be the most effective. Though this does not agree with the results from the PI-RADS study, which could well be down to the fact that the two tests were performed on different datasets, it still shows that the SMT model is more sensitive and specific to tumours than the traditional ADC model.

The limitations of this study were that the diffusion acquisition was not optimised for this model and that the dataset was not large enough to conclusively prove how useful the SMT is.

There are a number of ways in which to further investigate the ability of SMT to produce more accurate measures of microstructural anisotropy and more effective methods for classifying suspicious lesions. First, the number of b-values and the number of gradient directions should be further optimised. This study evaluated the method on 3 gradient directions, which is relatively few, and a larger number of b-values. In order for this technique to become clinically viable these values must be prospectively investigated and fixed. Then a multi-reader study into how repeatible these values are would be important, as a poor repeatibility means that a method cannot be rolled out to other centres and be effective.

3.3 Summary

This chapter set out to investigate two models for better quantifying tissue heterogeneity in the prostate and then utilise them in detecting and grading prostate cancer.

The models addressed different flawed assumptions that are made when using the Apparent Diffusion Coefficient model, one that the tissue within a pixel all has the same

diffusivity and the other that diffusivity is isotropic in the prostate. These two assumptions do not agree with what histology tells us about the microstructure of the prostate so the question is whether accounting for them can improve our ability to detect and grade prostate cancer.

This chapter has shown that even with the non-optimal datasets used, both models provide new and interesting information on the prostate and both show promise in being able to provide information that would be useful to clinicians when examining prostate cancer.

Chapter 4

T2 Modelling

In the past there has been an extensive amount of work done in the prostate looking into diffusion MRI for quantifying tissue microstructure. The diffusion measure that is currently used in routine clinical practice is ADC, chosen for its simplicity and robustness to noise. However, it does not provide quantitative information on the constituent compartments of the prostate tissue, instead giving an average diffusivity value across an entire pixel. Nor is ADC fully reliable in being able to give us information about the severity of a particular tumour [10]. EPI-based diffusion scans also suffer from distortion and signal pile up, which can lead to inaccurate results and sometimes render scans unusable.

The current cutting edge of clinical practice, multi-parametric MRI (mp-MRI) [151], utilises ADC, T2, and DCE MRI (as shown in Figure 4.1), acknowledging the fact that one modality isn't able to accurately inform a clinical decision with the techniques of

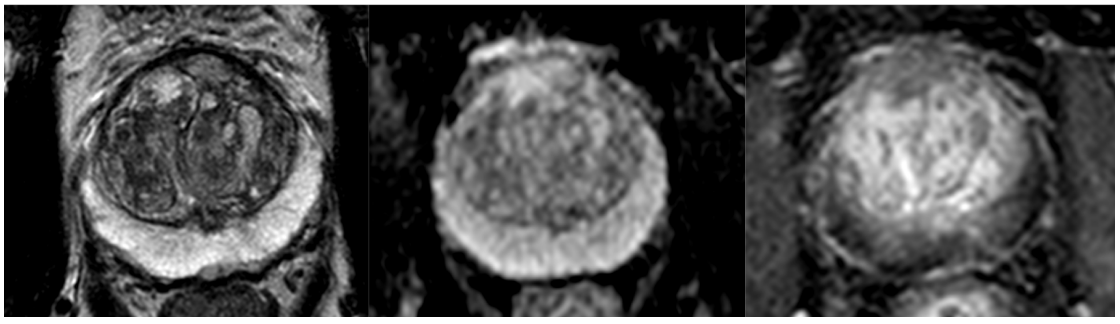


FIGURE 4.1: Example of the three components of multi-parametric MRI, axial T2, apparent diffusion coefficient (ADC) and dynamic contrast enhanced (DCE) images.

today. However, one of the issues with mp-MRI is that currently it has high sensitivity but low specificity, with over 40% of lesions classified as indeterminate. Furthermore, inter-reader agreement is highly variable, with one study having it as low as 55-65% [140]. These issues are the motivation behind looking for a quantitative technique that provides more infomation than simple ADC and T2 images in order to be able to distinguish tumour from normal tissue in more difficult cases.

This chapter looks into alternative methods for quantifying tissue microstructure by imposing models onto multi-echo T2 images instead. Whilst quantitative T2 modelling is not new in itself, its application to the prostate is a small but expanding field of research. Most studies into T2 in the prostate use the anatomical T2 weighted image, effectively a homogeneous single compartment model. This, like ADC, is simple but provides no direct quantification of the underlying tissue composition. Therefore, similar to the brain using Myelin Water Imaging (MWI) [101], this chapter investigates the possibility of using a multi-compartment model on multiple-echo T2 (ME-T2) data.

4.1 Model Selection

As described in Chapter 2, there are a number of models that have previously been applied to ME-T2 data, most examples of which are used in images of the brain. In choosing which model to use in the prostate, it must first be understood what microstructure is present within the tissue. ME-T2 imaging in the brain is designed to separate the myelin water compartment from the rest of the brain in order to detect changes in the thickness of the myelin sheath surrounding neurons. Myelin has a different T2 value to the rest of the brain so MWI separates the tissue into two compartments. When looking at the prostate, the three main tissue compartments are lumen, stroma and epithelia. Though the stroma and epithelia serve very different functions within the prostate they have similar T2 values in comparison to the large luminal space. This means that a two compartment model is also appropriate for the prostate, one quantifying lumen and one the stroma and epithelia combined. Blood vessels, the other main component of prostate tissue, are expected to make insignificant contributions to the overall signal[134, 136].

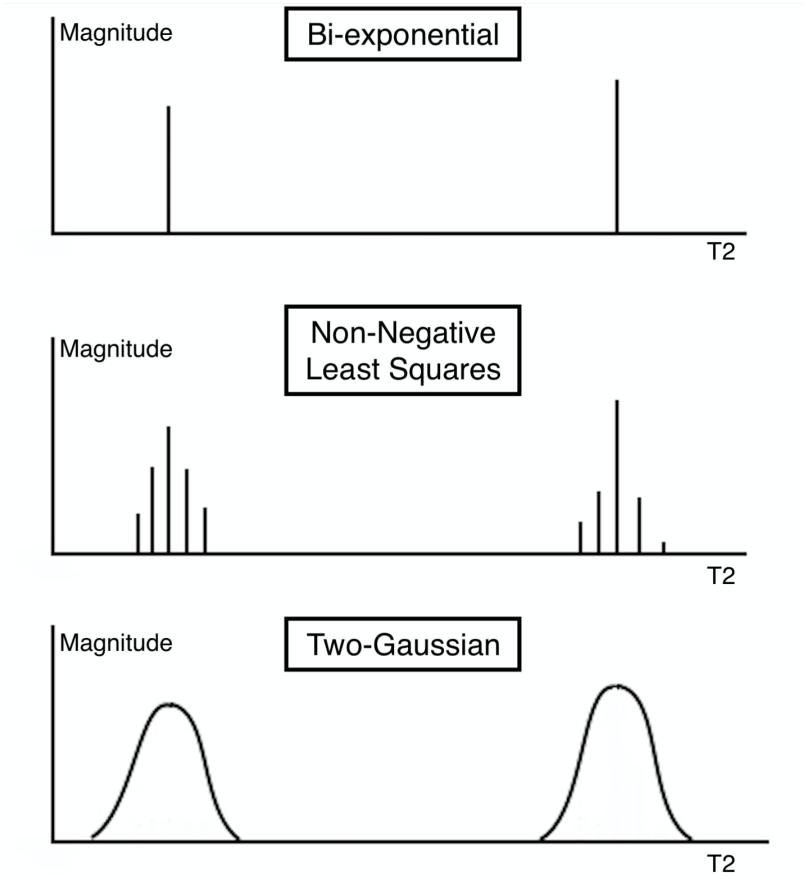


FIGURE 4.2: Simple diagrams representing a bi-exponential, NNLS and constrained two-Gaussian fit respectively for a two-compartment tissue.

In order to model the separate T2 compartments in the prostate, the choices are a bi-exponential model, a non-negative least squares (NNLS) fit [101] or a statistical model. The bi-exponential model has proven unable to distinguish the myelin compartment in the brain [51] so was ruled out when looking for the luminal component in the prostate.

Referring back to the literature on multi-compartment T2 modelling in the brain, most use an unconstrained NNLS algorithm [101, 134, 136]. This method fits an arbitrarily large number of exponentials to the signal decay curve which represent a distribution of T2 values. It has proven effective in the brain at measuring myelination and has also been used in the prostate [133, 135]. However, the problem with the NNLS method lies in the unconstrained nature of the fit. The large number of exponentials fit to the decay curve lead to an underdetermined problem where there are far fewer inputs than outputs in the model.

In order to assess the NNLS fitting method a preliminary experiment was designed to

test how many exponentials were needed in the NNLS model to properly detect two separate T2 compartments in prostate tissue. The prostate was acquired with a 64-echo sequence (echo spacing 25ms) and 20 pixels from healthy regions of interest in the PZ of each of 5 subjects were selected. Across all of the pixels it was necessary to fit at least 80 discrete T2 values between 0-1000ms to be able to find some degree of intra-compartment heterogeneity in a prostate tissue sample.

To have a robust fit it is necessary to have more datapoints than unknowns. Acquiring a large number of echoes per line of k-space in vivo leads to an impractical scan time. Hence a limitation of this fitting method is its underdetermined nature if used with a clinically plausible acquisition.

One way of addressing this is by using a regularisation term to smooth the NNLS peaks and make the estimates more reliable. By introducing terms such as an energy minimisation regulariser the delta-peaks that NNLS produces can be smoothed out and the optimisation can be made more robust. Using a regulariser to minimise the energy of the curvature of the spectrum, the optimisation becomes:

$$\chi^2 = min\left[\sum_{i=1}^{N}\left|\sum_{j=1}^{M} A_{ij}s_j - y_i\right|^2 + \mu\sum_{j=1}^{M}|s_{j+2} - 2s_{j+1} + s_j|^2\right]. \tag{4.1}$$

A is an NxM matrix containing values of $e^{-t_i/T_{2j}}$, s is an Mx1 vector of weights, y is the Nx1 vector of real signal values and μ is the smoothing factor. Larger values of μ lead to smoother and flatter T2 distributions whereas $\mu = 0$ gives the original NNLS fit.

The downside to using these regularisation terms is that the regulariser itself needs a weighting μ in the algorithm which may be difficult to determine. By increasing the weighting the peaks are smoothed more and the variance of the distributions increases so a meaningful value of intra-compartment heterogeneity is hard to attain, though the areas of the peaks might be better preserved.

Despite the underdetermined nature of the NNLS method the outputs in the prostate have consistently shown two compartments [57, 82, 152]. The reason so many methods use NNLS fitting is that it is unconstrained and can pick up subtleties that a more

constrained model might not. However, the fact that in the prostate it repeatedly produces two-compartment T2 distributions hints at the fact that, whilst it improves upon a bi-exponential model, the complexity and freedom added by the many exponentials in the NNLS fit is too much and a constrained model could be a better middle ground.

Since a distribution of T2 values obviously has an advantage over the bi-exponential and multi-exponential models, but the NNLS method is too unconstrained for the number of datapoints that can feasibly be collected, it makes logical sense to try to put in place a small number of constraints on the fitting of the T2 distribution that will remove the problem of the fitting being underdetermined without placing unduly strict assumptions on the tissue. There are a large number of constraints that can be put in place, such as constraining the range of T2 values of each compartment or the number of compartments.

The model that will be investigated in this study places two simple constraints on the fit. One is that there are two T2 compartments to be found within the prostate which, as has been discussed previously, is an accepted assumption given NNLS fitting results [134, 136].

The second assumption placed on the model is on the shape of the T2 distribution for each of the compartments. In NNLS fits each of the compartments is distributed around a mean value with a quick drop off either side. The two candidates for this shape were Gaussian and Gamma distributions, described respectively by the equations:

$$\mathcal{N}(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right) \tag{4.2}$$

where μ is the mean of the distribution and σ^2 is the variance, and

$$\Gamma(x|a,b) = \frac{1}{\Gamma(a)b^a} x^{a-1} e^{-\frac{x}{b}} \tag{4.3}$$

where a is the shape parameter, b is the scale parameter and $\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}dx$. The Gamma distribution has the benefit of being zero for all values of x<0, unlike the Gaussian.

For this work, the choice between the Gamma and Gaussian distributions was made based upon the assessment of the mean squared error of five in-vivo prostate datasets. By fitting both Gaussian and Gamma distributions to the signal and comparing the mean squared errors for each this experiment found that on a pixel by pixel basis there was no statistically significant difference, using a Kruskal-Wallis test, between the LWF values produced. This means that there was little advantage of one over the other. As it was commonly used in the literature, and due to the fact that it was mathematically easier to implement, it was decided that the Gaussian distribution would be included in the model.

These two restrictions lead to the T2 distribution being modelled as two Gaussians, reducing the number of output variables from 80-100, depending on the number of exponentials used in the NNLS fit, to 6. The T2 distribution was modelled as:

$$p(T_2) = M_0\left(\frac{\alpha}{\sqrt{2\pi\sigma_1^2}}exp\left(-\frac{(T_2-\mu_1)^2}{2\sigma_1^2}\right)+\frac{1-\alpha}{\sqrt{2\pi\sigma_2^2}}exp\left(-\frac{(T_2-\mu_2)^2}{2\sigma_2^2}\right)\right) \tag{4.4}$$

where M_0 is the overall magnitude of the signal, α is the relative height of the short T2 peak in relation to the long peak, μ_1 and μ_2 are the means of the two peaks and σ_1 and σ_2 are the standard deviations of the two peaks. This reduction in the number of output variables makes the model much easier and quicker to fit than NNLS.

The start points of the constrained distribution were established by using the results of a simple biexponential fit. The biexponential fit should give a good start point for the means of the two distributions. The mean T2 values are constrained to be above zero and $\mu_1 < 200ms < \mu_2$.

4.1.1 Comparing constrained and unconstrained fitting

The original NNLS fitting is unconstrained and allows for any number of compartments but it can lead to an underdetermined problem. The constrained two Gaussian model is more stable but the implicit assumptions must be justified. A comparison of the two would give an insight into which provides more information on the tissue microstructure of the prostate.

4.1.1.1 Simulation

A simple method for comparing two tissue models is to create simulated signals based on assumptions about the underlying tissue, add noise and then process these as individual pixels, comparing the output values with those that were originally used in the tissue simulation. Simulations are designed to be a quick and easy way of testing prior assumptions and getting a first approximation to the effects in real tissue.

First, a number of simulation tissues were created, each with different microstructural compositions, so that the fitting methods could be tested accurately across the whole range of values that would typically arise in the prostate. Two underlying tissue simulations were tested, one created using a two Gaussian T2 distribution and the other using a two delta-function T2 distribution, where a delta function is an infinitely thin peak at a single value of T2. This was done to make the comparison a fairer one given the fact that a model fitting distributions that are of the same form as those used to create the simulated tissue will naturally fit the data better. Using values from prior literature [82, 152], tissue simulations were created for each of 25, 50 and 75 ms for the short T2 component and 300, 550 and 800 ms for the long T2 component. From these tissue simulations a signal decay was created using the equation:

$$S(TE) = S_0 \int_0^\infty p(T_2) exp(-TE/T_2) dT_2 \tag{4.5}$$

where $p(T_2)$ is the simulated T2 distribution. The distribution for a two Gaussian is described in Equation 4.4 and a two delta-function distribution is described by:

$$p(T_2) = M_0\big(\alpha\delta(\mu_1) + (1-\alpha)\delta(\mu_2)\big) \tag{4.6}$$

where $\delta(\mu)$ is a delta function at a T2 value of μms.

Then the simulated signal was sampled using 64 echoes and an echo spacing of 25ms, as has already been seen in previous works in the prostate [134, 136]. Then, in order to make the simulated signal more realistic, white Gaussian noise with a specified Signal to Noise Ratio (SNR) of 100 was added to the signal in order to mimic the random noise

	Simulated LWF		
Fit	0.2	0.1	0
Constrained	0.1995	0.1123	0.0021
Unconstrained	0.1980	0.0970	0.0004

TABLE 4.1: LWF estimates using a two-delta tissue model. These are mean values across varied values of μ_1 and μ_2.

	Simulated LWF		
Fit	0.2	0.1	0
Constrained	0.2006	0.1122	0.0010
Unconstrained	0.1976	0.0971	0.0004

TABLE 4.2: LWF estimates using a two-Gaussian tissue model. These are mean values across varied values of μ_1 and μ_2.

that is present in all MRI images. Noise in modulus images is Rician but here, because of the good SNR, it can be approximated as Gaussian. In order to compare the two modelling methods the mean LWF value was calculated for each model over the given tissue types. The LWF of the model was used for comparison as this has been shown previously to be the best parameter of the LWI model for differentiating between normal tissue and tumour [134].

The results of this experiment are shown in Table 4.1. It shows the mean LWF values for the simulated tissue across different values of μ_1 and μ_2, calculated using LWI from a 64-echo simulated signal. These values show that, when using either a two-Gaussian or two-delta tissue model, there is little difference in the accuracy of the LWF estimate when using either the constrained (two-Gaussian) fit or the unconstrained (NNLS) fit.

This section of work has been carried out in order to inform the modelling technique that should be used in the continuing study of LWI. This simulation study has shown that there is no benefit to using a NNLS fitting in estimating the LWF. There are a number of downsides to using the NNLS fit, not least the time it takes to fit a single voxel when compared to the two-Gaussian method, so this experiment has confirmed that constraining the model of the tissue T2-distribution is both appropriate and advantageous.

4.1.1.2 Phantom

In order to further investigate whether the NNLS fit can be replaced by a constrained two-Gaussian fit, a phantom study was carried out. In this study a set of Eurospin TO5 gel phantoms were imaged with a 64-echo T2 sequence with TE values from 10-40ms. The phantoms chosen for the short and long T2 compartments were 50ms and 373ms respectively. This is because previous work in the prostate [136] considered anything below 200ms to be stroma and epithelia and anything above 200ms to be glandular lumen. Though this long T2 value used here is slightly lower than some previous mean estimates, the range of values is large and this combination allows us to test on the more difficult cases.

Much like the phantom study conducted by Storås et al. [152], the artificial combination of two phantom signals was carried out in order to mimic the composition of the two main T2 compartments found in the prostate. For this study, identical ROIs were drawn in the 50ms and 373ms phantom regions. Then the corresponding pixels in each of the ROIs were normalised to the magnitude of the initial echo, combined with a ratio of 20% long T2 to 80% short T2, and then multiplied by the mean of the magnitudes of the two corresponding pixels. The combination process was carried out in this way because the ratio of the signals would be incorrect if the signals were not normalised due to the fact that in the model the ratio is of normalised signals. This 80%-20% split is intended to mimic a LWF value of 0.2. This combination of datasets resulted in 103 signal decays. Using both the NNLS fitting and the two Gaussian model described in Section 4.1, the signal decays were fit and model parameters produced. The LWF values calculated using each method were then compared for each pixel.

Table 4.3 shows the mean estimated LWF values for a variety of TE values using both the NNLS and two-Gaussian fit. It shows that the estimates for LWF are little different for most echo spacings when using different echo train lengths. Interestingly, none of these results are particularly close to the 'actual' value of 0.2 but this may be down to the fact that it is diffucult to use a combination of separate signals to mimic tissue accurately.

		TE(ms)								
NE	Fitting	10	14	18	22	26	30	34	38	40
64	NNLS	0.137	0.147	0.139	0.134	0.133	0.119	0.114	0.119	0.061
	2-Gauss.	0.146	0.151	0.141	0.136	0.134	0.119	0.110	0.096	0.042
32	NNLS	0.083	0.121	0.126	0.125	0.129	0.116	0.109	0.103	0.069
	2-Gauss	0.087	0.134	0.133	0.133	0.135	0.120	0.112	0.107	0.071
16	NNLS	0.051	0.073	0.072	0.068	0.101	0.090	0.089	0.089	0.062
	2-Gauss	0.072	0.083	0.099	0.098	0.124	0.107	0.101	0.105	0.082
8	NNLS	0.000	0.013	0.039	0.048	0.070	0.055	0.055	0.050	0.045
	2-Gauss	0.001	0.063	0.072	0.070	0.074	0.060	0.060	0.056	0.050

TABLE 4.3: LWF values for each fitting method using 64-echo sequences with various TE values (Ground truth LWF = 0.2).

The results also show that TE values that are either too short or too long do not produce accurate LWF estimates. This is because a short TE will not adequately measure the long T2 signal components, such as the luminal space, whereas a long TE will not adequately measure the short T2 signal components, such as epithelia and stroma.

Most importantly, these results further back up the hypothesis that the NNLS fitting method can be replaced by the two-Gaussian method without a substantial loss in our ability to estimate the fractional volume of lumen within prostate tissue. As long as the echo spacings are chosen correctly there is little difference between the two in this phantom dataset.

Combining the results of the simulation and phantom studies, the conclusion of this section of work is that the NNLS fitting method performs little better in estimating the LWF than the two-Gaussian model. Combined with the fact that the NNLS method takes longer to fit, these results suggest that the NNLS fitting should be replaced by the two-Gaussian fit in further studies.

4.2 Sequence Design

As with all MR imaging, the pulse sequence used when generating the data to model the underlying tissue is important. In order to fit a complex model there needs to be a very careful choice made about the pulse sequence used to acquire that data. This section of the work aims to optimise the acquisition to a point where it can be used on patients

within a reasonable amount of time and produce good results when used in conjunction with the LWI model.

4.2.1 Type of sequence

First, there is a variety of different types of sequence to use. The simplest is to repeat the Spin-Echo sequence for all TE values and points in k-space. A quicker way to acquire multiple echoes is to use the MESE sequence in which for each 90° excitation a series of echoes are acquired using a train of re-focussing pulses, with the phase encoding constant across a train. Another way to attain the same results would be to use a steady state sequence much like that used in the mcDESPOT method described earlier. This is a quicker acquisition that circumvents the imperfect refocusing of other multi-echo methods. A recent work also utilised a modified 3D GRASE sequence in acquiring data for LWI [30], showing that in general the accuracy of the estimates was similar to the original LWI method provided SNR>100 and LWF<0.5.

The mcDESPOT model was ruled out due to the complexity of the acquisition and modelling and the fact that the main interest of using mcDESPOT in the brain is because it can take into account water exchange between different cell compartments, which we can assume to be negligible between the lumen and the rest of the tissue due to the lumen's large size relative to the surrounding cells [56]. Using separate Spin-Echo sequences would take a significantly longer time than MESE with little information gained so, whilst this is very easy to implement and has a high SNR, it is unrealistic to use clinically. Hence MESE was chosen as the pulse sequence that would be carried out in the rest of the studies.

4.2.2 Parameters of sequence

In acquiring a MESE scan the two main parameters over which to optimise are the number of echoes and the echo spacing. Both affect the range of T2 values that are sampled during the scan. This is important because each tissue component, whether it be lumen, stroma, or epithelia, has a different range of T2 values. When looking for

multiple compartments within the tissue, the correct range of T2 values must be sampled. If the echo spacing is too small then the longer components will not be adequately investigated (for a fixed number of echoes) whereas if the echo spacing is too large then the shorter components will not be sampled with enough datapoints to model reliably. When choosing the number of echoes, more will theoretically always be better but this is limited by the imperfect refocusing in scanners and comes with a longer scan time as a result of the increased TR due to scanner limitations on how much energy can safely be absorbed by the tissue in a given time. The resolution of the image is also important because there is a trade off between the resolution of the image, the number of echo times that are sampled and the scan time.

4.2.2.1 Simulation

In order to ascertain the combination of echo number and echo spacing that best investigates the prostate, first a simple simulation was used. When choosing an ME-T2 sequence to probe the range of different microstructures in the prostate a sequence needs to be devised that is a compromise between sampling the long and short T2 compartments adequately.

In this experiment a simulated signal using a 2-Gaussian distribution with ground truth parameters was created using a two Gaussian T2 distribution. Three different mean short T2 values (μ_1, 25-75ms), three mean long T2 values (μ_2, 300-800 ms) and three LWF values (0-0.2) were tested. The range is based upon previous work looking into ME-T2 imaging in the prostate [136]. In this experiment the range 0-20% was used because it has been seen in other MRI studies [136, 152]. White noise was added to the simulated signal with a Signal to Noise Ratio (SNR) value of 100, a representative value from the scanner used.

In order to test which acquisition was best across all of the tissue simulations, the simulated signals were sampled using a number of different schemes. Values of both echo spacing and the number of echoes were varied between 15-40ms and 32-64 echoes respectively, values based upon practical considerations of the scanner and the underlying tissue being examined. To each of these acquisitions a two Gaussian model was fitted.

	Echo Spacing (ms)					
LWF	15	20	25	30	35	40
0	0.0074	0.0060	0.0045	**0.0037**	0.0110	0.0110
0.1	0.1255	0.1260	**0.1248**	0.1286	0.1277	0.1358
0.2	0.2378	0.2380	0.2290	**0.2278**	0.2344	0.2679

TABLE 4.4: Mean LWF values for simulation using 32-echo scan.

	Echo Spacing (ms)					
LWF	15	20	25	30	35	40
0	**0.0030**	0.0135	0.0033	0.0165	0.0314	0.0033
0.1	0.1172	**0.1043**	0.1190	0.1219	0.1250	0.1325
0.2	0.2231	**0.2144**	0.2190	0.2308	0.2294	0.2349

TABLE 4.5: Mean LWF values for simulation using 64-echo scan.

The mean LWF across all values of μ_1 and μ_2 was then calculated in order to reveal how different echo spacings and underlying LWF values affect the fit. Whichever echo spacing provides the best results over the range of parameters will be a good first estimate of the optimum sequence.

The results of this experiment are shown in Tables 4.4 and 4.5. The closest estimates (shown in bold) for each value of LWF show that for a 32 echo sequence the best echo spacing to optimise the sequence over all values of underlying LWF is 30ms and for a 64 echo sequence it is 20ms. A spacing of 30ms would give a 32 echo sequence a range of 960ms which is long enough to explore the long T2 components without being too large to miss the small components. It makes sense that the optimum echo spacing of the 64 echo sequence would be lower, reflecting the fact that both sequences acquire over a similar range of T2 values when using the optimum spacing.

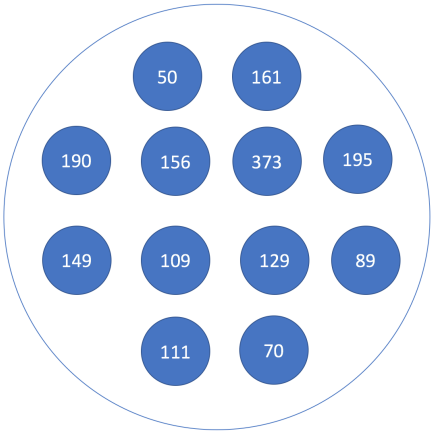
Comparing the 32 and 64 echo results, it can be seen that there is little benefit to MSE and LWF accuracy in using 64 echoes. Considering that the 64-echo scans are more difficult to acquire, since they are not readily available on most scanners, it is apparent that a 32-echo sequence will be adequate for a constrained version of LWI.

4.2.2.2 Phantom Study

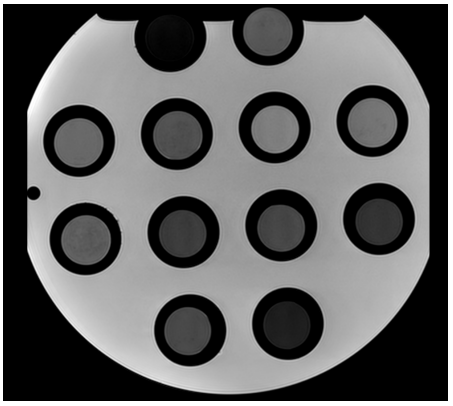
A further experiment was designed in which a number of these combinations were tested using a tissue phantom. A frame containing twelve Eurospin TO5 gel phantoms with

varying T2 values (ranging from 50-373ms) was scanned using a 64-echo sequence in which the echo spacing was varied. The echo spacing (TE) was tested for nine values in the range 10-40 ms. The lower bound was set to 10ms as below this the scanner would not explore the longer compartment and the higher bound was set to 40ms because the shortest T2 compartment that was expected in the prostate had been previously calculated to be 43-94ms [82, 152], meaning that an echo spacing of more than 40ms would struggle to accurately probe the short T2 microenvironment. As in Section 4.1.1.2, one phantom had an intrinsic T2 value of 50ms and the other 373ms.

Figure 4.3(a) shows the arrangement of the phantom tubes and their respective T2 values whilst Figure 4.3(b) shows the signal of a single echo at TE = 286ms to give a feel for image resolution and to show the varying signals for different phantom tubes. These experiments were carried out at room temperature. The temperature of the room is estimated to be between 20°C and 22°C, but was not measured at the time of the experiment.



(a) Reference T2 values in milliseconds calculated using monoexponential fit to data.



(b) Axial T2 image at TE = 100ms.

FIGURE 4.3: Images of phantom experiment.

Acquiring 64 echoes required a software patch to be applied to the scanner prior to the scan and removed afterwards. The purpose of assessing the number of echoes was to decide whether the 64 echoes required by the current LWI, using the unconstrained NNLS fit, could be reduced. Lowering the number of echoes to 32 would make the

TE (ms)	10	14	18	22	26	30	34	38	40	Ref.
μ_1 (ms)	62.9	53.5	51.3	51.0	**49.3**	53.0	50.7	52.2	59.1	50
μ_2 (ms)	1887.2	422.3	388.3	378.2	350.6	366.5	357.7	**373.9**	663.9	373
LWF	0.087	0.134	0.133	0.133	**0.135**	0.120	0.112	0.107	0.071	0.2

TABLE 4.6: Results using 32-echo sequence.

TE (ms)	10	14	18	22	26	30	34	38	40	Ref.
μ_1 (ms)	57.5	52.0	**50.5**	50.5	49.5	53.2	51.1	53.7	65.3	50
μ_2 (ms)	436.3	357.9	358.8	361.6	354.0	**370.1**	365.9	419.0	1581.3	373
LWF	0.146	**0.151**	0.141	0.138	0.134	0.119	0.110	0.096	0.042	0.2

TABLE 4.7: Results using 64-echo sequence.

sequence more clinically viable. Therefore in this experiment 32 and 64 echoes were tested.

As in Section 4.1.1.2, ROIs of the same size were drawn in regions of 50ms and 373ms T2 and the regions were combined on a pixel by pixel basis. As a result of this combination of datasets 103 signal decays were produced. Using the two Gaussian model described in Section 4.1, estimates of mean, standard deviation and magnitude for both the long and short T2 compartments were produced.

In order to choose the optimal TE, three metrics were assessed. Reference T2 values for each phantom were calculated as the average T2 value from a mono-exponential fit of each pixel, due to the fact that the reference values from the manufacturer can change with time. The difference between the two reference T2 values of the phantoms and the T2 values produced by the LWI model (μ_1 and μ_2) show which echo spacing led the model to produce the most accurate representation of each individual phantom.

Then the LWF was calculated. This is an estimate of the percentage of the total signal originating from the long T2 component, representing the lumen in prostate images. The accuracy of the LWF estimate is important because earlier literature shows this value to be particularly good at detecting prostate cancer [136]. The difference between the calculated LWF and the actual ratio of long to short T2 components, which was set at 0.2, was calculated to see which echo spacing also optimised the estimation of this parameter.

Tables 4.6 and 4.7 show the results of this experiment. The aim of this experiment is to select an optimum echo spacing for both 32 and 64 echo acquisitions and compare the two in terms of their ability to accurately model the tissue.

It is interesting to note that, with the exception of the extreme TE values, the change in TE does not significantly affect any of these values. However, as an optimum must be chosen, the two best values for the 32-echo sequence appear to be 26ms or 30ms and the two best for 64-echoes appear to be 14ms or 18ms. Interestingly the overall ranges for these values are 832ms and 960ms for the 32-echo and 896ms and 1152ms for the 64-echo acquisition. This shows us that both sequences are finding their optimum T2 range to be in the region of $\sim$ 1000ms.

Comparing the 32 and 64 echo values, the 64-echo data seems to have a slight advantage when approximating the LWF but this difference is relatively small. None of the LWF values are particularly accurate in this experiment, which could be a product of the fact that the signal combination can't quite mimic the tissue accurately. However, the inaccuracy in the LWF values in this phantom experiment is not considered to be a significant problem as the more important aspect is LWI's ability to differentiate between tumour and normal tissue.

The 64 echo sequence is only possible by installing a Clinical Science Key or software patch onto the scanner so, as 32 echoes gives a good result in this experiment, it would be advantageous to use it instead of the 64 echo scan. This phantom data suggests that the step from 64 to 32 echoes is justified in that the best estimates of the LWF, μ_1 and μ_2 for each are barely different.

A choice of echo time must be made on the echo spacing that is to be used on the 32-echo acquisition. This echo spacing will be used for a study on 50-100 subjects, after which an analysis will be made of the subjects and any alterations to the sequence can be made based on a larger dataset. The 32-echo sequence has an optimum T2 range of 960ms in simulation and $\sim$ 1000ms in phantom. In order to trade off these values, and to take account of the range used in previous studies [136], a range of 1000ms was chosen to investigate further, giving an echo spacing of 31.25ms. This seemed a sensible choice given the limited data combined with the need to shorten the sequence acquisition.

4.3 Preliminary comparison of T2 and diffusion modelling

In fitting a multi-compartment model to multi-echo T2 data the assumption is made that one Gaussian peak represents stromal and epithelial cells whilst the other represents luminal space. In order to show that it can predict tumour grade it would be interesting to first investigate whether the model is good at quantifying microstructure.

The aim of this experiment was to investigate whether this multi-compartment T2 model could provide information about tissue microstructure. To do this it was compared to the VERDICT diffusion model[120], which separates tissue compartments and has shown to be broadly consistent with histological estimates of the same variables [136]. The justification behind this comparison is that the VERDICT model has been related back to histological features so is used as a proxy for histology in this experiment due to the lack of biopsy and prostatectomy data available in this preliminary investigative study.

4.3.1 Methods

T2 Sequence	MESE
Scanner	Philips Achieva 3T MRI
No. of echoes	32
Echo spacing (ms)	31.25
TR (ms)	8956
Acq. voxel size (mm)	2x2x4
FOV (mm)	180x180
Duration ($mm:ss$)	05:50
Diffusion Sequence	PGSE single shot EPI
Scanner	Philips Achieva 3T MRI
b values (s/mm^2)	0, 90, 500, 1500, 2000
TR/TE (ms)	2000-3707/ 50–80
Voxel size (mm)	1.3x1.3x5
FOV (mm)	220x220
Duration ($mm:ss$)	12:57

TABLE 4.8: Scan parameters of multi-echo T2 and diffusion sequences.

The parameters of the multi-echo T2 and VERDICT diffusion scans are included in Table 4.8. Five subjects were imaged for this study. Both sequences were optimised under different time constraints which has led to a difference in voxel sixe, but this is not considered to be an issue due to the fact that the statistical tests are conducted across

ROIs rather than on a voxel-by-voxel basis. A board certified radiologist contoured the five lesions observed in two patients, which were confirmed to be cancers following targeted biopsy. Regions of interest (ROIs) were also placed in areas of normal tissue in each of the five subjects. This gives five tumours and five areas of benign tissue in the study.

The VERDICT model was fit to the diffusion data as described previously[120]. For each voxel six T2-related parameters were estimated: overall scaling (M_0), relative fraction between the two compartments (α), mean and standard deviation of the short T2 peak (μ_1 and σ_1) and mean and standard deviation of the long T2 peak (μ_2 and σ_2). The Luminal Water Fraction (LWF), as discussed in previous work[133], was also processed for each voxel after fitting. It is calculated as the area under the long T2 component divided by the area under both components and, much like the Myelin Water Fraction in the brain, has the potential to be a clinically relevant quantity if the link is proven histologically.

Maps of the VERDICT parameters (cell radius (R), cellularity, extra-cellular extra-vascular fraction (fEES), intra-cellular fraction (fIC) and vascular fraction (fVASC)) and the parameters of the multi-compartment T2 model were produced as described in Panagiotaki et al. [121]. Correlations between mean parameter values for both tumour and normal ROIs were computed for both models, alongside a p-value to quantify the significance of the correlation. Using a 95% confidence, a p-value less than 0.01 was taken to be significant (Bonferroni corrected for five comparisons). For this, a simple Pearson's correlation was carried out.

4.3.2 Results

Table 4.9 shows the correlations between each of the model parameters and their p-values. The significant correlations are highlighted in bold. There is a strong positive correlation between fEES and LWF, between R and LWF and between fEES and μ_2.

Figure 4.4 contains maps of the parameter LWF for two tumour regions, showing that tumour (outlined in red) has a lower value of LWF than the highly luminal peripheral

	R		Cellularity		fEES		fIC		fVASC	
	ρ	p	ρ	p	ρ	p	ρ	p	ρ	p
M_0	-0.052	0.887	-0.506	0.136	0.330	0.352	0.061	0.867	-0.615	0.058
α	0.590	0.073	-0.060	0.870	-0.178	0.623	0.407	0.243	-0.526	0.118
μ_1	-0.326	0.358	-0.140	0.699	0.546	0.102	-0.439	0.204	0.037	0.920
μ_2	-0.605	0.064	-0.171	0.637	0.737	0.015	-0.579	0.080	0.009	0.981
σ_1	-0.263	0.462	-0.192	0.596	0.391	0.264	-0.305	0.391	0.009	0.980
σ_2	-0.177	0.624	0.117	0.748	0.123	0.734	-0.027	0.940	-0.140	0.700
LWF	-0.749	0.013	-0.289	0.417	0.720	0.019	-0.606	0.063	0.092	0.800

TABLE 4.9: Correlation and p-values comparing multi-compartment T2 and VERDICT models.

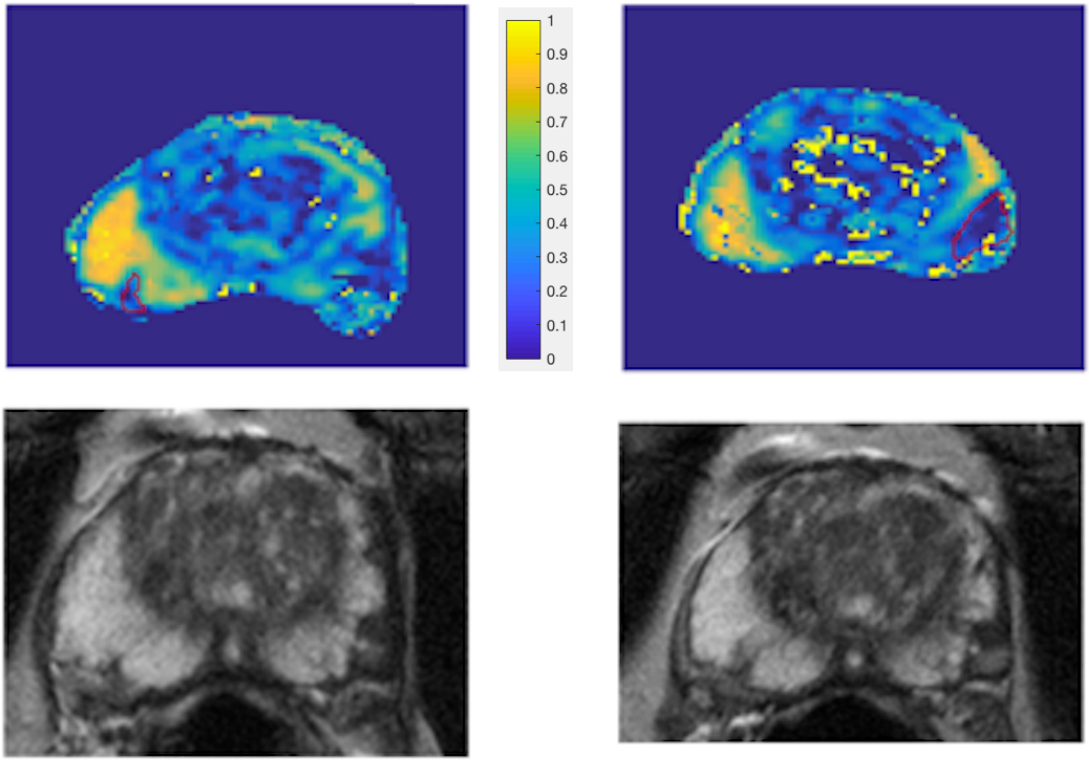


FIGURE 4.4: Maps of LWF and axial T2 for two example tumours. The tumours are outlined in red in the first row.

zone. This is likely to correspond with a decreased luminal tissue component, one of the hypothesised hallmarks of prostate cancer [32].

4.3.3 Discussion and Conclusion

These results suggest that the LWF value and the VERDICT diffusion model are showing some of the same phenomena in the prostate tissue. An interpretation of the results

is that in PZ tumours there is an increase in cell radius R, with a corresponding reduction in extra-cellular extra-vascular space fEES and a decreased luminal water content. Importantly, there is agreement between the two models about the changes happening within the tissue. Though it has only been carried out on a small number of patients it has shown that there this method shows promise as a way of non-invasively quantifying the relative volumes of the tissue compartments. Building on that, if as Chatterjee et al. [32] say the correlation between compartment volumes and Gleason grade are directly related, this result also suggests that the LWF could be able to distinguish between tumour grades.

In diagnostic imaging both T2W and diffusion images are acquired to assess different tissue properties, suggesting they carry complementary information. The differences in both the LWI and VERDICT models further suggest that the combination of the two could provide more information than one alone. The VERDICT model separates epithelia, contained within the fIC, from the lumen and stroma, represented by the fEES. By using the VERDICT model to separate epithelia from stroma and lumen and LWI to separate lumen from stroma and epithelia it seems that a model specifically quantifying each of these compartments individually could be attained.

The limitations of this experiment were that it was only carried out on a small number of subjects and that this correlation, though going some way to showing that LWI does model the microstructure well, is no substitute for actually comparing the metric against histology.

This experiment was published at the 25th meeting of the International Society of Magnetic Resonance in Medicine in 2017. Subsequent to this experiment being carried out a paper was published correlating LWI to histology [136], showing that the LWF strongly correlates with the fractional amount of luminal space in the prostate.

4.4 Differentiating tumour from normal tissue

Having discovered that LWI does seem to agree with diffusion microstructural models, next is a preliminary investigation into whether it can differentiate between tumour and

normal tissue. The ultimate aim is to non-invasively measure the grade of a particular lesion but the first step is detection.

An experiment was devised to test whether the differences of the parameters of LWI between tumour and normal tissue were significant. The five subjects from the initial pilot study were imaged using the ME-T2 scan parameters in Table 4.8. For each subject an area of both tumour and normal tissue were contoured by a radiologist based on information from biopsy taken after the acquisition, which is subject to the usual errors in the localisation of targeted biopsies. From this the mean value of each of the parameters of the model was calculated for each ROI.

To test the ability of the model to differentiate tumour from normal tissue a Kruskal-Wallis test was carried out. The Kruskal-Wallis test is a non-parametric test analysing whether a number of samples come from the same distribution. The reason for using a non-parametric test is that it does not assume a normal distribution of parameter values. Though the model used here does assume two Gaussians within the T2 distribution, the parameters that are produced by that model cannot themselves be assumed to be normally distributed.

	Healthy		Tumour	
	Mean	Range	Mean	Range
M_0	1.973	1.380-2.497	1.667	1.207-2.145
α	0.619	0.331-0.912	0.856	0.502-0.983
μ_1	0.074	0.058-0.090	0.072	0.060-0.082
μ_2	0.511	0.338-1.042	0.705	0.343-1.837
σ_1	0.003	0.001-0.007	0.003	0.001-0.006
σ_2	0.005	0.002-0.007	0.005	0.002-0.007
A_1(x10^8)	0.193	0.000-3.087	0.000	0.000-0.000
A_2	0.754	0.123-1.487	0.252	0.026-0.916
LWF	0.390	0.090-0.640	0.145	0.016-0.517
ADC (x10^{-4})	7.657	5.972-9.256	6.240	3.941-9.241

TABLE 4.10: Mean values and range for model parameters.

The LWF, μ_1, and μ_2 parameters all showed significant differences between normal tissue and tumour (p = 0.009, 0.047, 0.028 respectively). The mean values and ranges of the parameters are shown in Figure 4.10. As a proof of concept study on a small number of patients, this result shows that the LWI parameters, and in particular the LWF, can indeed help to detect prostate cancer. The fact that a ME-T2 sequence can be used for

this purpose is noteworthy. Furthermore, it goes some way to showing that the simplified LWI model, constrained to be two Gaussians, is as effective in detecting tumour as the more complicated NNLS method used in previous works. Other works with 64 echoes and a more complex NNLS fitting [134, 136] have shown to be effective but this work suggests that LWI derived parameters are also useful in detecting tumour when a 32 echo sequence is used along with a less complicated fitting method.

4.5 Summary

The aim of this chapter was to investigate which combination of model and acquisition best described the tissue microstructure of the prostate. The ability of this combination to both describe the microstructure and detect tumours was then assessed in a small number of subjects.

The first conclusion of this chapter was that the NNLS fit used in past works could be replaced with a simpler two-Gaussian model. Then the optimum ME-T2 acquisition was found to be a 32-echo sequence with a range of 1000ms and an echo spacing of 31.25ms.

Preliminary datasets then suggested that the LWI model using this combination of model and sequence correlated with tissue microstructure and could differentiate normal tissue from tumour.

Chapter 5

Simplified Luminal Water Imaging for the Detection of Prostate Cancer from Multi-Echo T2 MR Images

The work in Chapter 4 provided an acquisition and processing scheme that can differentiate between PCa and normal tissue. This chapter seeks to combine the optimised ME-T2 sequences with the new fitting methods to evaluate whether they can detect and grade PCa in a larger number of patients. This work has been accepted for publication in the Journal of Mangetic Resonance Imaging.

5.1 Abstract

Luminal Water Imaging (LWI) suffers less from imaging artifacts than the diffusion-weighted imaging used in multi-parametric MRI of the prostate. LWI obtains multi-compartment tissue information from a multi-echo T2 dataset. The purpose of this work is to compare a simplified LWI technique with Apparent Diffusion Coefficient (ADC) in classifying lesions based on groupings of PI-RADS v2 scores. Secondary aims

were to investigate whether LWI differentiates between histologically confirmed tumour and normal tissue as effectively as ADC, and whether LWI is correlated to the multi-compartment parameters of the Vascular, Extracellular, and Restricted Diffusion for Cytometry in Tumours (VERDICT) diffusion model.

Sixty-five male patients aged 49-79 were scanned as a subset of a larger prospective study. A 32-echo T2 and a 6 b-value diffusion sequence (0,90,500,1500,2000,3000 s/mm2) were used at 3T. Regions of interest were placed by a board-certified radiologist in areas of lesion and benign tissue and given PI-RADS v2 scores. Receiver Operating Characteristic and logistic regression analyses were performed.

LWI classifies tissue as PI-RADS 1,2 or PI-RADS 3,4,5 with an Area Under Curve (AUC) value of 0.779, compared to 0.764 for ADC. LWI differentiated histologically confirmed malignant from non-malignant tissue with AUC, sensitivity and specificity values of 0.81, 75% and 87%, compared to 0.75, 83% and 67% for ADC. None of the AUC values for LWF were statistically significantly different to those of ADC but this still shows that LWI can produce the same results as ADC with none of the downsides that affect diffusion-weighted imaging.

The microstructural basis of the LWI technique is further suggested by the correspondence with the VERDICT diffusion-based microstructural imaging technique, with the α, $A1$, $A2$, and LWF showing significant correlations. LWI alone can predict PI-RADS v2 score groupings and detect histologically confirmed tumours with an ability similar to ADC alone without the limitations of diffusion-weighted MRI. This is important given that ADC has an advantage in these tests as it already informs PI-RADS v2 scoring. LWI also provides multi-compartment information that has an explicit biophysical interpretation, unlike ADC.

5.2 Introduction

The suspicion of PCa is typically assessed using multi-parametric Magnetic Resonance Imaging (mp-MRI) images, including T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI) and dynamic contrast-enhanced (DCE) measurements [151], which are

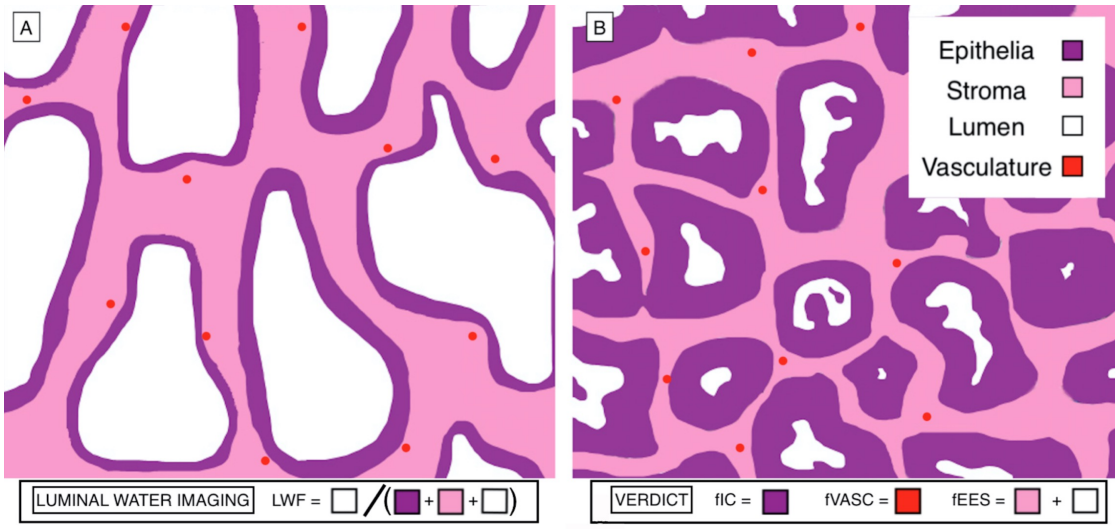


FIGURE 5.1: Diagram of prostate histology in A) Benign tissue B) Malignant adenocarcinoma. The four separate micro-environments present in the prostate are shown in the key. Using LWI, the short-T2 component is made up of the stroma and epithelia and the long-T2 component is made up of the lumen (the T2 of the vasculature is not considered to have a significant effect on the LWI model). Using the VERDICT model, the intra-cellular compartment is made up of epithelia, the extracellular-extravascular compartment is made up of stroma and lumen and the vascular compartment is made up of the vasculature.

then scored using an ordinal scale, most commonly the Prostate Imaging Reporting and Data System (PI-RADS) v2 scheme [119, 167]. mp-MRI provides sensitivity and specificity values of between 87-93% and 41-47% respectively in the detection of clinically significant prostate cancer [6].

Current limitations of mp-MRI are that its specificity when detecting tumours is low [6] and there is only a moderate inter-reader agreement across all lesions (55-65%) [140]. PI-RADS v2.0 on a 1.5T scanner also classifies 35.6-44.7% of lesions as indeterminate [6], meaning that the method cannot distinguish whether a large number of cases are clinically significant prostate cancer or not, though in practice the classification may vary depending on the radiologist, the patient population and the image quality.

Furthermore, Echo Planar Imaging (EPI) based diffusion scans often suffer from distortion, signal pile-up, or stretching artefacts [106], leading to less accurate results and in some circumstances unusable images. Different studies have tried to solve some of these issues, particularly with regards to reducing the proportion of lesions classified as indeterminate [27, 127], but none has conclusively solved them.

The tissue within the prostate has three major components, luminal space, epithelial cells and stromal cells, as shown in Figure 5.1. The luminal space stores the fluid produced by the surrounding epithelial cells and the stromal cells form a matrix surrounding these compartments, giving the prostate structure and forcing the fluid out of the prostate during ejaculation.

The Luminal Water Imaging (LWI) technique models two compartments, one the luminal space with a distribution of long T2 values and the other both the stromal and epithelial cells with a distribution of short T2 values. The VERDICT technique models three compartments, the intracellular compartment, which represents the restricted diffusion of the epithelial cells, the extracellular-extravascular compartment, which represents hindered diffusion within the luminal space and stroma, and the vascular compartment. The diffusivity values for the intracellular and extracellular-extravascular compartments are assumed to be the same and of a lower value than the pseudo-diffusivity of the vascular compartment.

The prostate is made up of a number of separate tissue zones. Each is made up of different proportions of gland, stroma and epithelia. The central zone consists of large, irregularly shaped glands, cuboidal epithelial cells and compact stromal tissue. The peripheral zone contains small, regularly arranged glands lined by columnar epithelial cells and surrounded by loosely interwoven stroma. The transition zone, composed of two lobules either side of the urethra, closely resembles the peripheral zone tissue.

Storas et al. 2008 [152] found that a multi-echo T2 (ME-T2) sequence is consistently better at probing tissue microstructure than a single echo T2 sequence, showing that in the prostate a mono-exponential fit is only adequate for describing the underlying tissue in 10% of subjects. Sabouri et al. [136] implemented a method to investigate quantitative T2 imaging in the prostate, proposing the Luminal Water Imaging (LWI) technique, which produces an estimate of the fractional volume of luminal water in each voxel of the prostate, the Luminal Water Fraction (LWF). Sabouri et al. have shown that there is a good correlation between LWF and histologically measured luminal fractional volume [136] and determined that LWI shows promise in being able to detect PCa and predict Gleason score [134].

However, the fitting method used has a high number of degrees of freedom, potentially making it vulnerable to noise and local minima. They used an echo train with 64 echoes which is not available on all scanners. In addition, compared with using fewer echoes, a 64 echo train may be more vulnerable to the cumulative effects of imperfect refocusing pulses and have a higher specific absorption rate.

This study proposes a simplified Luminal Water Imaging model fitting combined with a more clinically feasible 32-echo pulse sequence that has a lower specific absorption rate than 64-echo acquisitions and a potentially reduced cumulative error due to imperfect refocusing. A fitting on simulated data seeks to show that the simplified model and shorter sequence give similar values to the original LWI model.

By investigating how LWI compares to ADC, this work seeks to prove that equivalent information can be attained through multi-compartment modelling using multi-echo T2 data. In order to achieve this, the main aim of this work is to investigate whether LWI can distinguish between different PI-RADS v2 scores as well as ADC. Secondary aims are to investigate whether LWI can differentiate between histologically confirmed tumour and normal tissue [3] as well as ADC and whether LWI is correlated to the parameters of the Vascular, Extracellular, and Restricted Diffusion for Cytometry in Tumours (VERDICT) diffusion model [120].

5.3 Materials and Methods

This study introduces a new multi-echo T2 acquisition coupled with a new fitting method for LWI and compares the ability of this LWI to ADC in predicting clinically relevant groupings of PI-RADS v2 scores. There is also a comparison with the VERDICT microstructural diffusion method to confirm that LWI correlates with microstructural information from other techniques.

5.3.1 Simulation

A simulation was used to compare alternative methods for fitting the Luminal Water Fraction. Signals from a tissue with a distribution of T2 values were simulated. Datasets typical of both acquisition schemes were used (Sabouri: TE=25ms, NE=64, this study: TE=31.25ms, NE=32) and two separate datasets were created on different assumptions about the underlying T2 distribution, one using two Gaussian peaks and the other using two delta peaks. The values used for the shorter T2 compartment were 20, 50 and 80ms and for the longer T2 compartment were 300, 550 and 800ms. Using a value of Signal to Noise Ratio (SNR) of 100 and for ground truth LWF values of 0, 0.1 and 0.2 (a range typical of both tumour and normal tissue [136]), the LWF was calculated from fits to each set of conditions.

5.3.2 Patient Selection

In-vivo data were acquired on a sub-cohort of 65 patients from a larger prospective study [78]. The study received UK Research Ethics Committee approval on 23rd December 2015. Patients were recruited between September 2016 and October 2017 and provided written informed consent following a minimum 24-h period of consideration. The patient inclusion criteria were (1) men sent for prostate mp-MRI following previous biopsy more than 6 months previously and (2) biopsy naive men presenting a clinical suspicion of prostate cancer. Patient exclusion criteria included (1) men unable to have an MRI scan, or in whom artefact would reduce the quality of the MRI, (2) men unable to give informed consent, (3) previous treatment (prostatectomy, radiotherapy, brachytherapy) of prostate cancer, (4) on-going hormonal treatment for prostate cancer, and (5) previous biopsy within 6 months of scheduled mp-MRI. Five subjects were excluded throughout the course of this study, four due to MR contraindications and one due to a technical fault. In this paper an assessment is made on a broad range of subjects, including men for whom it was decided a biopsy was not necessary.

5.3.3 MRI Acquisitions

Subjects were scanned on a 3.0T scanner (Philips Achieva; Philips Medical Systems, Best, The Netherlands) using a 32-channel cardiac coil. A multi-echo spin-echo sequence with an echo spacing of 31.25ms and TR of 8956ms was used. The other parameters were: Number of echoes = 32; Field of View = 180x180 mm; Acquired voxel size = 2x2x4mm; Scan duration = 5min 50s. DWI was acquired for VERDICT fitting with Single Diffusion Encoding (SDE) single shot EPI sequences over 6 b-values (0,90,500,1500,2000,3000 s/mm2). TR/TE = 2000-3707/50-80 ms; Field of View = 220x220 mm; Voxel size = 1.3x1.3x5mm; Scan duration = 12:57 [120]. A standard mp-MRI protocol was also conducted on these patients as detailed in previous work [78]. Voxel sizes of the ME-T2 and diffusion acquisitions were different due to the time constraints that each was designed under.

5.3.4 Regions of Interest and histologic examination

A radiology research fellow with five years of experience in prostate mp-MRI reporting (FG) contoured ninety-seven suspected areas of either malignant or benign tissue. The lesions had previously been located in the mp-MRI images and the primary lesions were then located and contoured in a single slice of the 93.75ms echo of the ME-T2 image with no knowledge of the LWI maps. In the case of malignant tissue the entire lesion was outlined whereas in healthy tissue the ROI from the lesion was copied into a region of healthy tissue in the same prostate zone. The 93.75ms echo was chosen for its similar echo time to a traditional axial T2 weighted prostate image ($\sim$100ms). Ninety-eight regions of interest (ROIs) were also contoured on the corresponding slice of the ADC maps. These ROIs were drawn to directly correspond to those contoured on the ME-T2 images. The numbers of ROIs were slightly different in the ADC and ME-T2 images due to two ME-T2 images having not been correctly acquired and one ADC image having a large artifact in the region of interest.

Each ROI was assigned a PI-RADS v2 score based upon the standard mp-MRI images, acquired in addition to the VERDICT and ME-T2 images. PI-RADS v2 is a method

of scoring tissue on a scale of 1-5, with 1 meaning clinically significant cancer is highly unlikely to be present and 5 meaning clinically significant cancer is highly likely to be present. In the peripheral zone (PZ) of the prostate, where the majority of tumours arise, this scoring is primarily informed by diffusion images, with axial T2 and DCE images used when the diffusion image is indeterminate. Histological grading using a targeted transperineal template biopsy was available on a subset of 31 of the ROIs. The Gleason grading system was used to evaluate the biopsy tissue samples [3].

Our primary analysis uses ROIs drawn without knowledge of the LWF maps to avoid bias. If LWF maps are used directly in the future, an indication of reproducibility between readers will be informative. In a sub-study on 16 patient data sets, two separate readers each drew 16 ROIs on the LWF maps themselves and the median parameter values were evaluated using a correlation and Bland-Altman analysis.

5.3.5 Data processing

Sabouri et al [136] used a regularised non-negative least squares (NNLS) algorithm to fit a multi-exponential model. The NNLS algorithm fits a large number of exponentials (>100) to the signal decay curve, including a regularising term in the error minimisation to compensate for problems associated with having a large number of unknown variables compared to the number of echoes. A large number of exponentials also seems inappropriate because in previous works [134, 136, 152], only two compartments are usually identified in the prostate gland.

Hence, for this work we have constrained the model to two compartments, each with a Gaussian probability distribution of T2 values. The choice of Gaussians is a mathematically simple choice and, based upon preliminary analysis, the difference between Gaussian, log-Gaussian and gamma distributions made little difference to the overall fit of the signal decay. Hence a two-Gaussian model was fitted to the individual T2 signal decay curves using a least squares regression. These two Gaussian distributions model the tissue as a combination of a luminal compartment with long T2 values and a compartment composed of stroma and epithelia with shorter T2 values. The probability density p over T2 value T2 in a particular pixel is given by

$$p(T2) = \frac{\alpha}{\sigma_1\sqrt{2\pi}} exp\Big(- \frac{(T2-\mu_1)^2}{2\sigma_1^2}\Big) + \frac{1-\alpha}{\sigma_2\sqrt{2\pi}} exp\Big(- \frac{(T2-\mu_2)^2}{2\sigma_2^2}\Big), \tag{5.1}$$

with the signal intensity S at each echo time TE subsequently calculated as

$$S = M_0 \int_0^\infty p(T2) exp\Big(- \frac{TE}{T2}\Big) dT2. \tag{5.2}$$

This model fitting minimises the mean squared error between the actual signal and the modelled signal, using the Levenberg-Marquardt algorithm to optimise over six parameters, the absolute signal magnitude (M_0), the magnitude ratio between the two peaks (α), the means of the two components (μ_1 and μ_2), and the variances of the two components (σ_1 and σ_2).

The values of μ_1 and μ_2 were constrained to be 0-200ms and 200-3000ms respectively, none of the other parameters were constrained. The starting values for the constrained model for M_0, μ_1 and μ_2 were calculated once for each subject by taking an average intensity for each echo over all pixels in the prostate and carrying out a bi-exponential fit on the averaged signal decay across the ROI. This study used the short and long T2 values of the bi-exponential fit as initial estimates for μ_1 and μ_2 and the mean of the magnitudes of each exponential as the initial estimate for M_0. Initial values of σ_1 and σ_2 were set at $5\text{x}10^{-4}$, which was in the range of the standard deviation of the NNLS peaks in previous LWI fittings. This reduction in the number of parameters compared to the unconstrained model should provide a more reliable fit and operate using fewer data points, permitting echo trains with fewer echo signals.

The areas under the individual peaks, A_1 for the shorter T2 peak and A_2 for the longer T2 peak, were calculated by integrating the respective Gaussians using their magnitude, mean and variance. The Luminal Water Fraction (LWF) was then calculated as the fraction of the total area under the distribution curve attributed to the peak with the longer T2:

$$LWF = \frac{A_2}{A_1 + A_2} \tag{5.3}$$

For each of these parameters a map was created across the entire prostate. Then the ROIs produced earlier on the 93.75ms echo of the ME-T2 image were superimposed onto these parameter maps and the median values of these parameters were calculated for each ROI. All data were processed using Matlab (*MATLAB and Statistics Toolbox Release 2017a, The MathWorks, Inc., Natick, Massachusetts, USA*).

The VERDICT model, when applied to the prostate, is a three-compartment diffusion-based microstructural model that characterizes water diffusion into vascular, intracellular (IC) and extracellular-extravascular space (EES) compartments. The IC compartment has volume fraction (fIC), diffusivity (dIC) and cell radius (R) as parameters. The EES compartment has volume fraction (fEES) and diffusivity (dEES) as parameters. The vascular model has volume fraction (fVASC) and pseudo-diffusivity (P) as parameters.

5.3.6 Statistics

In order to create a 95% confidence interval for the simulation data, bias corrected and accelerated percentile bootstrapping was used on 1000 bootstrap samples. Differences were characterised between the median parameter values of ROIs with different PI-RADS v2 groupings of scores and determined using a logistic regression model combined with 5-fold cross validation. Three comparisons were made in this way. The comparison between the scores PI-RADS 1,2 vs PI-RADS 3,4,5 aims to divide those lesions needing further action from those that do not. The other two comparisons, PI-RADS 1,2 vs PI-RADS 3, and PI-RADS 3 vs PI-RADS 4,5, both aim to investigate the model's ability to discern between the three main categories of negative (1,2), indeterminate (3) and positive (4,5) disease. A p-value of $p<0.008$ was taken to be significant (Bonferroni corrected p-value for 95% significance).

The mean values for sensitivity, specificity and area-under-curve (AUC) values across the five-folds were also computed using a Receiver Operating Characteristic (ROC) analysis. Sensitivity and specificity values were calculated from the ROC analysis using an operating point with the shortest distance to the point of perfect discrimination. A

logistic regression was performed on the median values of those ROIs with a corresponding histological grading in order to distinguish malignant (Gleason 3+3 and above) from non-malignant tissue.

To detect significant statistical differences between the values of AUC for LWF and ADC when predicting PI-RADS v2 categories, a Kruskal-Wallis non-parametric statistical test was performed on the AUC values of each of the five folds of the cross validation. To detect significant statistical differences between the values of AUC for LWF and ADC when predicting Gleason categories, a Kruskal-Wallis non-parametric statistical test was performed on 1000 bootstrapped examples. For both the PI-RADS v2 and Gleason Score Kruskal-Wallis tests using a Bonferroni correction, a p-value less than 0.013 would suggest with 95% confidence that the null hypothesis (that the AUC values of ADC and LWF come from the same distribution) be rejected.

In order to assess the relationship between LWI and the VERDICT diffusion model for prostate [120] a Pearson's correlation coefficient was calculated between the VERDICT intra-cellular volume fraction (fIC) and each of the parameters of LWI separately. The fIC parameter was chosen due to the fact that it has previously shown significant difference between PCa and normal tissue [3] and that it represents the cellular compartment within the tissue and hence might be expected to negatively correlate with the LWF.

A standard Bland-Altman analysis was carried out on the median values within the subset of 16 ROIs to analyse the reproducibility of this method.

5.4 Results

The age range of the 65 patients was 49-79 years with a mean of 65. Of the T2 ROIs with PI-RADS v2 scoring, there were 31 PI-RADS 1, 32 PI-RADS 2, 18 PI-RADS 3, 5 PI-RADS 4 and 11 PI-RADS 5. Similarly for the diffusion ROIs there were 30 PI-RADS 1, 32 PI-RADS 2, 18 PI-RADS 3, 7 PI-RADS 4 and 11 PI-RADS 5. Of the histologically examined ROIs 16 were found to be benign, 3 were Gleason 3+3, 6 were 3+4, 3 were 4+3, 2 were 4+4, 1 was 4+5.

The simulation results in Table 5.1 show that the proposed two-Gaussian method with 32 echoes has a similar accuracy to the original LWI at determining the LWF over a range of ground truth tissue models.

Table 5.2 shows the results of the statistical tests for in-vivo PI-RADS v2 score groupings. Table 5.3 presents the comparison of the proposed method with histologically confirmed malignant lesions and shows values of 0.81 and 0.75 (AUC), 75% and 83% (sensitivity) and 87% and 67% (specificity) respectively for LWF and ADC.

Table 5.4 shows the results of Kruskal-Wallis tests designed to test whether there is a significant difference between the AUC values produced by ADC and those produced by LWF. These tests were performed on each of the four score groupings investigated. In none of the comparisons made can the results prove with 95% confidence that the ADC and LWF are producing significantly different AUC values.

Table 5.5 shows the correlations between the individual parameters of the LWI model and three parameters from the VERDICT diffusion model. With respect to fIC, significant positive correlations greater than 0.5 were seen for α and A1 and significant negative correlations less than -0.5 were seen for A2 and LWF. With respect to fEES, significant positive correlations greater than 0.5 were seen for A2 and LWF and significant negative correlations less than -0.5 was seen for α. None of the LWI parameters show a correlation greater than 0.5 or less than -0.5 with fVASC.

Figure 5.2 shows the LWF map for one patient alongside the axial-T2 and ADC maps from the same subject. Note the higher LWF in the peripheral zone (PZ), consistent with histological findings of large regular glandular lumen and loosely woven stroma in the PZ. Figure 5.3 shows graphs for the T2 distributions from single example voxels in healthy and cancerous tissue respectively. Clear changes in the distributions are visible.

The Bland-Altman analysis of the median LWF values of the subset of 16 ROIs is shown in Figure 5.4. It produced an r^2 value of 0.928, a bias of 0.013 and limits of agreement of -0.151 and 0.124, with a mean value of 0.179.

		Ground Truth LWF		
Pulse Sequence	Model Fitting	0	0.1	0.2
32-echo & 31.25ms echo spacing	Two-Gaussian	0.0010	0.1023	0.1943
32-echo & 31.25ms echo spacing	NNLS	0.0002	0.0901	0.1826
64-echo & 25ms echo spacing	Two-Gaussian	0.0028	0.1007	0.2007
64-echo & 25ms echo spacing	NNLS	0.0003	0.0910	0.1843

TABLE 5.1: Mean estimated LWF values for both the constrained and unconstrained models using different ground truth LWF values in simulation. These are the mean values using both the delta and Gaussian ground truth models over a range of μ_1 and μ_2 values.

Scores	1,2v3				3v4,5				1,2v3,4,5			
Test	p	AUC	Sen.	Spe.	p	AUC	Sen.	Spe.	p	AUC	Sen.	Spe.
LWF	0.005	0.786	63.3	84.0	0.013	0.867	80.0	76.7	0.000	0.881	80.6	78.7
ADC	0.395	0.655	48.7	80.9	0.081	0.747	73.3	79.3	0.035	0.691	59.4	77.4

TABLE 5.2: Three separate analyses using LWF and ADC to predict PI-RADS v2 scores. The p-value comes from the logistic regression model and the other statistics are from an ROC analysis.

Variable	AUC	Sensitivity	Specificity
LWF	0.81	75%	87%
ADC	0.75	83%	67%

TABLE 5.3: ROC analysis of LWF in detecting malignant lesions (Gleason 3+3 and above).

Score Groupings	p-value
PI-RADS 1,2 v 3	0.0758
PI-RADS 3 v 4,5	0.1246
PI-RADS 1,2 v 3,4,5	0.0758
Gleason 3+3 and above	0.7771

TABLE 5.4: p-values of Kruskal-Wallis tests between ADC and LWF for the ROC analyses performed on each of the four score groupings tested. A p-value of 0.013 means that the null hypothesis, that the ADC and LWF predictions have the same AUC values, can be rejected with 95% confidence.

		M_0	α	μ_1	μ_2	σ_1	σ_2	A_1	A_2	LWF
f_{IC}	ρ	0.394	0.620	-0.131	0.252	0.025	0.272	0.516	-0.602	-0.618
	p	0.000	0.000	0.246	0.024	0.824	0.015	0.000	0.000	0.000
f_{EES}	ρ	-0.232	-0.541	0.069	-0.274	0.058	-0.261	-0.364	0.583	0.564
	p	0.039	0.000	0.543	0.014	0.612	0.019	0.001	0.000	0.000
f_{VASC}	ρ	-0.170	-0.081	0.059	0.065	-0.097	-0.004	-0.162	0.002	0.038
	p	0.132	0.474	0.603	0.567	0.395	0.971	0.152	0.983	0.736

TABLE 5.5: Correlation between LWI parameters and the intra-cellular fraction (fIC), extracellular-extravascular fraction (fEES) and vascular fraction (fVASC) parameters of the VERDICT diffusion model.

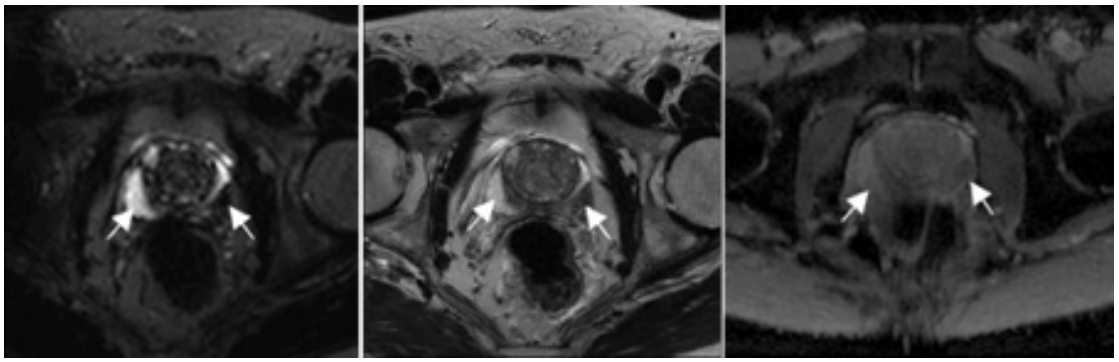


FIGURE 5.2: LWF map, axial T2 image and ADC map for one patient. The region of healthy tissue is highlighted by the arrow on the left of each image, the tumour by the arrow on the right of each image. This figure shows distortions in the PZ in the ADC image, highlighting a disadvantage of DWI over ME-T2 modelling.

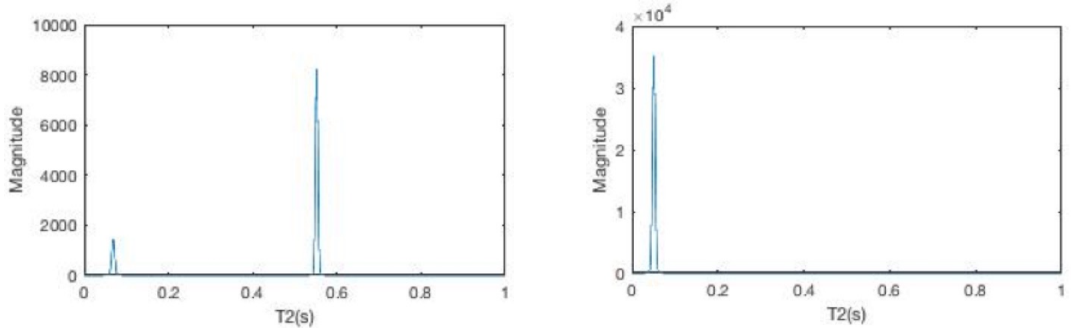
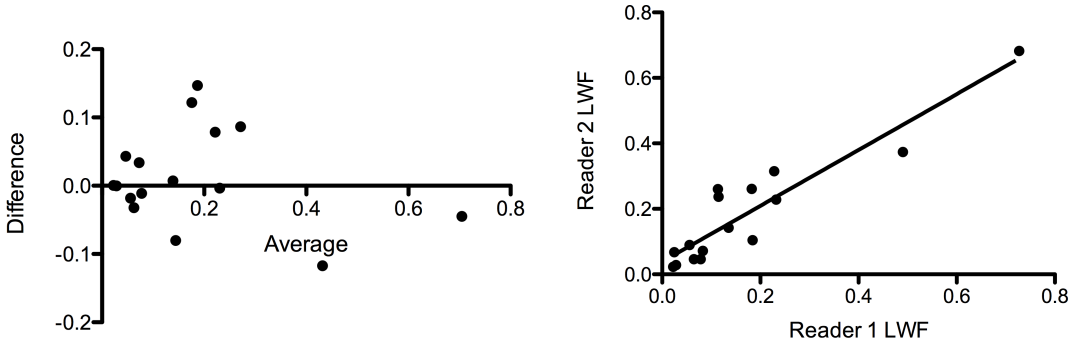


FIGURE 5.3: T2 distributions for an example pixel in a) normal tissue and b) tumour. These pixels are taken from the subject in Figure 5.2 in the regions highlighted by the arrows.



(a) Bland-Altman plot.

(b) Correlation plot.

FIGURE 5.4: Bland-Altman and correlation plots for median LWF values of ROIs from two separate readers.

5.5 Discussion

The main aim of this work was to compare a simplified LWI with ADC for the prediction of PI-RADS v2 scores. The LWF predictions showed similar AUC values to ADC,

suggesting that LWF is able to discriminate between clinically relevant PI-RADS v2 groupings of scores as well as ADC. Given that the PI-RADS v2 scoring scheme uses ADC as a major component and might thus be expected to favour ADC, it is interesting to find that LWI alone can predict PI-RADS v2 score as well as ADC alone. The ability of LWF to differentiate tumour from normal tissue using histologically examined ROIs further reinforces the idea that multi-echo T2 modelling shows promise as a method for detecting and grading PCa.

Our results suggest that LWI may be a useful tool in prostate cancer detection. If using the LWF measure in the future, ROIs may be drawn directly on LWF maps, in which case the repeatability between readers is an important consideration. The correlation and Bland-Altman analysis carried out on a subset of our patients provides an indication of the expected variability. A more extensive quantification of the value of LWI in prostate cancer detection requires larger and multi-centre studies and the results presented here can inform those studies.

The correlations between LWI and the VERDICT model are intended to give an insight into how LWI relates to tissue microstructure. An increase in LWF represents an increase in the fractional volume of luminal space within the prostate. As PCa typically invades the luminal space and reduces the luminal fractional volume, a reduced LWF value is expected within a tumour. The parameters with significant correlations greater than 0.5 in magnitude suggest that as the volume fraction of the luminal space decreases the volume fraction of the intra-cellular compartment increases and the volume fraction of the extracellular-extravascular compartment decreases [120]. These correlations with the VERDICT model parameters suggest that LWI is sensitive to the underlying tissue microstructure. Note that these correlations are different to the initial ones in Section 4.3. This is likely down to the large increase in the size of the dataset in this chapter and neither contradicts the conclusions of the other, with both pointing towards the relationship between th LWF parameter and the fractional volume of lumen in the prostate.

MRI is utilised in the prostate for the detection and staging of tumours. LWI allows for the collection of microstructural information without the distortion artifacts seen in

diffusion imaging. In the future, it might be that either less diffusion data needs to be acquired, or that complementary information from LWI improves the efficacy of mp-MRI given the different tissue compartments that they are designed to probe. Further sequence optimisation to increase coverage of the prostate and reducing the thickness of the slices would be beneficial prior to testing on a much larger number of subjects and across multiple centres. A future prospective assessment should also include a multiple reader study to quantify the variability introduced by radiological placement of ROIs.

Although a PI-RADS score can be allocated to any lesion, a limitation of this study is that not all patients subsequently received biopsy, limiting the ability to perform full histological correlations. Another limitation is that biopsy is prone to sampling error which can lead to the undergrading of tumours [69]. Sampling error could be reduced by using whole-mount histopathology but this would lead to a bias in the Gleason grades of the lesions studied due to the fact that radical prostatectomies are only carried out on subjects with more significant lesions. Furthermore, the subjects in this study were a subset of a larger prospective study [78], meaning that the analysis done in this paper was retrospective. A larger number of subjects could prove more conclusively the hypothesis that LWI performs at least as well as ADC in discerning the PI-RADS v2 score and Gleason score. This would allow for a reduced number of protocols for grading prostate cancer, reducing overall scan times. The AUC for the detection of PCa is also lower in this study than previous results [134], possibly due to previous studies using patients scheduled for retropubic prostatectomy, weighting the lesions towards more advanced tumours.

In conclusion, this work suggests that LWI is sensitive to the tissue microstructure and can be as effective as ADC in the classification of lesions using the PI-RADS V2 scores whilst providing images with minimal distortions.

Chapter 6

Optimisation for Clinical Use

The initial studies done into the optimisation of the acquisition of the ME-T2 sequence for modelling LWI were carried out on simulated data, gel phantoms and a small number of patients. This was intended to give a good approximation of the best values in order to begin the first studies on a large number of patients. However, now these data have been acquired they can be utilised to further perfect the acquisition.

One question that arises from the previous studies is how few echoes can be acquired without reducing the effectiveness of the LWI technique. Studies by other authors into LWI have all used 64-echo acquisitions so the original aim of the studies in this work was to reduce the echo sequence from 64 to 32 echoes, meaning that the sequence could then be used clinically on most scanners. Having acquired 32-echo data on 65 patients a study into the minimum number of echoes that can be used can now be carried out. In conjunction with another researcher, a study was made into the effects that the reduction in echo number has on the ability of the LWF to distinguish tumour from normal tissue.

Another area of the MESE acquisition that is of interest, as addressed in the brain with MWI by Prasloski et al. [124], is the issue of imperfect pulse refocusing. It is common for a scanner to have a flip angle that in practice deviates from the ideal of 180° due to B1 effects. This question is addressed in this chapter by investigating how a change in the flip angle affects the ability of the LWI to fit the tissue and test a method for correcting these errors.

Both of these experiments aim to use the available dataset to improve the acquisition and modelling for future studies. If large clinical studies are to be carried out using LWI then the whole process needs to be thoroughly investigated and the scan parameters fixed into the future.

6.1 Sequence optimisation

Chapter 4 details how the model and pulse sequence were initially chosen, using experiments to determine the optimum number of echoes and echo spacing of the multi-echo T2 sequence. These investigations were based on preliminary datasets, including simulations, phantoms and a small number of in-vivo cases, and were designed to give the best values possible for the larger T2 study included in Chapter 5.

However, now this study has been carried out, the data can be used to further refine the pulse sequence used for LWI modelling. This section aims to investigate whether a further reduction in the number of echoes in the sequence, reducing SAR and providing the opportunity to improve volume coverage or resolution, based on the fact that a reduction in echo number will lead to a decrease in scan time. This work was done in collaboration with Fiona Gong, also from the Centre of Medical Imaging at UCL. My contributions were to provide the image dataset, regions of interest, histology information and the Matlab code for LWI modelling.

Nineteen men were imaged using the sequence parameters in Table 4.8. All patients underwent targeted biopsy of suspicious lesions and the contralateral prostate. Following biopsy, a radiologist contoured 31 regions of interest (both healthy and lesion) on the clinical T2-weighted images. The LWF maps were generated as previously described using the 32-echo sequence as well as three other maps using the first 16, 8 and 6 echoes of the 32 echo dataset respectively. In reducing the number of echoes the range was also reduced, rather than keeping the range the same and removing alternating echoes, as the aim of this experiment was to reduce the TR and hence the scan time.

An ROC analysis was performed on the median LWF values from each ROI to evaluate the classification accuracy of LWF in differentiating between cancer and benign tissue.

In order to visualise how the reduction in echoes is affecting the fitting of the LWI model, an example pixel in both benign tissue and tumour is fit using 32, 16, 8 and 6 echoes and the total signal decay is shown alongside the model fit.

For the 19 patients in this dataset the ROC curves were similar when decreasing the number of echoes (Figure 6.1), with only a marginal decrease in AUC (0.85, 0.82, 0.81 and 0.80 for 32, 16, 8 and 6 echoes respectively). A further analysis needs to be done on what effect this small reduction in AUC would have on the incorrect diagnosis of patients but these results are encouraging.

For the example pixels, the goodness of the LWI fit was comparable between 32 and 16 echoes but was visibly reduced for 8 and 6 echoes in both benign tissue and cancer (Figures 6.2 and 6.3 respectively). These images show the full range of datapoints for the ME-T2 acquisition but each fitting is calculated based upon the appropriate number of echoes. LWF decreased with a reduction in the number of echoes for both benign and cancerous regions.

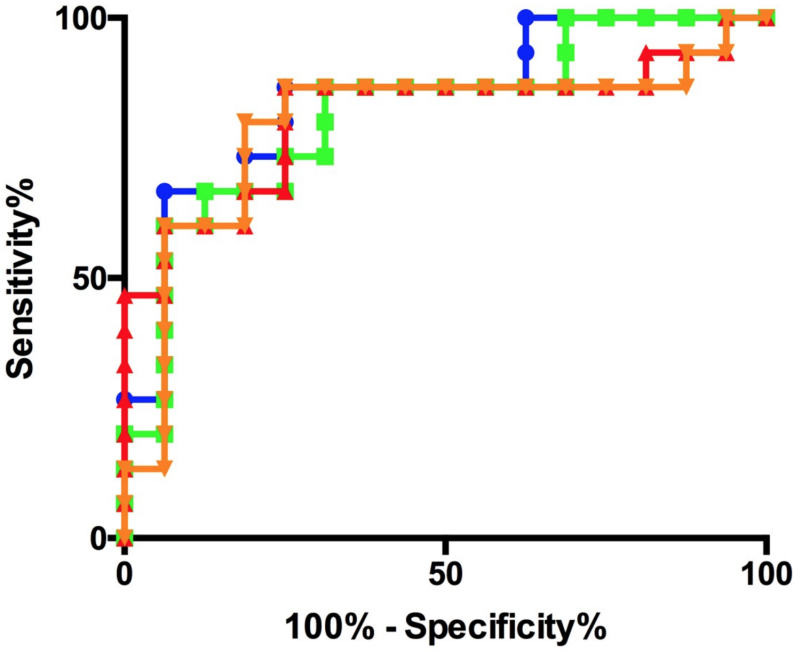


FIGURE 6.1: ROC curves for 32-echo (blue), 16-echo(green), 8-echo(red)and 6-echo (orange) datasets.

The optimisation of the protocol is important when moving towards a technique that can be used in clinic. These results show that a reduction in the echo train length has

only a small effect on the AUC of LWI in classifying tumours and normal tissue, despite the reduction in the goodness of fit shown in the example pixels. This suggests that the scan could be reduced from its current time of 5m 49s, which could enable greater coverage and/or enhanced resolution in the images. It is estimated that using a 6-echo acquisition a 6 minute scan could acquire 20 slices with a resolution of 1.6x1.6x4mm, an improvement on the 6 slices of 2x2x4mm resolution that was acquired for the previous studies in this work. This improvement in resolution would make it easier for clinicians to accurately locate lesions and hence make the method more likely to be taken up clinically. Future work is required to prospectively validate the clinical performance of a reduced number of echoes and to investigate how the SNR changes with these alterations to the number of echoes.

6.2 Imperfect pulse refocusing

Multi-echo spin-echo sequences are vulnerable to the imperfect refocusing of the RF pulses. A perfect refocusing pulse would cause the spins to flip exactly 180° but, due to inhomogeneities in the B_1 field, this angle can be altered. These field inhomogeneities lead to a signal decay that is no longer a true T2 decay as not all of the spins refocus at the same time, meaning that the maximum number of spins refocusing at any one time is altered.

The longer the echo train the greater this effect will be as the imperfect flip angles accumulate, which is a secondary motivation for the shortening of the MESE sequences that has been discussed in previous chapters. Furthermore, future work will aim to reduce the scan time of the ME-T2 acquisition in order to improve resolution and/or coverage. In order to do this, one option is to reduce the flip angle and then compensate for this using post-processing.

As shown in the brain when modelling the Myelin Water Fraction [124], correcting for these imperfect pulses can lead to more accurate model parameters. Hence this chapter seeks to investigate the effects that imperfect refocusing may have on the effectiveness of LWI.

6.2.1 Background

Imperfect refocusing is a problem when using quantitative modelling on T2 data because an alteration in intensity across the signal decay that increases with TE, which would ordinarily be attributed to lower or higher mean T2 values for each component within the tissue, could actually be down to imperfect refocusing. If any technique is to be used clinically it must be able to work multiple times on the same scanner as well as across other centres but because of slight differences in the B_1 fields of each scanner some of the quantitative parameters produced by the model would be inherently different. The B1 field in the prostate depends upon the properties of the surrounding tissue. Whilst the scanner will try to calibrate the B1, it cannot do this perfectly.

Figure 6.4 shows an example in the brain of how different refocusing flip angles affect the shape of the signal decay. Though the difference between the curves may visually seem small for minor differences in the flip angle, the nature of the fitting models means that these deviations can cause large differences in the model parameters, particularly in the first few echoes where all of the signal from the short T2 component is found.

In order to understand the method for correcting these errors it is important to first look at what happens to spin populations within the patient during the scan. When an RF pulse is applied to a population of spins that are all in phase, the resultant magnetisation can be decomposed into several components:

- M_+ - rephasing transverse magnetisation
- M_- - dephasing transverse magnetisation
- M_z - longitudinal magnetisation

When the refocusing pulses are exactly 180° echoes are produced as in Figure 6.5. As seen in a MESE sequence, when the flip angle deviates from the ideal value the magnetisation is split between the three components, as shown in Figure 6.6. Not only does this reduce the size of the echo by reducing the size of the refocusing component but it means that the dephasing and longitudinal components could cause extra, unwanted

echoes. In practice, every RF pulse can be thought of as acting like three pulses, a 0°, a 90° and a 180° pulse.

Under the assumption that the refocusing pulse is perfectly rectangular with amplitude B_1 and duration t_p, the refocusing angle α can be estimated as:

$$\alpha = \gamma B_1 t_p \tag{6.1}$$

where γ is the gyromagnetic ratio. If t_p is held constant for varying values of B_1, this equation can be rearranged through substitution as:

$$\alpha = 180 \frac{B_1}{B_1^{180^\circ}} \tag{6.2}$$

In order to get an estimate for the actual flip angle α in degrees, the value for $B_1/B_1^{180^\circ}$ must be acquired in each individual subject. This can be done using a number of different methods but in this work representative values were taken from 'Water Shift and B1' (WASABI) data acquired by a colleague for a separate CEST project [141]. B_1 maps were created in the prostate for ten subjects using images from a Philips Achieva 3T MRI scanner. In these subjects the B_1 field was found to be in the range 90-94% of the expected value when measured across a single-slice ROI encompasing the whole prostate. Hence, using a value of 0.92 for $B_1/B_1^{180^\circ}$ and using Equation 6.2 the flip angle α would be 165.6°.

Though it is possible to alter the pulse sequence to ensure that B_1 is equal to $B_1^{180^\circ}$, this is not a simple process so it would be a more reliable technique if the model could cope with this effect. A complex version of the Bloch rotation matrix [70] can be implemented using the Extended Phase Graph (EPG) algorithm. This algorithm calculates the states of a number of the different spin populations (M_+,M_- and M_z) as they undergo refocusing and relaxation during the MESE sequence.

After each refocusing pulse the proportions of the spins in each one of these populations changes. Hence the EPG algorithm iterates through each pulse to calculate the population proportions as they change through time and from that calculates the expected

		Assumed flip angle					
NE	LWF	100°	120°	140°	160°	180°	200°
	0.2	0.0517	0.1486	0.1802	0.1909	0.1995	0.1909
32	0.1	0.0431	0.0813	0.0902	0.0920	0.0985	0.0920
	0	0.0100	0.0111	0.0007	0.0004	0.0008	0.0005
	0.2	0.1127	0.1676	0.1857	0.1927	0.2017	0.1927
64	0.1	0.0670	0.0875	0.0926	0.0940	0.0986	0.0940
	0	0.0013	0.0011	0.0008	0.0007	0.0011	0.0007

TABLE 6.1: LWF values for simulation using 32 and 64 echo scan data respectively for an actual flip angle of 180°.

proportion of the spins that will rephase correctly. The equations for calculating the actual signal intensity is described in detail in a previous work [124].

The size of the effect that the EPG algorithm shown previously [124] has on the modelling of the T2 spectrum depends upon how far the refocusing angle differs from 180°. To test the effects, a number of experiments were devised whereby the results of the LWI fitting were compared with and without the use of the EPG algorithm.

6.2.2 Simulation

First, an experiment was set up whereby simulation pixels were created to investigate whether using the EPG algorithm significantly changes the estimates of the LWF. These results should inform us as to whether the LWI technique is sensitive to imperfect pulse refocusing and to what degree various actual flip angles affect the model.

The underlying tissue was simulated using a two Gaussian T2 distribution with simulated flip angles of 160° and 180°. The values of LWF were varied between 0-0.2 and two pulse sequences were simulated, one with 32 echoes and a TE of 31.25ms and the other with 64 echoes and a TE of 25ms. The flip angle assumed by the model itself was also varied between 100° and 200° in order to see how changing the assumed angle in the EPG affected the approximation of the LWF. All other values were fixed (SNR=100, $T2_{short}$=50ms, $T2_{long}$=500ms).

The EPG algorithm was applied to the datapoints prior to the LWI modelling. The EPG algorithm also requires a T1 value, which was set at 1500ms due to previous studies

		Assumed flip angle					
NE	LWF	100°	120°	140°	160°	180°	200°
32	0.2	0.0780	0.1464	0.1593	0.2015	0.2077	0.2015
	0.1	0.0549	0.0763	0.0837	0.0971	0.1019	0.0971
	0	0.0176	0.0153	0.0046	0.0265	0.0230	0.0342
64	0.2	0.1114	0.1455	0.1722	0.2032	0.2061	0.2032
	0.1	0.0693	0.0802	0.0850	0.0999	0.1018	0.0999
	0	0.0171	0.0108	0.0043	0.0531	0.0354	0.0356

TABLE 6.2: LWF values for simulation using 32 and 64 echo scan data respectively for an actual flip angle of 160°.

having shown that this is approximately the mean T1 value for PZ tissue [87]. This value would differ throughout the prostate so the most accurate way to do this would be to register a map of T1 values to the 32-echo images and use pixelwise T1 values. However, preliminary test suggest that deviations of approximately 200ms, which is the estimated standard deviation of the T1 values in the PZ [87], do not substantially change the LWF values output by the fitting algorithm when using the EPG.

The results in Table 6.1 show the estimated LWI values when the flip angle of the simulated data is set to 180° and the results in Table 6.2 show the estimated LWI values when the flip angle of the simulated data is set to 160°. In Table 6.1, the EPG is only having a small effect on the LWF values produced by the LWI model when the flip angle deviates between 160-200°. Outside of these angles, particularly for higher LWF values, the EPG is having a much larger effect. Similarly for Table 6.2, changing the angle used by the EPG model by ±20° from the ideal 180° case shows that for small deviations in the flip angle the EPG is having little impact on the ME-T2 signal decay, and hence the LWF value that is produced by the LWI model.

When comparing Table 6.1 and 6.2, the most accurate LWI estimates when the actual flip angle is 180° are when the EPG also assumes a flip angle of 180°. Similarly, the most accurate LWI estimates when the actual flip angle is 160° are when the EPG also assumes a flip angle of 160° when the LWF is 0.2, but this is not the case as the intrinsic simulated LWF decreases.

When taking an imaging technique from an experimental one to clinical use, multi-centre studies must be carried out in order to test how it works on different scanners.

As each scanner and subject is different it is important that LWI works for various flip angles. This investigation points to the conclusion that imperfect refocusing leads to insignificant changes in LWF when compared to the difference between PCa and normal prostate tissue as long as the scanner's actual flip angle is within the range 160-200°.

6.2.3 In-vivo

65 subjects were imaged using the T2 sequence scan parameters in Figure 4.8. For each subject an ROI of both tumour and normal tissue were contoured and maps of the two Gaussian model parameters were created using the EPG algorithm with different values of the assumed flip angle. The mean LWF parameter values for each ROI were produced and compared.

Then a comparison was made between the LWI parameter values produced using the EPG algorithm and those without it. The ability of the mean LWF values to differentiate between PI-RADS v2 score groupings was assessed using both a logistic regression and ROC analysis.

The p-values from the logistic regression and the AUC, sensitivity and specificity of each test from the ROC analysis are shown in Table 6.3. The AUC values in particular show that if the EPG model is used with an assumed angle of 200° then the AUC of the test increases when differentiating between PI-RADS 1,2v3 and PI-RADS 3v4,5. However, as the true flip angles are unknown, all that can be concluded is that when EPG corrections are applied, the analyses vary enough to suggest that further investigation is required.

The purpose of this study was to look at the effects on the LWI model of imperfect refocusing. The results show that use of the EPG model does affect the ability of the LWI model to differentiate between different PI-RADS score groupings. Unlike the simulation data, these results suggest that the imperfect refocusing problem should be addressed in future studies.

Different scanners have different refocusing angles and, though it should be exactly 180°, could easily vary from 160-200° between scanners and subjects. This would mean that the sensitivity and specificity of LWI in differentiating tumour from normal tissue could

Angle	PI-RADS Groupings	p	AUC	Sens. (%)	Spec. (%)
200°	1,2 v 3	0.030	0.804	80.7	87.2
	3 v 4,5	0.025	0.907	93.3	92.0
	1,2 v 3,4,5	0.000	0.802	71.7	83.4
180°	1,2 v 3	0.005	0.786	63.3	84.0
	3 v 4,5	0.013	0.867	80.0	76.7
	1,2 v 3,4,5	0.000	0.881	80.6	78.7
160°	1,2 v 3	0.015	0.719	82.7	65.0
	3 v 4,5	0.037	0.878	81.7	88.3
	1,2 v 3,4,5	0.000	0.798	72.4	80.1
140°	1,2 v 3	0.489	0.594	71.4	61.3
	3 v 4,5	0.012	0.850	90.0	83.3
	1,2 v 3,4,5	0.011	0.706	81.0	65.9
120°	1,2 v 3	0.742	0.636	69.2	76.3
	3 v 4,5	0.427	0.742	78.3	78.3
	1,2 v 3,4,5	0.511	0.600	56.1	78.6
100°	1,2 v 3	0.476	0.609	60.1	82.1
	3 v 4,5	0.264	0.748	87.0	85.3
	1,2 v 3,4,5	0.726	0.617	76.5	61.0

TABLE 6.3: A comparison of the ability of LWI parameter values produced using the EPG algorithm to differentiate between PI-RADS v2 score groupings.The p-values are from a logistic regression and the AUC, sensitivity and specificity values are from an ROC analysis.

be changing on the order of 93.3% to 80.0% and 92.0% to 76.7% respectively when differentiating PI-RADS 3v4,5 by using the EPG with an angle of 200°.

6.2.4 Discussion

The EPG work was designed to investigate that imperfect refocusing is affecting the T2 signal and that, by correcting for that, better quantitative estimates of tissue microstructure are achieveable. The simulation study suggested that the EPG had little effect on the parameters of the LWI model when the actual flip angle deviated between 160-200°. However, the larger in-vivo study provides a good indication that the use of the EPG model with assumed flip angles 20° either side of the ideal, whilst the p-values remain below 0.05, could strongly affect the clinical utility of the technique. Hence the results indicate that fully taking into account the true flip angle could lead to better fitting. Whilst operating at a nominal 180° provides reasonable results without having to make corrections, this work suggests that imperfect flip angles should be taken into account to maximise the performance of LWI.

This experiment has a number of drawbacks. The flip angle calculated using the WASABI method is an approximation and the flip angle also varies spatially due to inhomogeneities in the B_1 field so, even if the average flip angle were correct, the only way to perfectly correct for imperfect refocusing is to calculate a flip angle for each pixel individually. However this study seems to suggest that this is not necessary if the variation is small. The EPG model is also subject to assumptions and approximations that mean it is not perfect at calculating the correct signal decays for a given flip angle. This experiment is modelling the effect of changing the assumed flip angle, not the actual flip angle. Hence, whilst the experiment does not change the actual flip angle it should give an idea of sensitivity to mis-matched actual and assumed flip angles.

This work has shown that there is a need for further investigation into using methods for correcting imperfect pulse refocusing in scanners. This should take the form of a multi-center study as the benefits of the EPG to improving the reproducibility of LWI needs to be investigated. Comparing the capabilities of LWI both with and without imperfect refocusing correction will show how LWI varies across scanners and whether the EPG method can correct for this.

6.3 Summary

The objective of this chapter was to use the 65 patient dataset acquired to improve the imaging and model fitting to maximise the ability of LWI to detect and grade prostate cancer. The methods for doing this were twofold, investigating how far the echo train could be shortened whilst still giving good LWI estimates and looking into whether imperfect refocusing is having an effect on the LWI estimates.

The first experiment has shown that it is possible to reduce the unmber of echoes acquired still further from 32 echoes down to as little as 6 echoes. This is a strong result for the LWI method as it shows how robust it is to a reduction in the dataset. It also opens up the possibility of speeding up the scans and possibly improving upon resolution. This would make the multi-echo T2 images look a lot closer to the anatomical T2 images that clinicians currently use.

Given that the results when detecting and grading prostate cancer are already matching the current best in our dataset, this method can only improve in its ability to perform these tasks when a larger dataset is used. These works lay the ground work for an optimised acquisition and a perfected LWI modelling.

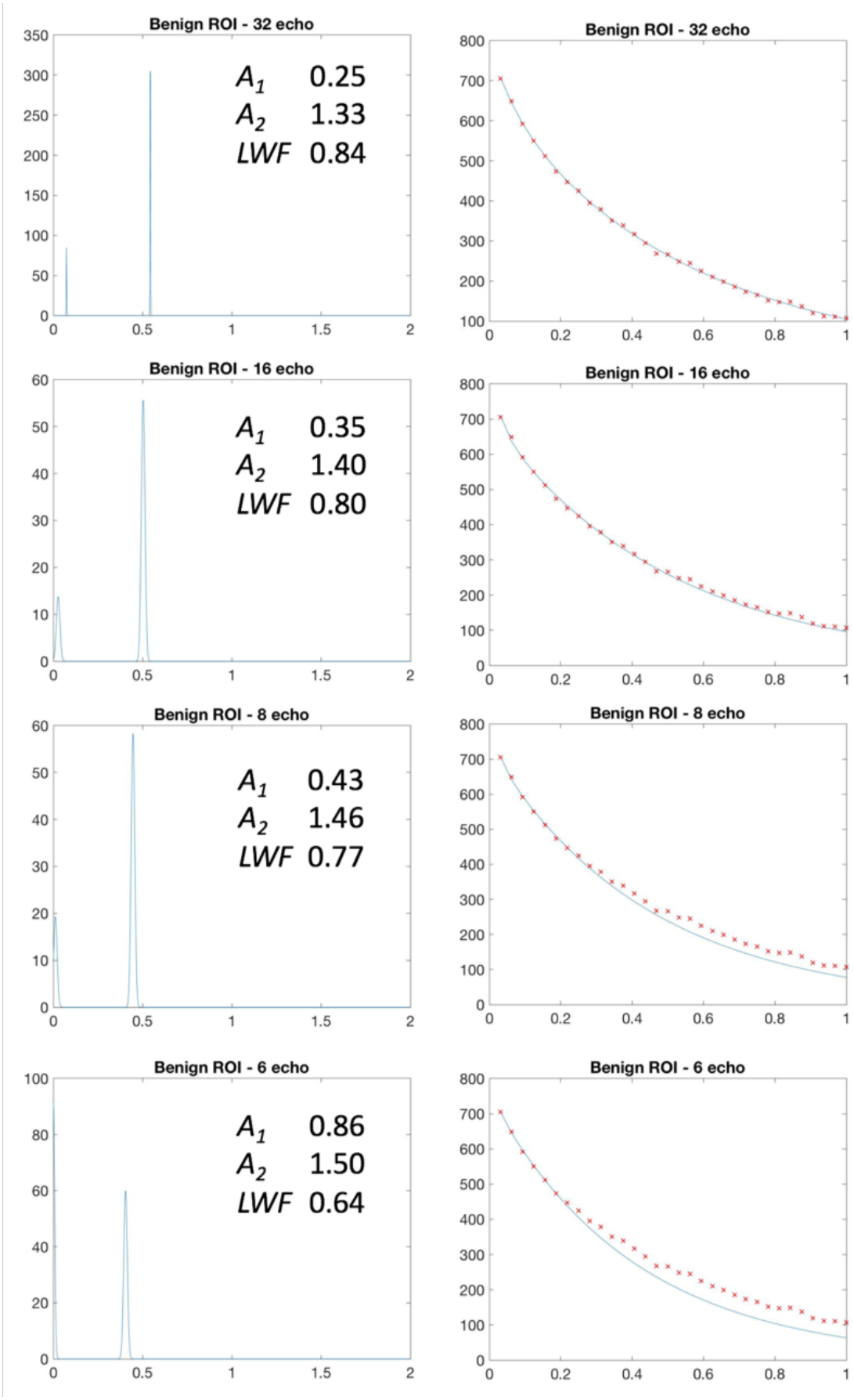


FIGURE 6.2: T2 signal and LWI fitting for pixel of benign tissue. The images on the left are the fitted T2 distribution of the pixel and the image on the right is the T2 signal decay alongside the model fit.

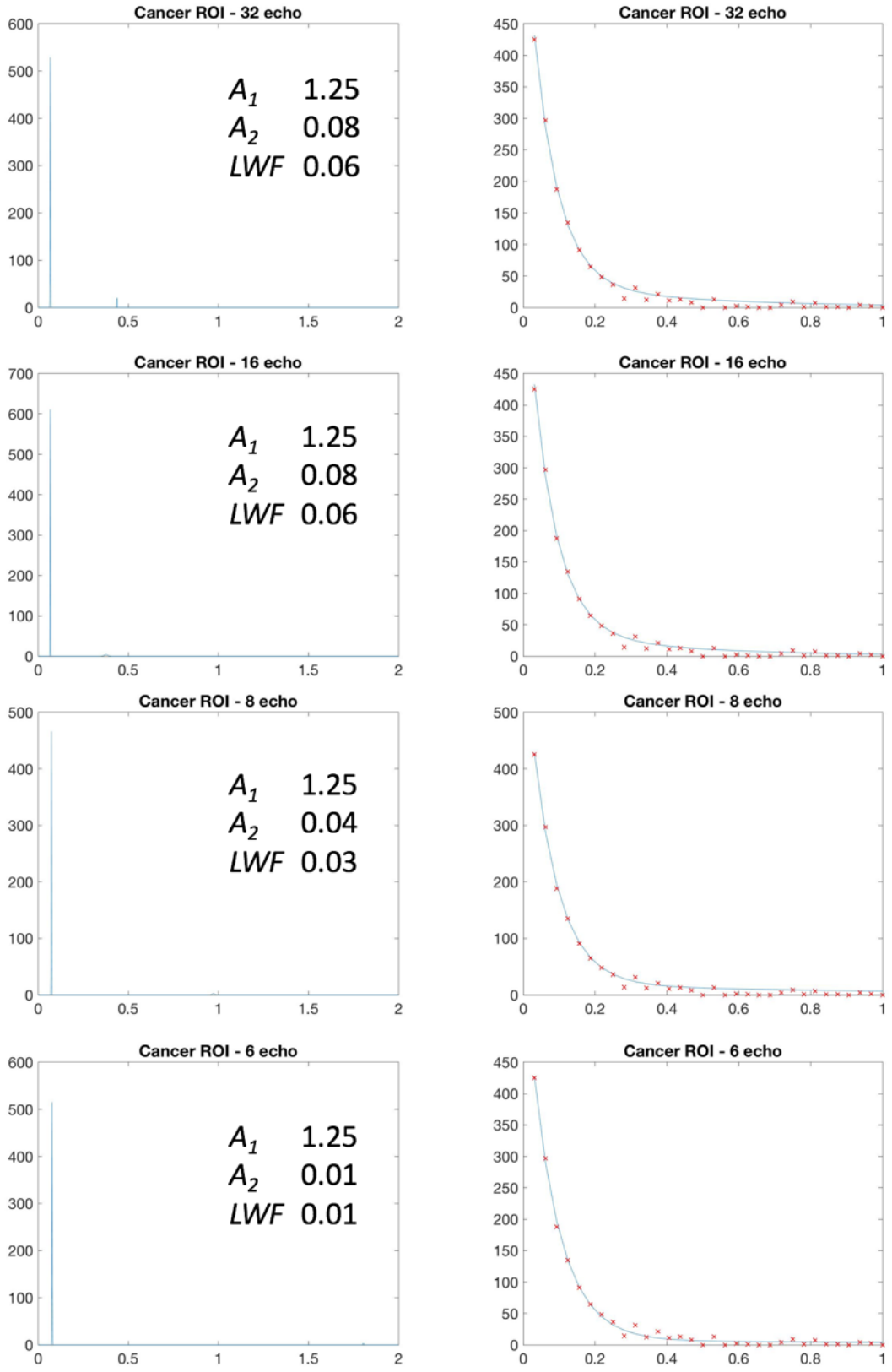


FIGURE 6.3: T2 signal and LWI fitting for pixel of cancerous tissue.The images on the left are the fitted T2 distribution of the pixel and the image on the right is the T2 signal decay alongside the model fit.

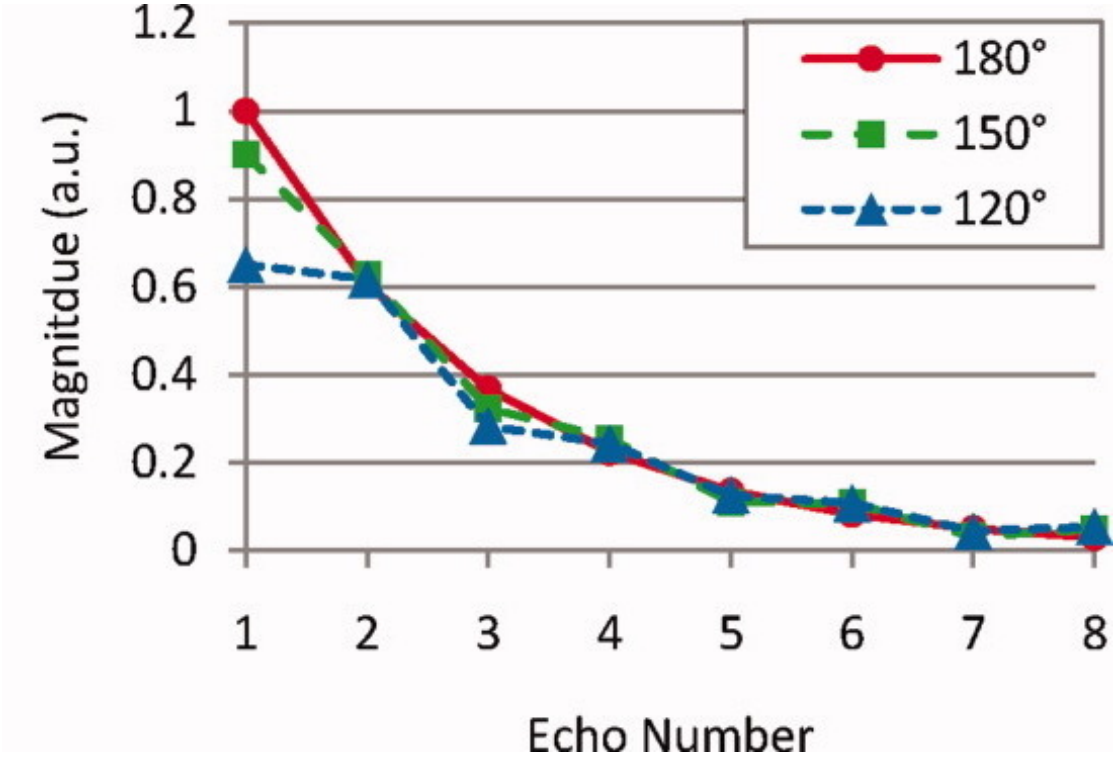


FIGURE 6.4: First eight echoes of the signal decay of the myelin component for three refocusing flip angles. All of the magnitudes are normalized to the value of the first echo for the 180° case (reproduced from Prasloski et al. [124] with permission).

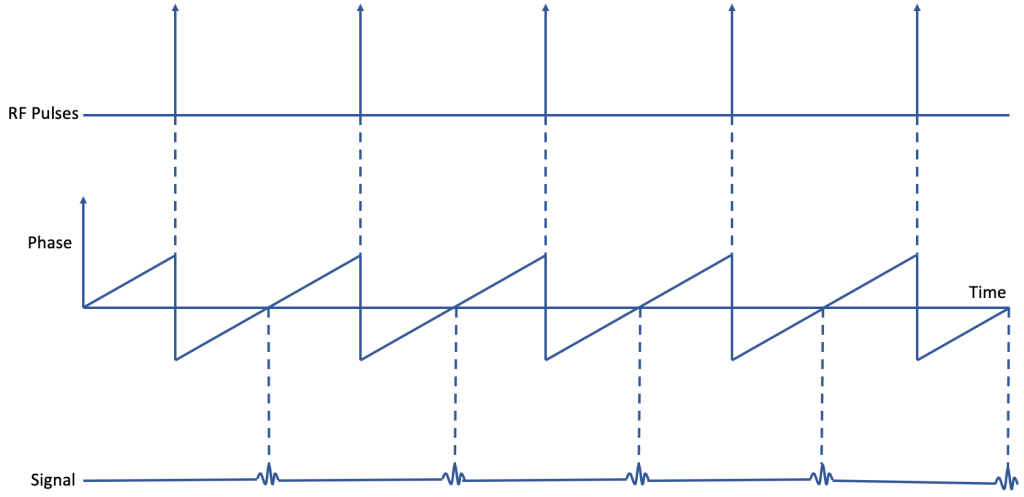


FIGURE 6.5: A graph showing the first five pulses of a CPMG pulse train with perfect 180° pulse refocusing, including the phase graph and the resultant signal showing the five corresponding echoes.

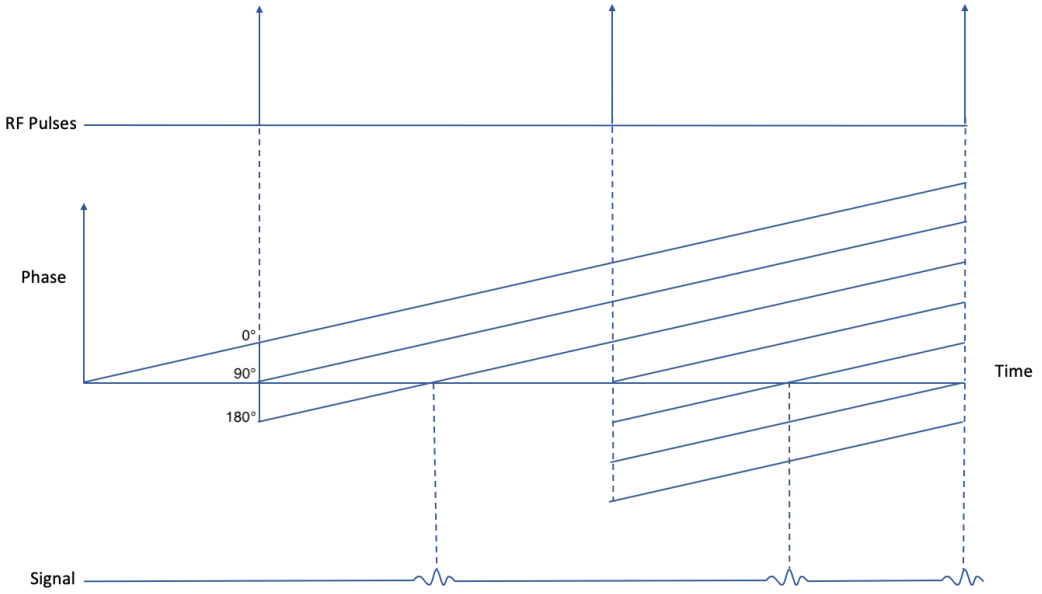


FIGURE 6.6: A graph showing the first three pulses of a CPMG pulse train with imperfect pulse refocusing, including the phase graph and the resultant signal.

Chapter 7

Machine Learning for Prostate Cancer Detection

Machine learning is a rapidly growing technique in classification tasks. The work contained in this thesis so far is the classification of regions of interest into a number of categories. This work looks at using machine learning for these classification tasks because they remove previous model assumptions on the data.

This chapter first presents a comparison of machine learning techniques in order to assess which holds the most promise for classifying tumours using T2 data. Then a larger study into the use of the best of these techniques in classifying regions of interest is carried out. Parts of this work were presented at the 26th Postgraduate Symposium of the British Chapter of the International Society of Magnetic Resonance in Medicine

7.1 Comparison of Machine Learning Techniques

Multi-parametric MRI (mp-MRI), which includes T2-weighted, diffusion and DCE measurements [151], provides good sensitivity and specificity in the detection of clinically significant prostate cancer [6]. However, the accuracy of mp-MRI is limited by variability in and the subjectivity of human radiological interpretation of images. Overall, mp-MRI classifies a significant proportion of lesions as indeterminate, meaning that the

method cannot tell whether a large number of cases are clinically significant prostate cancer. Storas et al. [152] found that a multi-echo T2 sequence can separate the effects of different tissue micro-environments. This work examines the ability of both Support Vector Machines and simple Neural Networks to differentiate tumour from normal tissue using multi-echo T2 data. The cutting-edge machine learning techniques, such as convolutional neural networks, could be effective in this problem due to their ability to take spatial information into account but this study aims to use simpler techniques as a proof of concept.

7.1.1 Methods

A small experiment was first carried out in order to decide which machine learning technique showed most promise out of the subset of simple methods chosen to investigate. A sub-cohort of nineteen patients from a larger study met our inclusion criteria [78] as previously discussed in Section 5.3.2. These acquisitions were carried using the T2 sequence detailed in Table 4.8. A board-certified radiologist (FG) contoured twelve lesions across nineteen patients, using histopathological reports to contour on a single echo of the multi-echo images. Nineteen regions of interest (ROIs) were also placed in areas of normal tissue in each of the subjects, all of which were histologically sampled and did not show cancer. Two algorithms were run voxel-wise on the 32-echo data, consisting of 3962 voxels from thirty-one ROIs. A neural network with one hidden layer consisting of thirty-two inputs and ten fully connected neurons was trained, validated and tested on a voxel by voxel basis, classifying each voxel as either significant tumour (defined as Gleason 3+4 or above) or not. The same classification was then carried out using a support vector machine on both the multi-echo data and the T2 values calculated from a mono-exponential fit to the multi-echo data. Five-fold cross-validation was applied when training and testing the models. In order to compare the models, the average Receiver Operating Characteristic (ROC) curve across all folds was plotted for each classification algorithm. The accuracy was assessed using the Area-Under-Curve (AUC) value, as well as the sensitivity and specificity of the operating point.

Method	AUC	Sens. (%)	Spec. (%)
Neural network	0.82	84	72
SVM	0.54	59	50
SVM (mono-exponential)	0.62	88	46
mp-MRI (quoted from [6])	-	88	45
LWI (quoted from Chapter 5)	0.81	75	87

TABLE 7.1: The average results of the cross-validated ROC analyses along with results from other studies.

7.1.2 Results & Conclusion

As shown in Table 7.1, the multi-echo T2 sequence combined with a neural network classification has similar sensitivity and specificity in the detection of clinically significant prostate cancer (defined here as Gleason 3+4 and above) to the LWI model investigated in Chapter 5. Higher specificity means that fewer normal regions are falsely predicted to be significant tumour. These values suggest that multi-echo T2 imaging shows promise in detecting clinically significant prostate cancer using a simple imaging technique that is less prone to artefacts than diffusion sequences. This improved sensitivity and specificity may come from the fact that the multi-echo data can more accurately represent multiple compartments than ADC or traditional axial T2 images. This experiment has shown that neural networks could be more effective in detecting prostate cancer than SVMs and so were selected for further research.

In order to better investigate the difference between the machine learning techniques a prospective study chould be designed using a larger dataset in which the radiologist is blinded to previous histopathology, and more sophisticated neural networks could also be applied.

7.2 Neural networks for scoring of tumours

Following on from the work in Section 7.1, in this section of work neural networks were investigated further as a way of detecting prostate cancer. Neural networks are versatile and robust machine learning algorithms used in a wide range of classification tasks, as described in Section 2.6. The use of neural networks is not common in prostate MRI

and no works have used them on either multi-echo T2 data or Luminal Water Imaging parameters. The aim of this work is to investigate whether neural networks trained on only the multi-echo T2 signal values can produce comparable results to current diffusion imaging methods and the LWI method from previous chapters. Neural networks were optimised for each dataset and then tested for their ability to differentiate pixels within ROIs graded PI-RADS 1 & 2 from pixels within ROIs graded PI-RADS 3,4 & 5 using an ROC analysis. The multi-echo T2 dataset was found to be as effective as the LWI parameters in differentiating between these PI-RADS scores and more effective than either the diffusion-weighted images or ADC maps.

7.2.1 Introduction

As discussed in previous chapters, mp-MRI is one of the best methods available for detecting and grading tumours non-invasively using MRI but it suffers from issues such as low specificity. This leads to a large number of men unnecessarily undergoing biopsy, a painful and invasive procedure. The clinical need for new modalitles that better distinguish between different grades of tumour is clear.

Two ways in which to alter and improve a particular imaging method are either to change the acquisition or to change the model imposed upon the acquired data. Better data can provide more information on the underlying tissue but is of limited use if the signal model is too simple to capture it. However, a more intricate model is of limited use if the data are not rich enough to support it in a robust and repeatable way.

Whilst studies have been made on using neural networks on voxel-wise diffusion or T2-weighted data to detect and grade prostate tumours, this study has extended these investigations to multi-echo T2 data. It has been shown [174] that using convolutional neural networks on mp-MRI images to produce lesion probability maps could have similar sensitivity and better specificity than a 6-core prostate biopsy in localising prostate lesions. Yang et al. then go on to show that these methods also show promise in being able to grade these lesions accurately. However, this technique is yet to be repeated and still only uses the images acquired in a standard mp-MRI acquisition.

This chapter investigates the use of machine learning on 32-echo T2 and diffusion data. In particular, this study compares the ability of different combinations of MR data in classifying pixels in an image into clinically relevant groupings of PI-RADS v2 scores using neural networks. Neural networks are tested on the data from 6 b-value diffusion acquisitions, 32-echo T2 acquisitions, ADC maps and the parameters of the LWI model. This hypothesis of this work is that the use of neural networks on a range of different acquisitions and maps removes the need to impose a model on the data whilst providing comparable detection and grading of PCa.

7.2.1.1 Patient Selection

A sub-cohort of 24 patients was used from a larger prospective study [78] and on 23rd December 2015 received UK Research Ethics Committee approval. Patients were recruited between September 2016 and October 2017 and provided written informed consent on the day of the trial intervention following a minimum 24-h period of consideration. The patient inclusion criteria were included in a previous study [78]. 46 subjects were excluded, 1 due to a technical fault, 4 due to MR contraindications, and a further 41 due to poor registration between the diffusion and T2 images.

7.2.2 MRI Acquisitions

Subjects were scanned on a 3.0T scanner (Philips Achieva; Philips Medical Systems, Best, The Netherlands) using a 32-channel cardiac coil. A multiple b-value DWI was acquired with a Single Diffusion Encoding (SDE) single shot EPI sequence over 6 b-values of 0,90,500,1500,2000,3000 s/mm2. TE/TR = 50-80/2000-3707 ms; Voxel size = 1.3x1.3x5mm; Field of View = 220x220 mm; Scan duration = 12:57. A multi-echo spin-echo sequence was acquired with parameters: TE/TR = 31.25/8956 ms; Number of echoes = 32; Acquired voxel size = 2x2x4mm; Field of View = 180x180 mm; Scan duration = 5:50.

7.2.2.1 Regions of Interest and histologic examination

A board-certified radiologist with 5 years of experience in prostate mpMRI reporting (FG) contoured 98 regions of interest (ROIs) of either malignant or benign tissue on the ADC maps using previous information on the location of lesions, including from previous biopsy where available. 97 ROIs were also contoured on the 93.75ms echo of the ME-T2 image, chosen for its similar echo time to a traditional axial T2 weighted prostate image (∼100ms). For all ROIs, both benign and lesion, a PI-RADS v2 score was assigned. PI-RADS v2 is a method of scoring tissue on a scale of 1-5, with 1 meaning clinically significant cancer is highly unlikely to be present and 5 meaning clinically significant cancer is highly likely to be present. A subset of 19 patients also had histology information available.

7.2.2.2 Image Transformation

In order to use neural networks to classify pixels using both diffusion and T2 data combined, the images must be correctly aligned so the pixels in the diffusion image match the pixels in the T2 image. The main source of misalignment between DWI and T2-weighted images is inhomogeneities in the B0 field. These inhomogeneities cause echo planar imaging (EPI) based images to become distorted when compared to T2-weighted images. In order to undistort the diffusion images a B0 map was acquired as part of the scan and, using the bandwidth/pixel ratio calculated by the scanner, the shifting of the EPI-based diffusion imags due to these inhomogeneities could be calculated. After distortion correction the DWI images were then resliced into the same plane and resolution as the T2-weighted images using 3D interpolation with a separable kernel that accounted for in-plane voxel sizes and slice thickness.

Once the distortion was removed and the images had been resliced each DWI and T2-weighted image was assessed against the ROIs drawn on the T2-weighted image. In 41 subjects this process of unwarping and reslicing did not provide adequately registered images. This study relies on pixel-by-pixel predictions, meaning that a pixel in a diffusion image needs to be registered to a pixel in a T2 image. As the process of registering images

more thoroughly is a complex and time-consuming process, for this study the decision was made to visually reject those subjects for which the images were poorly aligned. This led to the number of subjects in the study reducing to 24, which contained 37 ROIs in all.

7.2.2.3 Neural networks

Neural networks were used in this work due to their versatility. 'A feedforward network with a single layer is sufficient to represent any function, but the layer may be infeasibly large and may fail to learn and generalize correctly' [61].

The structure of a neural network needs to be optimised for a task if it is to be effective. Two important parameters of a neural network are the number of layers it contains and the number of nodes contained within each layer. This was taken into account by testing a wide range of neural networks in order to find the optimum for each training dataset.

For each set of parameters the number of layers was iterated between 1-4 and the number of nodes in a given layer was iterated between 1-61. In a neural network the number of nodes in a particular layer need not be the same as that of other layers but, as changing the number of nodes in each layer greatly increases the complexity of the optimisation, the number of nodes was the same for each layer.

A fully-connected feed-forward network with sigmoid hidden neurons and softmax output neurons was created in Matlab (*MATLAB and Statistics Toolbox Release 2017a, The MathWorks, Inc., Natick, Massachusetts, USA*). Matlab was chosen for its simple and easy to us ML applications and the fully connected feed-forward algorithm was used as it is a simple and versatile technique. The scaled conjugate gradient backpropogation method was used to update the weights and biases in the network and the performance of the network was calculated using cross entropy.

The network is run pixelwise on the signal and parameter values so no spatial information is used. Each of the datasets is normalised in order for each of the input variables to have a zero mean and unit variance. Five datasets were separately tested in this way:

- 32-echo T2 - the signal values of the 32-echo acquisition.
- 6 b-value diffusion - the diffusion acquisition signal values averaged over 3 orthogonal gradient directions.
- LWI - the 6 fitted parameters of the LWI model along with the A_1, A_2 and LWF parameters calculated after the fitting.
- ADC - the apparent diffusion coefficient value calculated from the diffusion acquisition.
- T2 & diffusion signal intensities - a combination of the 32-echo T2 and 6 b-value diffusion acquisitions.

7.2.2.4 Statistics

In order to reduce the risk of the results being affected by the contents of the training, validation and test sets, cross validation was used. The patients were split up into five groups, which would consist of the five folds of the cross validation. Splitting by patient rather than splitting up into five equally sized groups of pixels ensures that pixels from the test patients are not seen in the training set, ensuring against overfitting in order to make the results more reproducible in the future. Whilst rotating through using each of the five folds of the cross validation as the test set, the rest of the data was itself split into training and validation sets over which the parameters of the fit, including the numbers of layers and nodes in the network, were optimised.

Mean sensitivity, specificity and area-under-curve (AUC) values across the 5-folds were computed using a Receiver Operating Characteristic (ROC) analysis. Sensitivity and specificity values were calculated from the ROC analysis using an operating point with the shortest distance to the point of perfect discrimination. A comparison between the scores PI-RADS 1,2 vs PI-RADS 3,4,5 separates lesions needing further action from those that do not.

Dataset	No. of Layers	No. of nodes	AUC	Sens. (%)	Spec. (%)
32-echo T2	3	2	0.874	85.3	86.8
6 b-val. diffusion	4	2	0.808	59.7	93.8
LWI	2	52	0.889	77.7	88.8
ADC	4	22	0.785	75.6	75.4
T2 & diffusion	4	42	0.774	58.6	92.1

TABLE 7.2: For each of the 5 datasets, presented in the table are the average results of the cross-validated ROC analyses alongside the number of layers and nodes in the neural network that produced these values.

7.2.3 Results

The results of the cross-validated ROC analysis are shown in Table 7.2. They show that both the 32-echo T2 values and the LWI parameters have a similar AUC when used by a neural network to distinguish between PI-RADS 1 & 2 and PI-RADS 3,4 & 5. Both of the diffusion datasets have lower AUC values.

Figure 7.1 shows the signal decays of 575 pixels from ROIs of both suspected healthy tissue and suspected tumour. The colours in this graph are based on the prediction from the final neural network, with blue representing PI-RADS 1&2 pixels and red representing PI-RADS 3,4&5 pixels. Similarly, Figures 7.2 and 7.3 show how these predictions performed against the ground truth PI-RADS v2 scores.

Figure 7.4 shows an example of the predictions made by the neural network. It correctly predicts the location of the suspected tumour in the left PZ. This image seems to show that the algorithm is predicting that all of the TZ is tumour, which is misleading because the neural network is mostly trained on PZ data in this experiment and hence does not correctly interpret TZ pixels.

7.2.4 Discussion

The results in Table 7.2 show that using a neural network trained on the 32-echo images alone to predict whether a lesion requires further action or not (PI-RADS 3-5 and PI-RADS 1-2 respectively) is as effective as using the parameters of the LWI model in the same way. This is an interesting result because it means that it can be used as a model-free method that requires no tissue assumptions and works on the image data

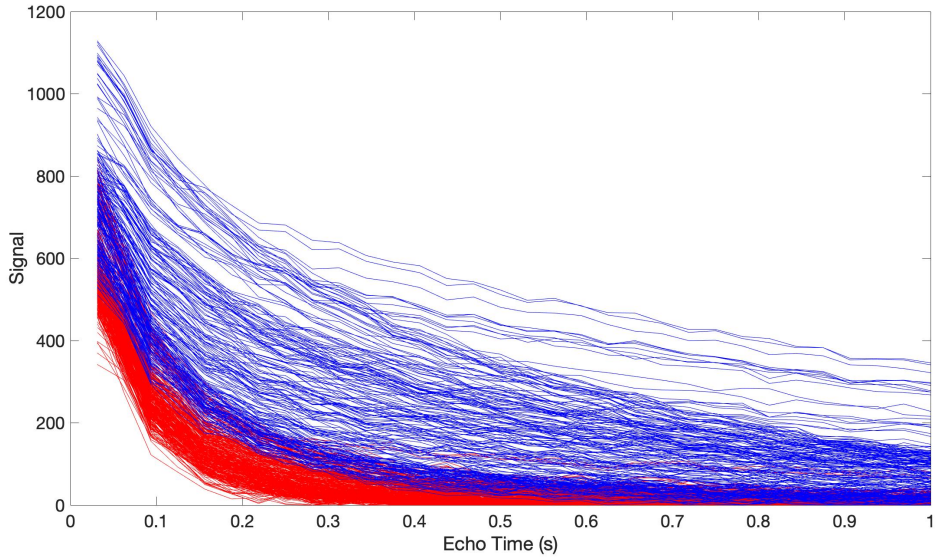


FIGURE 7.1: All signal decays for test set plotted against time. Red are the pixels predicted to be PI-RADS 3,4&5 and blue are the pixels predicted to be PI-RADS 1 & 2.

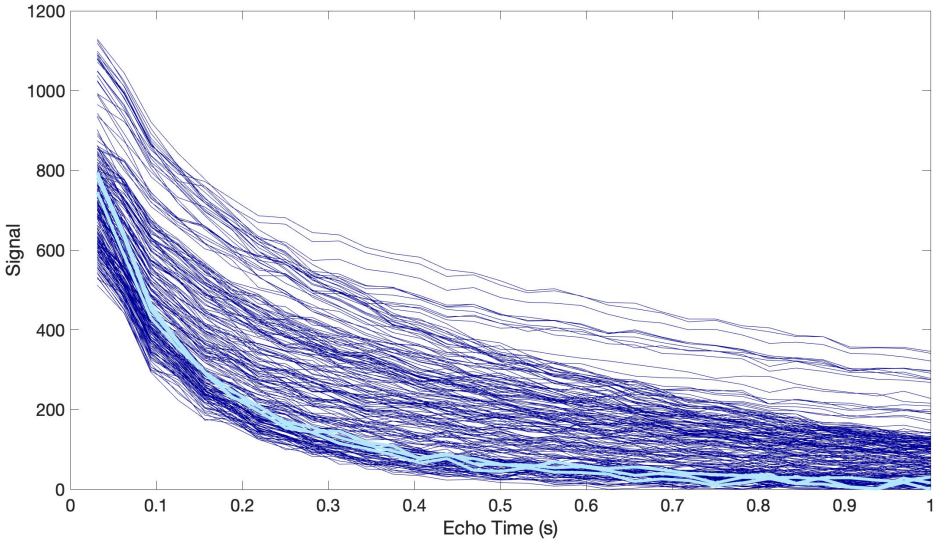


FIGURE 7.2: This plot depicts those pixels that were predicted to be normal tissue. The darker lines are pixels that were predicted correctly and the lighter lines those for which the neural network prediction was wrong.

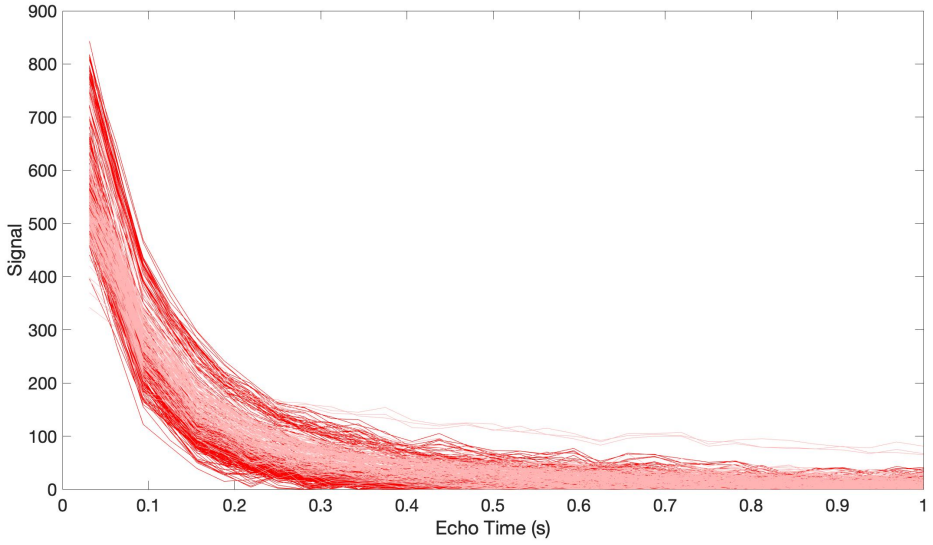
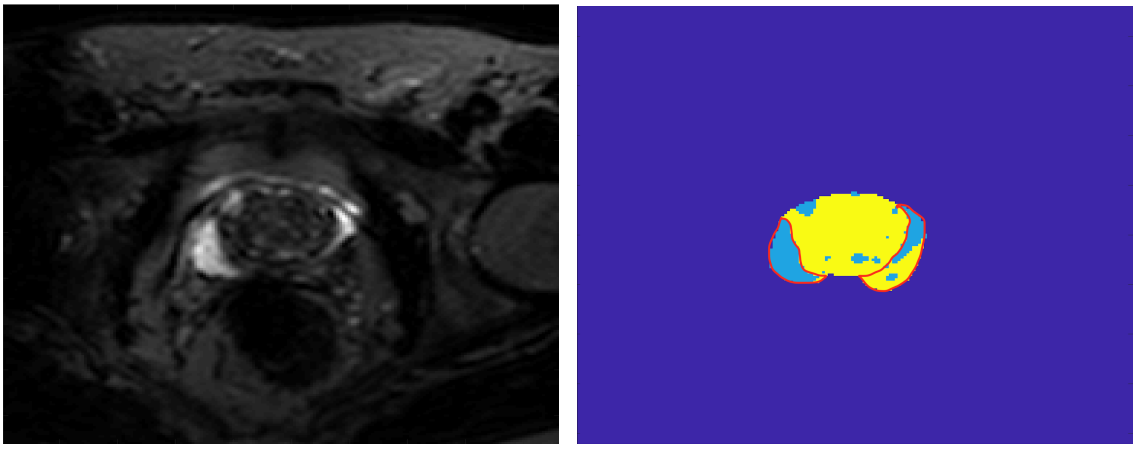


FIGURE 7.3: This plot depicts those pixels that were predicted to be tumour. The darker lines are pixels that were predicted correctly and the lighter lines those for which the neural network prediction was wrong.



(a) (b)

FIGURE 7.4: (a) This is a T2 weighted image of a prostate containing a large tumour in the left PZ. (b) The final neural network, using the 32-echo data alone, has predicted which pixels are normal tissue and which are tumour (tumour = yellow, normal = blue). The algorithm is trained on PZ tumours so although it is predicting the TZ to be mostly tumour this is misleading because the algorithm has not seen any TZ data. The PZ is highlighted in red.

itself. Once the model is correctly trained it will take less time to process an individual image than it would fitting the LWI model.

Table 7.2 also highlights some of the weaknesses of neural networks. When both the 32-echo T2 data and the 6 b-value diffusion data is combined the ability of the network

to differentiate between suspected healthy tissue and suspected tumour decreases. This is because as the number of input variables increases to 38 when they are combined it becomes more likely that the neural network will overfit the data and hence have poor results on the test sets. Furthermore, the size of the networks fit to the data reveals limitations of this study. The optimal network size on the 32-echo T2 and 6 b-value diffusion datasets is small, fitting only three and four hidden layers respectively. An explanation for this is that the larger neural networks are overfitting the training data and not generalising well to the test set. The fact that the largest dataset, the combination of T2 and diffusion datasets, is optimised by a network with 4 hidden layers of 42 nodes each but only has an AUC of 0.774 suggests that none of the networks are able to adequately model the data without more data.

Figure 7.1 shows that in general the pixels with the longer T2 values, and hence slower decays, are classed as normal tissue. However, due to the overlap in the middle it can be seen that the neural network is performing a more complex prediction than just selecting a T2 cutoff value. Figures 7.2 and 7.3 show that almost all of the erroneous predictions seem to be understandable given that the wrongly predicted pixels are visually hard to distinguish from those that are correctly predicted.

Figure 7.4 shows that the neural network looks to be correctly predicting the location of lesions in an example image purely based upon the 32-echo T2 data. As the network is trained only on regions of interest, rather than whole images, this goes some way to showing that the model is working correctly outside of these ROIs as well, as long as the tissue is in the same zone of the prostate. One thing to note is the network's prediction that all of the TZ tissue is tumour. This is due to the dataset, as none of the patients included in this study had tumours in the TZ, meaning that the neural network was not able to learn how to interpret TZ tissue.

One limitation to this study is that the dataset is by no means perfect for properly assessing the performance of a neural network. Typically these machine learning techniques are used on thousands if not millions of data points, so the size of this dataset is a limitation. The small size of the dataset means that the deeper networks are prone to overfitting the data, making them less effective when used on the test set. With

a larger dataset deeper networks could be used and better results could be obtained. Furthermore, the choice of structure of the neural network in this case was not the only one. Utilising different types of network could enhance it's ability to better differentiate between suspected lesion and suspected benign tissue. Another limitation is that the predictions in this experiment are scores from radiologists which are based on imaging rather than being histologically validated.

7.3 Summary

Overall this investigation into machine learning has shown that neural networks can classify tumours and benign tissue. In order to build on this result and further investigate the use of neural networks on multi-echo T2 data, future work would involve setting up a larger prospective study looking into the ability of this method to differentiate tumour from normal tissue. Histology is the current gold standard of lesion classification so, in order to fully validate this technique, the data must contain tissue information gained by either biopsy or prostatectomy. There is no correct minimum number of datapoints that are needed to use a neural network but in general the more datapoints the better if overfitting is to be avoided. If this method were to be used clinically, the number of datasets that the neural network is trained on would have to be larger. A wider variety of pathologies would also be necessary.

Chapter 8

Conclusion and Future Work

8.1 Conclusion

At the beginning of this project, when reviewing the literature surrounding quantitative tissue modelling in prostate cancer, three things were clear. One, the models that were in routine clinical use were very simple, which is to their benefit, but was still not specific enough to differentiate cancer from normal tissue in a proportion of cases. Two, many of the models that were being developed in the prostate to solve this problem were overly complicated, meaning that the dataset needed to robustly fit them was large. Three, models were being implemented in other regions of the body, particularly the brain, that were applicable to the prostate but were not being translated quickly.

The overall aim of this thesis was to improve the sensitivity and specificity of MRI methods in the detection and grading of prostate cancer through creating and testing tissue models that took into account the heterogeneous nature of prostate microstructure without requiring an infeasible acquisition through the model being overly complicated.

A number of opportunities arose from the literature review. Whilst studying diffusion imaging it became apparent that neither diffusivity distributions nor diffusion anisotropy were taken into account in the current cutting-edge models. Using distributions of diffusivity values rather than a single ADC means that at a cost of fitting one extra parameter an insight into just how heterogeneous the tissue is can be gained. Diffusion anisotropy

is not in use for the simple reason that orientational heterogneity in the prostate makes this difficult but, using novel techniques taken from the modelling of axons in the brain, would also reveal more information about the underlying tissue changes between normal prostate microstructure and tumour. However, as well as opportunities in diffusion MRI, the literature also showed the possibility of using quantitative T2 modelling to detect and grade tumours. These methods did exist in modelling myelin in the brain but were unused in the prostate until Sabouri et al. published in 2017 [136].

These opportunities perfectly fit the aim of the thesis in that they all sought to improve our quantification of prostate microstructure through a combination of advanced acquisitions and better signal modelling without fitting a large number of parameters.

As discussed in the introduction, the 'Imaging Biomarker Roadmap for Cancer Studies' [115] provides a framework for developing imaging biomarkers. As shown in Appendix A, in order to bridge the first translational gap for any imaging technique there first needs to be suitable data. Chapter 4 set out to decide upon the data that was to be acquired. By performing experiments on simulated data, gel phantoms and in-vivo this work concluded that the LWI fitting should be carried out using a constrained two-Gaussian fit, that the acquisitino should be reduced to 32-echoes and that the echo spacing should be fixed at 31.25ms. After coming to these conclusions a dataset of 65 patients was then acquired using these findings.

Once a suitable dataset exists, the roadmap then suggests that the imaging technique must undergo both technical validation and biological and clinical validation. After doing initial technical tests into the hardware and software considerations, ethical approval and tolerability, the LWI acquisition was considered clear to go ahead in patients. Biological and clinical validation was the route that was taken at this point due to the fact that clinical validation would be intriguing enough to motivate future research into repeatibility and reproducability studies but not necessarily vice versa.

The end of Chapter 4 went on to correlate the LWI model to the VERDICT diffusion model, successfully linking some microstructural compartments, and then found in a small dataset that the LWF parameter of the LWI model is significantly different between tumour and normal tissue. These small experiments were designed to give a first estimate

as to the strength of LWI. Chapter 5 was the main bulk of the clinical validation. This work looked into the sensitivity and specificity of LWI in both differentiating tumour from normal tissue and then discerning between different groupings of PI-RADS scores, showing that for both of these tasks the LWF parameter performs as well or better than ADC without the limitations of an EPI-based diffision acquisition. This chapter then goes on to show through correlations with the VERDICT diffusion model that LWI is sensitive to tissue microstructure.

After carrying out a clinical validation study on LWI, the next objective on the roadmap is to look into technical validation. The first step towards technical validation is a repeatibility study. To this end, part of Chapter 5 is devoted to a study in which a subset of 16 of the subjects were used in a study into the repeatibility of the LWF parameter. Analysis showed that there was a strong correlation between the median LWF values of ROIs produced by two separate readers.

Using the 65 subject dataset, in Chapter 6 the acquisition was then further investigated. One study showed that the number of echoes could be reduced to as low as 6 echoes whilst barely reducing the ability of the LWF to detect prostate cancer whilst another showed that correcting for imperfect refocusing can improve the efficacy of LWI.

Chapter 7 then looked into machine learning techniques for discriminating between PI-RADS v2 score groupings, showing that neural networks used on either the 32-echo signal intensities or the LWI parameters are more effective than ADC.

The main impact of this thesis has been to take LWI from an unknown technique to one that has been both clinically and technically validated, providing an optimised 32-echo sequence and a robust fitting method.

8.2 Contributions

This thesis made a number of key contributions:

- A constrained LWI fitting that is faster to perform and works on small datasets of as little as six echoes. This method works as well or better than diffusion imaging

for all classifications performed, correlates with tissue microstructure, is repeatible in the dataset used here and, with the use of imperfect refocusing correction, could work across multiple centers.

- An ME-T2 sequence that is optimised for the LWI method. This acquisition could feasibly be as short as 6 echoes, meaning that the resolution could be much like that of an anatomical T2 image within a reasonable time frame. This drastic reduction in the number of echoes needed to accurately fit the LWI makes the technique substantially more likely to be clinically accepted.

- A machine learning technique that, when used on ME-T2 data, is significantly more sensitive and specific to tumours than ADC and removes the need to fit tissue models at all.

- Preliminary research into two intriguing diffusion modelling techniques that give novel information into the microstructure of the prostate and, even without optimised datasets, can match ADC for tumour detection.

8.3 Future Work

As set out in the Chapter 1, in order to take LWI from a new imaging biomarker to a validated clinical technique (as defined by the Imaging Biomarker Roadmap for Cancer Studies in Appendix A.1) it first has to have its technical and biological performance quantified in a small number of centres. This would involve carrying out similar studies to that in Chapter 5 across a number of centres and assessing how the results vary across centres. This will give a good estimation of how well the technique would perform when being used across numerous sites in the future.

In order to improve the LWI technique itself the first thing that should be further looked into is the balance of SAR, resolution, coverage, SNR and the number of echoes acquired. For clinical use the scan needs to be strong in all of these areas so a further optimised sequence must be tried. In particular, the current LWI imaging protocol does not cover the entire prostate at the current imaging resolution so it is important to increase the

coverage to the whole prostate so that when this technique is investigated in multiple centres there is no need to have prior knowledge of the location of a tumour or suspicious region. Practically, this would mean keeping the same slice thickness but increasing the coverage in the axial direction to approximately 70mm whilst keeping the scan to under around six minutes so the it is both tolerable and within SAR reccomendations. Work has already being carried out in this space, as shown by the work of Gong et al. in Section 6.1 (on which I am an author), showing that a reduction in echo number reduces SAR and imaging time but does slightly reduce the accuracy of the LWI model.

In the longer term, as the sequence is always SAR limited, the acquisition could benefit from the reduction of the flip angles from 180°. This would reduce the restrictions that SAR places on the sequence, opening up the possibility of further improvements to the rest of the sequence. These imperfect pulses could then be corrected for by using post-processing methods such as the Extended Phase Graph. Compressed sensing and multi-band excitation could also be further investigated.

Once the first translational gap has been bridged, further research should then look into multi-centre studies looking into the sensitivity, specificity and reproducibility of LWI, its effect on diagnostic accuracy and its relationship to patient outcome. Much like the mp-MRI study conducted by Ahmed et al. [6], this study would be a prospective, multi-centre study containing hundreds of subjects.

The further investigation of the machine learning techniques of Chapter 7 should have three aims. One would be to increase the size of the dataset, as machine learning techniques often struggle to work on the size of the datasets that have been used in this work. Alongside this, more sophisticated machine learning techniqes could be used to better differentiate between grades of tumour. This would likely initially take the form of larger neural networks as the size of the dataset increases but this could then develop onto using convolutional neural networks to indirectly include spatial information, such as the size and shape of the tumour, when predicting tumour grade. The third area in which further investigation should be done involves utilising other imaging modalities, such as DCE-MRI, to see whether they can improve the diagnostic accuracy of the machine learning algorithms.

Appendix A

Imaging Biomarker Roadmap

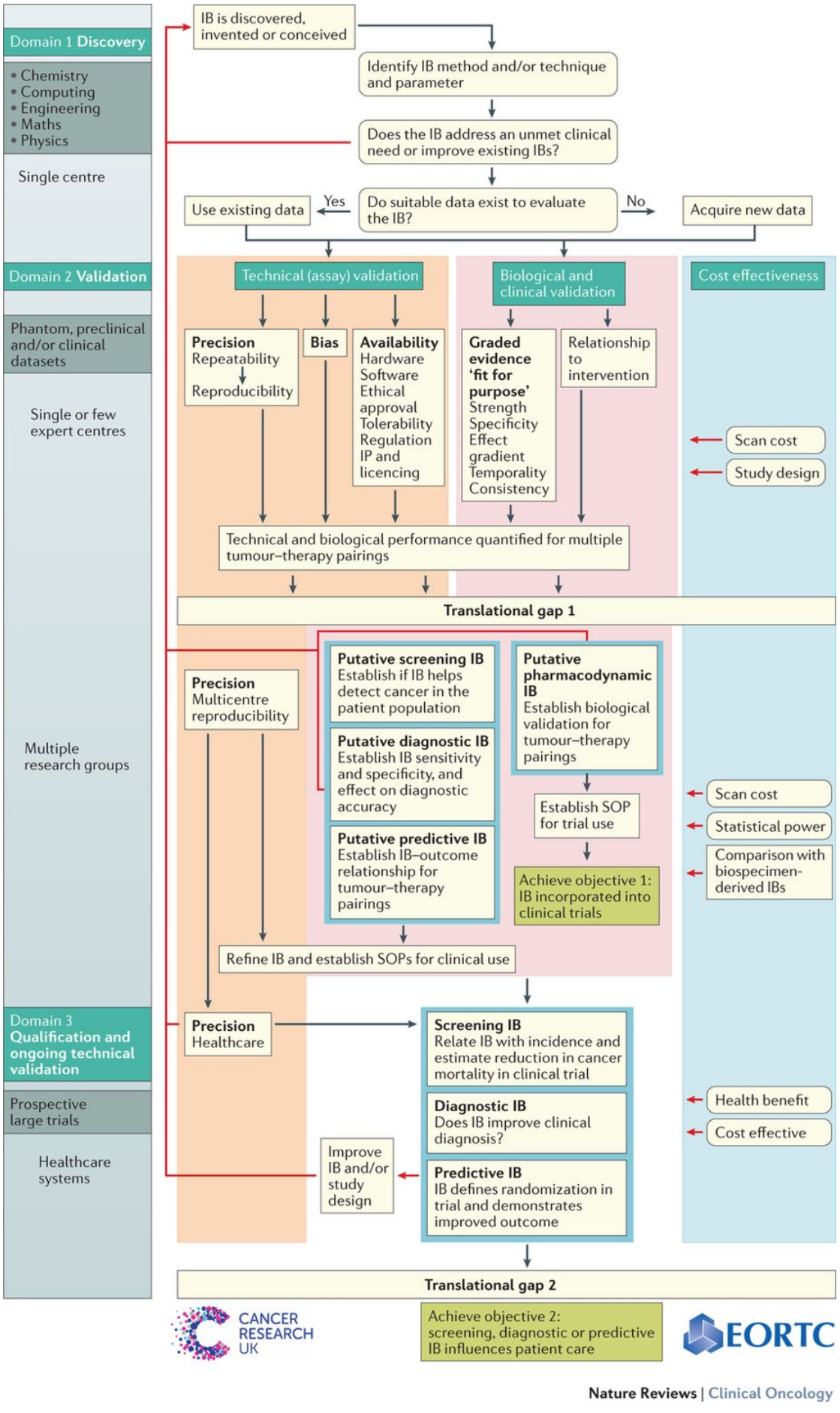


FIGURE A.1: The imaging biomarker roadmap for cancer studies (reproduced from O'Connor et al. [115] under a creative commons license.)

Bibliography

[1] Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *The Journal of Chemical Physics*, 42(1):288, jul 1965.

[2] Deaths Registered in England and Wales - Office for National Statistics, 2015.

[3] The 2014 international society of urological pathology (isup) consensus conference on gleason grading of prostatic carcinoma. *The American Journal of Surgical Pathology*, 40(2), 2016.

[4] M. Abd-Alazeez, H. U. Ahmed, M. Arya, S. C. Charman, E. Anastasiadis, A. Freeman, M. Emberton, and A. Kirkham. The accuracy of multiparametric MRI in men with negative biopsy and elevated PSA level-Can it rule out clinically significant prostate cancer? *Urologic Oncology: Seminars and Original Investigations*, 32(1):4517–4522, 2014.

[5] M. Adhyam and A. K. Gupta. A Review on the Clinical Utility of PSA in Prostate Cancer. *Indian journal of surgical oncology*, 3(2):120–9, jun 2012.

[6] H. U. Ahmed, A. El-Shater Bosaily, L. C. Brown, R. Gabe, R. Kaplan, M. K. Parmar, Y. Collaco-Moraes, K. Ward, R. G. Hindley, A. Freeman, A. P. Kirkham, R. Oldroyd, C. Parker, and M. Emberton. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *The Lancet*, 389:815–822, 2017.

[7] E. Alonso-Ortiz, I. R. Levesque, and G. B. Pike. MRI-Based Myelin Water Imaging: A Technical Review. *Magnetic Resonance in Medicine*, 73:70–81, 2015.

[8] T. Andrews, J. L. Lancaster, S. J. Dodd, C. Contreras-Sesvold, and P. T. Fox. Testing the three-pool white matter model adapted for use with T2 relaxometry. *Magnetic Resonance in Medicine*, 54(2):449–454, aug 2005.

[9] J.-P. Armspach, D. Gounot, L. Rumbach, and J. Chambron. In vivo determination of multiexponential T2 relaxation in the brain of patients with multiple sclerosis. *Magnetic Resonance Imaging*, 9:107–113, 1991.

[10] S. Barbieri, M. Brönnimann, S. Boxler, P. Vermathen, and H. C. Thoeny. Differentiation of prostate cancer lesions with high and with low Gleason score by diffusion-weighted MRI. *European Radiology*, 27(4):1547–1555, jun 2016.

[11] J. O. Barentsz, J. Richenberg, R. Clements, P. Choyke, S. Verma, G. Villeirs, O. Rouviere, V. Logager, and J. J. Fütterer. ESUR prostate MR guidelines 2012. *European Radiology*, 22(4):746–757, Apr 2012.

[12] K. M. Bennett, J. S. Hyde, S. D. Rand, R. Bennett, H. G. J. Krouwer, K. J. Rebro, and K. M. Schmainda. Intravoxel distribution of DWI decay rates reveals C6 glioma invasion in rat brain. *Magnetic Resonance in Medicine*, 52(5):994–1004, nov 2004.

[13] K. M. Bennett, K. M. Schmainda, R. T. Bennett, D. B. Rowe, H. Lu, and J. S. Hyde. Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magnetic resonance in medicine*, 50(4):727–34, oct 2003.

[14] L. K. Bittencourt, J. O. Barentsz, L. C. D. de Miranda, and E. L. Gasparetto. Prostate MRI: diffusion-weighted imaging at 1.5T correlates better with prostatectomy Gleason Grades than TRUS-guided biopsies in peripheral zone tumours. *European radiology*, 22(2):468–75, feb 2012.

[15] D. Bonekamp, S. Kohl, M. Wiesenfarth, P. Schelb, J. P. Radtke, M. Götz, P. Kickingereder, K. Yaqubi, B. Hitthaler, N. Gählert, T. A. Kuder, F. Deister, M. Freitag, M. Hohenfellner, B. A. Hadaschik, H.-P. Schlemmer, and K. H. Maier-Hein. Radiomic machine learning for characterization of prostate lesions with mri: Comparison to adc values. *Radiology*, 289(1), 2018. PMID: 30063191.

[16] E. Bonet-Carne, A. Daducci, E. Panagiotaki, E. Johnston, N. Stevens, D. Atkinson, S. Shonit Punwani, and D. Alexander. Non-invasive quantification of prostate cancer using AMICO framework for VERDICT MR. *International Society for Magnetic Resonance in Medicine (ISMRM)*, 2016.

[17] E. Bonet-Carne, M. Tariq, H. Pye, M. Appayya, A. Haider, C. Bayley, J. Jacobs, A. Freeman, D. Hawkes, G. Shaw, W. Hayley, D. Alexander, S. Punwani, and E. Panagiotaki. Histological Validation of in-vivo VERDICT MRI for Prostate using 3D Personalised Moulds. Joint Annual Meeting ISMRM-ESMRMB 2018, 2018.

[18] R. M. Bourne, A. Bongers, A. Chatterjee, P. Sved, and G. Watson. Diffusion anisotropy in fresh and fixed prostate tissue ex vivo. *Magnetic resonance in medicine*, oct 2015.

[19] R. M. Bourne, N. Kurniawan, G. Cowin, T. Stait-Gardner, P. Sved, G. Watson, S. Chowdhury, and W. S. Price. Biexponential diffusion decay in formalin-fixed prostate tissue: Preliminary findings. *Magnetic Resonance in Medicine*, 68(3):954–959, sep 2012.

[20] R. M. Bourne, N. Kurniawan, G. Cowin, T. Stait-Gardner, P. Sved, G. Watson, and W. S. Price. Microscopic diffusivity compartmentation in formalin-fixed prostate tissue. *Magnetic Resonance in Medicine*, 68(2):614–620, aug 2012.

[21] R. M. Bourne, N. Kurniawan, G. Cowin, P. Sved, and G. Watson. 16 T Diffusion microimaging of fixed prostate tissue: Preliminary findings. *Magnetic Resonance in Medicine*, 66(1):244–247, jul 2011.

[22] R. M. Bourne, N. Kurniawan, G. Cowin, P. Sved, and G. Watson. Microscopic diffusion anisotropy in formalin fixed prostate tissue: Preliminary findings. *Magnetic Resonance in Medicine*, 68(6):1943–1948, dec 2012.

[23] R. M. Bourne, E. Panagiotaki, A. Bongers, P. Sved, G. Watson, and D. C. Alexander. Information theoretic ranking of four models of diffusion attenuation in fresh and fixed prostate tissue ex vivo. *Magnetic Resonance in Medicine*, 72(5):1418–1426, nov 2014.

[24] N. Branger, T. Maubon, M. Traumann, J. Thomassin-Piana, N. Brandone, S. Taix, J. Touzlian, S. Brunelle, G. Pignot, N. Salem, G. Gravis, and J. Walz. Is negative multiparametric magnetic resonance imaging really able to exclude significant prostate cancer? the real-life experience. *BJU International*, 119(3):449–455.

[25] F. Bray, J. Ferlay, I. Soerjomataram, R. Siegel, L. Torre, and A. Jemal. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin, in press.*, 8(6), 2018.

[26] M. Brizmohun, E. Johnston, A. Latifoltojar, J. O'Callaghan, E. Bonet-Carne, U. Ferizi, B. Yvernault, H. Pye, D. Patel, J. Clemente, et al. The intracellular component of verdict (vascular, extracellular, and restricted diffusion for cytometry in tumors) mri distinguishes gleason 4 pattern better than apparent diffusion coefficient. Joint Annual Meeting ISMRM-ESMRMB 2018, 2018.

[27] M. Brizmohun Appayya, H. S. Sidhu, N. Dikaios, E. W. Johnston, L. A. Simmons, A. Freeman, A. P. Kirkham, H. U. Ahmed, and S. Punwani. Characterizing indeterminate (Likert-score 3/5) peripheral zone prostate lesions with PSA density, PI-RADS scoring and qualitative descriptors on multiparametric MRI. *The British Journal of Radiology*, 91(1083):20170645, 2018. PMID: 29189042.

[28] Canadian Cancer Society. The prostate. [Online; accessed November 16, 2018].

[29] H. Carr. Steady-State Free Precession in Nuclear Magnetic Resonance. *Physical Review*, 112(5):1693–1701, 1958.

[30] R. W. Chan, A. Z. Lau, G. Detzler, V. Thayalasuthan, R. K. Nam, and M. A. Haider. Evaluating the accuracy of multicomponent t2 parameters for luminal water imaging of the prostate with acceleration using inner-volume 3d grase. *Magnetic Resonance in Medicine*, 81(1):466–476.

[31] R. Chang, R. Kirby, and B. Challacombe. Is there a link between bph and prostate cancer? *The Practitioner*, 256(1750):13–17, 2012.

[32] A. Chatterjee, G. Watson, E. Myint, P. Sved, M. McEntee, and R. M. Bourne. Changes in Epithelium, Stroma, and Lumen Space Correlate More Strongly with Gleason Pattern and Are Stronger Predictors of Prostate ADC Changes than Cellularity Metrics. *Radiology*, 277(3):751–762, dec 2015.

[33] A. P. Chen, D. Xu, R. Henry, A. Qayyum, J. Kurhanewicz, and D. B. Vigneron. Diffusion Tensor Imaging of the Prostate Following Therapy. In *ISMRM*, San Francisco, 2003.

[34] M. M. Cheung, E. S. Hui, K. C. Chan, J. A. Helpern, L. Qi, and E. X. Wu. Does diffusion kurtosis imaging lead to better neural tissue characterization? A rodent brain maturation study. *NeuroImage*, 45(2):386–392, 2009.

[35] C.-L. Chin, F. W. Wehrli, S. N. Hwang, M. Takahashi, and D. B. Hackney. Biexponential diffusion attenuation in the rat spinal cord: Computer simulations based on anatomic images of axonal architecture. *Magnetic Resonance in Medicine*, 47(3):455–460, mar 2002.

[36] E. Chiou, F. Giganti, E. Bonet-Carne, S. Punwani, I. Kokkinos, and E. Panagiotaki. Prostate cancer classification on verdict dw-mri using convolutional neural networks. In Y. Shi, H.-I. Suk, and M. Liu, editors, *Machine Learning in Medical Imaging*, pages 319–327, Cham, 2018. Springer International Publishing.

[37] K. A. Christensen, D. M. Grant, E. M. Schulman, and C. Walling. Optimal determination of relaxation times of fourier transform nuclear magnetic resonance. Determination of spin-lattice relaxation times in chemically polarized species. *The Journal of Physical Chemistry*, 78(19):1971–1977, sep 1974.

[38] E. Cornel, G. Smits, G. Oosterhof, H. Karthaus, F. Deburyne, J. Schalken, and A. Heerschap. Characterization of human prostate cancer, benign prostatic hyperplasia and normal prostate by in vitro 1h and 31p magnetic resonance spectroscopy. *The Journal of urology*, 150(6):2019—2024, December 1993.

[39] S. De Santis, A. Gabrielli, M. Bozzali, B. Maraviglia, E. Macaluso, and S. Capuani. Anisotropic anomalous diffusion assessed in the human brain by scalar invariant indices. *Magnetic Resonance in Medicine*, 65(4):1043–1052, apr 2011.

[40] N. B. Delongchamps, M. Rouanne, T. Flam, F. Beuvon, M. Liberatore, M. Zerbib, and F. Cornud. Multiparametric magnetic resonance imaging for the detection and localization of prostate cancer: combination of T2-weighted, dynamic contrast-enhanced and diffusion-weighted imaging. *BJU International*, 107(9):1411–1418, may 2011.

[41] S. C. Deoni, B. K. Rutt, T. Arun, C. Pierpaoli, and D. K. Jones. Gleaning multicomponent T1 and T2 information from steady-state imaging data. *Magnetic Resonance in Medicine*, 60(6):1372–1387, dec 2008.

[42] S. C. Deoni, B. K. Rutt, and T. M. Peters. Rapid combinedT1 andT2 mapping using gradient recalled acquisition in the steady state. *Magnetic Resonance in Medicine*, 49(3):515–526, mar 2003.

[43] S. C. L. Deoni, T. M. Peters, and B. K. Rutt. High-resolutionT1 andT2 mapping of the brain in a clinically acceptable time with DESPOT1 and DESPOT2. *Magnetic Resonance in Medicine*, 53(1):237–241, jan 2005.

[44] M. D. Does and J. C. Gore. Compartmental study ofT1 andT2 in rat brain and trigeminal nerve in vivo. *Magnetic Resonance in Medicine*, 47(2):274–283, feb 2002.

[45] M. R. Engelbrecht, H. J. Huisman, R. J. F. Laheij, G. J. Jager, G. J. L. H. van Leenders, C. A. Hulsbergen-Van De Kaa, J. J. M. C. H. de la Rosette, J. G. Blickman, and J. O. Barentsz. Discrimination of Prostate Cancer from Normal Peripheral Zone and Central Gland Tissue by Using Dynamic Contrast-enhanced MR Imaging. *Radiology*, 229(1):248–254, oct 2003.

[46] M. R. Engelbrecht, P. Puech, P. Colin, O. Akin, L. Lemaître, and A. Villers. Multimodality Magnetic Resonance Imaging of Prostate Cancer. *Journal of Endourology*, 24(5):677–684, may 2010.

[47] J. I. Epstein, W. C. Allsbrook Jr, M. B. Amin, L. L. Egevad, I. G. Committee, et al. The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. *The American journal of surgical pathology*, 29(9):1228–1242, 2005.

[48] M. F. Falangola, J. H. Jensen, J. S. Babb, C. Hu, F. X. Castellanos, A. Di Martino, S. H. Ferris, and J. A. Helpern. Age-related non-Gaussian diffusion patterns in the prefrontal brain. *Journal of Magnetic Resonance Imaging*, 28(6):1345–1350, dec 2008.

[49] D. Fehr, H. Veeraraghavan, A. Wibmer, T. Gondo, K. Matsumoto, H. A. Vargas, E. Sala, H. Hricak, and J. O. Deasy. Automatic classification of prostate cancer gleason scores from multiparametric magnetic resonance images. *Proceedings of the National Academy of Sciences*, 112(46):E6265–E6273, 2015.

[50] A. F. Frøhlich, L. Østergaard, and V. G. Kiselev. Effect of impermeable boundaries on diffusion-attenuated MR signal. *Journal of Magnetic Resonance*, 179(2):223–233, 2006.

[51] K. Gersonde, T. Tolxdorff, and L. Felsberg. Identification and characterization of tissues by t2-selective whole-body proton nmr imaging. *Magnetic Resonance in Medicine*, 2(4):390–401.

[52] P. Gibbs, G. P. Liney, M. D. Pickles, B. Zelhof, G. Rodrigues, and L. W. Turnbull. Correlation of ADC and T2 Measurements With Cell Density in Prostate Cancer at 3.0 Tesla. *Investigative Radiology*, 44(9):572–576, sep 2009.

[53] P. Gibbs, M. D. Pickles, and L. W. Turnbull. Diffusion imaging of the prostate at 3.0 tesla. *Invest Radiol*, 41(2):185–188, 2006.

[54] P. Gibbs, M. D. Pickles, and L. W. Turnbull. Diffusion imaging of the prostate at 3.0 tesla. *Investigative radiology*, 41(2):185–188, 2006.

[55] P. Gibbs, D. J. Tozer, G. P. Liney, and L. W. Turnbull. Comparison of quantitativeT2 mapping and diffusion-weighted imaging in the normal and pathologic prostate. *Magnetic Resonance in Medicine*, 46(6):1054–1058, dec 2001.

[56] N. Gilani, P. Malcolm, and G. Johnson. A model describing diffusion in prostate cancer. *Magnetic Resonance in Medicine*, jul 2016.

[57] N. Gilani, A. B. Rosenkrantz, P. Malcolm, and G. Johnson. Minimization of errors in biexponential T 2 measurements of the prostate. *Journal of Magnetic Resonance Imaging*, 42(4):1072–1077, oct 2015.

[58] G. F. Giskeødegård, H. Bertilsson, K. M. Selnæs, A. J. Wright, T. F. Bathen, T. Viset, J. Halgunset, A. Angelsen, I. S. Gribbestad, and M.-B. Tessem. Spermine and Citrate as Metabolic Biomarkers for Assessing Prostate Cancer Aggressiveness. *PLoS ONE*, 8(4):e62375, apr 2013.

[59] D. F. Gleason. Classification of prostatic carcinomas. *Cancer chemotherapy reports*, 50(3):125–128, 1966.

[60] X. Glorot, A. Bordes, and Y. Bengio. Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 315–323, 2011.

[61] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[62] S. J. Graham, P. L. Stanchev, and M. J. Bronskill. Criteria for analysis of multicomponent tissue T2 relaxation data. *Magnetic Resonance in Medicine*, 35(3):370–378, mar 1996.

[63] A. Graser, A. Heuck, B. Sommer, J. Massmann, J. Scheidler, M. Reiser, and U. Mueller-Lisse. Per-Sextant Localization and Staging of Prostate Cancer: Correlation of Imaging Findings with Whole-Mount Step Section Histopathology. *American Journal of Roentgenology*, 188(1):84–90, jan 2007.

[64] M. D. Greer, A. M. Brown, J. H. Shih, R. M. Summers, J. Marko, Y. M. Law, S. Sankineni, A. K. George, M. J. Merino, P. A. Pinto, P. L. Choyke, and B. Turkbey. Accuracy and agreement of piradsv2 for prostate cancer mpmri: A multireader study. *Journal of Magnetic Resonance Imaging*, 45(2):579–585.

[65] M. A. Haider, T. H. van der Kwast, J. Tanguay, A. J. Evans, A.-T. Hashmi, G. Lockwood, and J. Trachtenberg. Combined T2-Weighted and Diffusion-Weighted MRI for Localization of Prostate Cancer. *American Journal of Roentgenology*, 189(2):323–328, aug 2007.

[66] M. G. Hall and T. R. Barrick. From diffusion-weighted MRI to anomalous diffusion imaging. *Magnetic Resonance in Medicine*, 59(3):447–455, mar 2008.

[67] T. Hambrock, D. M. Somford, H. J. Huisman, I. M. van Oort, J. A. Witjes, C. A. Hulsbergen-van de Kaa, T. Scheenen, and J. O. Barentsz. Relationship between Apparent Diffusion Coefficients at 3.0-T MR Imaging and Gleason Grade in Peripheral Zone Prostate Cancer. *Radiology*, 259(2):453–461, may 2011.

[68] C. J. Harvey, J. Pilcher, J. Richenberg, U. Patel, and F. Frauscher. Applications of transrectal ultrasound in prostate cancer. *The British Journal of Radiology*, 85(special_issue_1):S3–S17, 2012. PMID: 22844031.

[69] I. Heidegger, V. Skradski, E. Steiner, H. Klocker, R. Pichler, A. Pircher, W. Horninger, and J. Bektic. High risk of under-grading and -staging in prostate cancer patients eligible for active surveillance. *PLOS ONE*, 10(2):1–12, 02 2015.

[70] J. Hennig. Multiecho imaging sequences with low refocusing flip angles. *Journal of Magnetic Resonance (1969)*, 78(3):397 – 407, 1988.

[71] P. A. Humphrey. Gleason grading and prognostic factors in carcinoma of the prostate. *Modern Pathology*, 17:292, feb 2004.

[72] S. Isebaert, L. Van den Bergh, K. Haustermans, S. Joniau, E. Lerut, L. De Wever, F. De Keyzer, T. Budiharto, P. Slagmolen, H. Van Poppel, and R. Oyen. Multiparametric MRI for prostate cancer localization in correlation to whole-mount histopathology. *Journal of Magnetic Resonance Imaging*, 37(6):1392–1401, jun 2013.

[73] Y. Itou, K. Nakanishi, Y. Narumi, Y. Nishizawa, and H. Tsukuma. Clinical utility of apparent diffusion coefficient (ADC) values in patients with prostate cancer: can ADC values contribute to assess the aggressiveness of prostate cancer? *Journal of magnetic resonance imaging : JMRI*, 33(1):167–72, jan 2011.

[74] I. Jambor, H. Merisaari, P. Taimen, P. J. Boström, H. Minn, M. Pesola, and H. J. Aronen. Evaluation of different mathematical models for diffusion-weighted imaging of normal prostate and prostate cancer using high b-values: A repeatability study. *Magnetic Resonance in Medicine*, 73(5):1988–1998, may 2015.

[75] J. H. Jensen and J. Helpern. Quantifying Non-Gaussian Water Diffusion by Means of Pulsed-Field-Gradient MRI. *Magnetic Resonance in Medicine*, 11, 2003.

[76] J. H. Jensen and J. A. Helpern. MRI quantification of non-Gaussian water diffusion by kurtosis analysis. *NMR in biomedicine*, 23(7):698–710, aug 2010.

[77] J. H. Jensen, J. A. Helpern, A. Ramani, H. Lu, and K. Kaczynski. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic resonance in medicine*, 53(6):1432–40, jun 2005.

[78] E. W. Johnston, H. Pye, E. Bonet-Carne, E. Panagiotaki, D. Patel, M. Galazi, S. Heavey, L. Carmona, A. Freeman, G. Trevisan, C. Allen, A. Kirkham, K. Burling, N. Stevens, D. Hawkes, M. Emberton, C. Moore, H. U. Ahmed, D. Atkinson, M. Rodriguez-Justo, T. Ng, D. Alexander, H. Whitaker, and S. Punwani. INNOVATE: A prospective cohort study combining serum and urinary biomarkers with novel diffusion-weighted magnetic resonance imaging for the prediction and characterization of prostate cancer.

[79] R. Jyoti, T. P. Jain, H. Haxhimolla, H. Liddell, and S. E. Barrett. Correlation of apparent diffusion coefficient ratio on 3.0â¯t mri with prostate cancer gleason score. *European Journal of Radiology Open*, 5:58 – 63, 2018.

[80] E. Kaden, T. R. Knösche, and A. Anwander. Parametric spherical deconvolution: inferring anatomical connectivity using diffusion MR imaging. *NeuroImage*, 37(2):474–88, aug 2007.

[81] E. Kaden, F. Kruggel, and D. C. Alexander. Quantitative mapping of the per-axon diffusion coefficients in brain white matter. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, pages n/a–n/a, may 2015.

[82] L. Kjaer et al. In vivo estimation of relaxation processes in benign hyperplasia and carcinoma of the prostate gland by magnetic resonance imaging. *Magnetic resonance imaging*, 5(1):23–30, 1987.

[83] P. Kozlowski, S. D. Chang, E. C. Jones, K. W. Berean, H. Chen, and S. L. Goldenberg. Combined diffusion-weighted and dynamic contrast-enhanced MRI for prostate cancer diagnosis—Correlation with biopsy and histopathology. *Journal of Magnetic Resonance Imaging*, 24(1):108–113, jul 2006.

[84] P. Kozlowski, S. D. Chang, R. Meng, B. Mädler, R. Bell, E. C. Jones, and S. L. Goldenberg. Combined prostate diffusion tensor imaging and dynamic contrast enhanced MRI at 3T–quantitative correlation with biopsy. *Magnetic resonance imaging*, 28(5):621–8, jun 2010.

[85] D. L. Langer, T. H. van der Kwast, A. J. Evans, A. Plotkin, J. Trachtenberg, B. C. Wilson, and M. A. Haider. Prostate Tissue Composition and MR Measurements: Investigating the Relationships between ADC, T2, Ktrans, ve, and Corresponding Histologic Features1. *http://dx.doi.org/10.1148/radiol.10091343*, 2010.

[86] D. L. Langer, T. H. van der Kwast, A. J. Evans, L. Sun, M. J. Yaffe, J. Trachtenberg, and M. A. Haider. Intermixed Normal Tissue within Prostate Cancer: Effect on MR Imaging Measurements of Apparent Diffusion Coefficient and T2—Sparse versus Dense Cancers. *Radiology*, 249(3):900–908, dec 2008.

[87] A. Latifoltojar, N. Dikaios, A. Ridout, C. Moore, R. Illing, A. Kirkham, S. Taylor, S. Halligan, D. Atkinson, C. Allen, M. Emberton, and S. Punwani. Evolution of multi-parametric MRI quantitative parameters following transrectal ultrasound-guided biopsy of the prostate. *Prostate Cancer And Prostatic Diseases*, 18:343, jul 2015.

[88] C. Laule, E. Leung, D. K. B. Li, A. L. Traboulsee, D. W. Paty, A. L. MacKay, and G. R. W. Moore. Myelin water imaging in multiple sclerosis: quantitative correlations with histopathology. *Multiple Sclerosis*, 12(6):747–753, 2006.

[89] C. Laule, I. M. Vavasour, S. H. Kolind, D. K. B. Li, T. L. Traboulsee, G. R. W. Moore, and A. L. MacKay. Magnetic Resonance Imaging of Myelin. *Neurotherapeutics*, 4:460–484, 2007.

[90] E. M. Lawrence et al. Evaluating Prostate Cancer Using Fractional Tissue Composition of Radical Prostatectomy Specimens and Pre-Operative Diffusional Kurtosis Magnetic Resonance Imaging. *PLOS ONE*, 11(7):e0159652, jul 2016.

[91] M. Lazar, J. H. Jensen, L. Xuan, and J. A. Helpern. Estimation of the orientation distribution function from diffusional kurtosis imaging. *Magnetic Resonance in Medicine*, 60(4):774–781, 2008.

[92] Y. LeCun, D. Touresky, G. Hinton, and T. Sejnowski. A theoretical framework for back-propagation. In *Proceedings of the 1988 connectionist models summer school*, volume 1, pages 21–28. CMU, Pittsburgh, Pa: Morgan Kaufmann, 1988.

[93] C.-Y. Lee, K. M. Bennett, and J. P. Debbins. Sensitivities of statistical distribution model and diffusion kurtosis model in varying microstructural environments: A Monte Carlo study. *Journal of Magnetic Resonance*, 230:19–26, 2013.

[94] F. Lee, P. J. Littrup, S. T. Torp-Pedersen, C. Mettlin, T. A. McHugh, J. M. Gray, G. H. Kumasaka, and R. D. McLeary. Prostate cancer: comparison of transrectal us and digital rectal examination for screening. *Radiology*, 168(2):389–394, 1988. PMID: 3293108.

[95] H. K. Lim, J. K. Kim, K. A. Kim, and K.-S. Cho. Prostate Cancer: Apparent Diffusion Coefficient Map with T2-weighted Images for Detection—A Multireader Study. *Radiology*, 250(1):145–151, jan 2009.

[96] G. P. Liney, M. Lowry, L. W. Turnbull, D. J. Manton, A. J. Knowles, S. J. Blackband, and A. Horsman. Proton mr t2 maps correlate with the citrate concentration in the prostate. *NMR in Biomedicine*, 9(2):59–64, 1996.

[97] G. P. Liney, L. W. Turnbull, M. Lowry, L. S. Turnbull, A. J. Knowles, and A. Horsman. In vivo quantification of citrate concentration and water T2 relaxation time of the pathologic prostate gland using 1H MRS and MRI. *Magnetic Resonance Imaging*, 15(10):1177–1186, 1997.

[98] C. Liu, R. Bammer, B. Acar, and M. E. Moseley. Characterizing non-gaussian diffusion by using generalized diffusion tensors. *Magnetic Resonance in Medicine*, 51(5):924–937, may 2004.

[99] H. Lu, J. H. Jensen, A. Ramani, and J. A. Helpern. Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging. *NMR in Biomedicine*, 19(2):236–247, apr 2006.

[100] A. L. MacKay, C. Laule, I. M. Vavasour, T. A. Bjarnason, S. H. Kolind, and B. Mädler. Insights into brain microstructure from the T2 distribution. *Magnetic Resonance Imaging*, 24(4):515–525, 2006.

[101] A. L. MacKay, K. Whittall, J. Adler, D. Li, D. Paty, and D. Graeb. In vivo visualization of myelin water in brain by magnetic resonance. *Magnetic Resonance in Medicine*, 31(6):673–677, jun 1994.

[102] A. Madabhushi and G. Lee. Image analysis and machine learning in digital pathology: Challenges and opportunities. *Medical Image Analysis*, 33:170 – 175, 2016. 20th anniversary of the Medical Image Analysis journal (MedIA).

[103] J. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1 – 36, 1998.

[104] G. Manenti, M. Carlani, S. Mancino, V. Colangelo, M. Di Roma, E. Squillaci, and G. Simonetti. Diffusion tensor Magnetic Resonance Imaging of prostate cancer. *Investigative Radiology*, 42(6):412–9, jun 2007.

[105] Y. Mazaheri, H. Hricak, S. W. Fine, O. Akin, A. Shukla-Dave, N. M. Ishill, C. S. Moskowitz, J. E. Grater, V. E. Reuter, K. L. Zakian, K. A. Touijer, and J. A. Koutcher. Prostate Tumor Volume Measurement with Combined T2-weighted Imaging and Diffusion-weighted MR: Correlation with Pathologic Tumor Volume. *Radiology*, 252(2):449–457, aug 2009.

[106] Y. Mazaheri, H. A. Vargas, G. Nyman, O. Akin, and H. Hricak. Image artifacts on prostate diffusion-weighted magnetic resonance imaging: Trade-offs at 1.5 tesla and 3.0 tesla. *Academic Radiology*, 20(8):1041 – 1047, 2013.

[107] L. N. Mazzoni, S. Lucarini, S. Chiti, S. Busoni, C. Gori, and I. Menchi. Diffusion-weighted signal models in healthy and cancerous peripheral prostate tissues: Comparison of outcomes obtained at different b-values. *Journal of Magnetic Resonance Imaging*, 39(3):512–518, mar 2014.

[108] J. E. McNeal. Normal histology of the prostate. *The American journal of surgical pathology*, 12(8):619–33, aug 1988.

[109] J. E. McNeal, E. A. Redwine, F. S. Freiha, and T. A. Stamey. Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread. *The American journal of surgical pathology*, 12(12):897–906, 1988.

[110] D. Mitchell, D. Burk, S. Vinitski, and M. Rifkin. The biophysical basis of tissue contrast in extracranial MR imaging. *American Journal of Roentgenology*, 149(4):831–837, oct 1987.

[111] G. R. W. Moore, E. Leung, A. L. MacKay, I. M. Vavasour, K. P. Whittall, K. S. Cover, D. K. Li, S. A. Hashimoto, J. Oger, T. J. Sprinkle, and D. W. Paty. A pathology-MRI study of the short-T2 component in formalin-fixed multiple sclerosis brain. *Neurology*, 55(10):1506–10, nov 2000.

[112] R. V. Mulkern, S. J. Haker, and S. E. Maier. On high b diffusion imaging in the human brain: ruminations and experimental insights. *Magnetic resonance imaging*, 27(8):1151–62, oct 2009.

[113] H. Narita, K. K. Tha, N. Hashimoto, H. Hamaguchi, S. Nakagawa, H. Shirato, and I. Kusumi. Mean kurtosis alterations of cerebral white matter in patients with schizophrenia revealed by diffusion kurtosis imaging. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 71:169–175, 2016.

[114] S. M. Noworolski, D. B. Vigneron, A. P. Chen, and J. Kurhanewicz. Dynamic contrast-enhanced MRI and MR diffusion imaging to distinguish between glandular and stromal prostatic tissues. *Magnetic resonance imaging*, 26(8):1071–80, oct 2008.

[115] J. P. B. O'Connor et al. Imaging biomarker roadmap for cancer studies. *Nature Reviews Clinical Oncology*, 14:169, oct 2016.

[116] J. Oh, E. T. Han, M. C. Lee, S. J. Nelson, and D. Pelletier. Multislice Brain Myelin Water Fractions at 3T in Multiple Sclerosis. *Journal of Neuroimaging*, 17(2):156–163, apr 2007.

[117] F. P. Oliveira and J. M. R. Tavares. Medical image registration: a review. *Computer Methods in Biomechanics and Biomedical Engineering*, 17(2):73–93, 2014. PMID: 22435355.

[118] K. Oshio, H. Shinmoto, and R. V. Mulkern. Interpretation of Diffusion MR Imaging Data using a Gamma Distribution Model. *Magnetic Resonance in Medical Sciences*, 13(3):191–195, 2014.

[119] A. R. Padhani, J. Weinreb, A. B. Rosenkrantz, G. Villeirs, B. Turkbey, and J. Barentsz. Prostate Imaging-Reporting and Data System Steering Committee: PI-RADS v2 Status Update and Future Directions. *European Urology*, oct 2018.

[120] E. Panagiotaki, R. W. Chan, N. Dikaios, H. U. Ahmed, J. O'Callaghan, A. Freeman, D. Atkinson, S. Punwani, D. J. Hawkes, and D. C. Alexander. Microstructural Characterization of Normal and Malignant Human Prostate Tissue With Vascular, Extracellular, and Restricted Diffusion for Cytometry in Tumours Magnetic Resonance Imaging. *Investigative Radiology*, 50(4):218–227, 2015.

[121] E. Panagiotaki, S. Walker-Samuel, B. Siow, S. P. Johnson, V. Rajkumar, R. B. Pedley, M. F. Lythgoe, and D. C. Alexander. Noninvasive quantification of solid tumor microstructure using VERDICT MRI. *Cancer research*, 74(7):1902–12, apr 2014.

[122] Y. Pang, B. Turkbey, M. Bernardo, J. Kruecker, S. Kadoury, M. J. Merino, B. J. Wood, P. A. Pinto, and P. L. Choyke. Intravoxel incoherent motion MR imaging for prostate cancer: An evaluation of perfusion fraction and diffusion coefficient derived from different b-value combinations. *Magnetic Resonance in Medicine*, 69(2):553–562, feb 2013.

[123] S. H. Polanec, T. H. Helbich, H. Bickel, G. J. Wengert, K. Pinker, C. Spick, P. Clauser, M. Susani, S. Shariat, and P. A. Baltzer. Quantitative Apparent Diffusion Coefficient Derived From Diffusion-Weighted Imaging Has the Potential to Avoid Unnecessary MRI-Guided Biopsies of mpMRI-Detected PI-RADS 4 and 5 Lesions. *Investigative radiology*, 53(12), 2018.

[124] T. Prasloski, B. Mädler, Q.-S. Xiang, A. L. MacKay, and C. Jones. Applications of stimulated echo correction to multicomponent T2 analysis. *Magnetic Resonance in Medicine*, 67(6):1803–1814, jun 2012.

[125] M. Quentin, G. Pentang, L. Schimmöller, O. Kott, A. Müller-Lutz, D. Blondin, C. Arsov, A. Hiester, R. Rabenalt, and H.-J. Wittsack. Feasibility of diffusional kurtosis tensor imaging in prostate MRI for the assessment of prostate cancer: Preliminary results. *Magnetic Resonance Imaging*, 32(7):880–885, 2014.

[126] A. Raj et al. Multi-Compartment T2 Relaxometry Using a Spatially Constrained Multi-Gaussian Model. *PLoS ONE*, 9(6):e98391, jun 2014.

[127] J. G. Rivas, F. Giganti, M. Alvarez-Maestro, M. J. Freire, V. Kasivisvanathan, L. Martinez-Pineiro, and M. Emberton. Prostate indeterminate lesions on magnetic resonance imaging - biopsy versus surveillance: A literature review. *European Urology Focus*, 2018.

[128] C. G. Roehrborn. Benign prostatic hyperplasia: an overview. *Reviews in urology*, 7 Suppl 9(Suppl 9):S3–S14, 2005.

[129] M. C. Roethke, T. A. Kuder, T. H. Kuru, M. Fenchel, B. A. Hadaschik, F. B. Laun, H.-P. Schlemmer, and B. Stieltjes. Evaluation of Diffusion Kurtosis Imaging Versus Standard Diffusion Imaging for Detection and Grading of Peripheral Zone Prostate Cancer. *Investigative Radiology*, 50(8):483–9, 2015.

[130] A. B. Rosenkrantz, X. Kong, B. E. Niver, D. S. Berkman, J. Melamed, J. S. Babb, and S. S. Taneja. Prostate cancer: comparison of tumor visibility on trace diffusion-weighted images and the apparent diffusion coefficient map. *AJR. American journal of roentgenology*, 196(1):123–9, jan 2011.

[131] A. B. Rosenkrantz, A. R. Padhani, T. L. Chenevert, D.-M. Koh, F. De Keyzer, B. Taouli, and D. Le Bihan. Body diffusion kurtosis imaging: Basic principles, applications, and considerations for clinical practice. *Journal of Magnetic Resonance Imaging*, 42(5):1190–1202, nov 2015.

[132] A. B. Rosenkrantz, E. E. Sigmund, G. Johnson, J. S. Babb, T. C. Mussi, J. Melamed, S. S. Taneja, V. S. Lee, and J. H. Jensen. Prostate Cancer: Feasibility and Preliminary Experience of a Diffusional Kurtosis Model for Detection and Assessment of Aggressiveness of Peripheral Zone Cancer. *Radiology*, 264(1):126–135, jul 2012.

[133] S. Sabouri, S. Chang, R. Savdie, J. Zhang, E. Jones, L. Goldenberg, and P. Kozlowski. Luminal water imaging: a novel MRI method for prostate cancer diagnosis. In *ISMRM*, page 2489, Singapore, 2016.

[134] S. Sabouri, S. D. Chang, R. Savdie, J. Zhang, E. C. Jones, S. L. Goldenberg, P. C. Black, and P. Kozlowski. Luminal water imaging: A new mr imaging t2 mapping technique for prostate cancer diagnosis. *Radiology*, 284(2):451–459, 2017. PMID: 28394754.

[135] S. Sabouri, L. Fazli, S. Chang, R. Savdie, E. Jones, L. Goldenberg, and P. Kozlowski. MR measurements of luminal water in prostate gland. In *ISMRM*, page 2487, Singapore, 2016.

[136] S. Sabouri, L. Fazli, S. D. Chang, R. Savdie, E. C. Jones, S. L. Goldenberg, P. C. Black, and P. Kozlowski. Mr measurement of luminal water in prostate gland: Quantitative correlation between mri and histology. *Journal of Magnetic Resonance Imaging*, 46(3):861–869.

[137] L. R. Schad, G. Brix, I. Zuna, W. Härle, W. J. Lorenz, and W. Semmler. Multiexponential proton spin-spin relaxation in MR imaging of human brain tumors. *Journal of computer assisted tomography*, 13(4):577–87, 1989.

[138] J. Scheidler, H. Hricak, D. B. Vigneron, K. K. Yu, D. L. Sokolov, L. R. Huang, C. J. Zaloudek, S. J. Nelson, P. R. Carroll, and J. Kurhanewicz. Prostate Cancer: Localization with Three-dimensional Proton MR Spectroscopic Imaging—Clinicopathologic Study. *Radiology*, 213(2):473–480, nov 1999.

[139] B. Scherrer, A. Schwartzman, M. Taquet, M. Sahin, S. P. Prabhu, and S. K. Warfield. Characterizing brain tissue by assessment of the distribution of anisotropic microstructural environments in diffusion-compartment imaging (DIAMOND). *Magnetic resonance in medicine*, 76(3), sep 2015.

[140] L. Schimmöller, M. Quentin, C. Arsov, R. S. Lanzman, A. Hiester, R. Rabenalt, G. Antoch, P. Albers, and D. Blondin. Inter-reader agreement of the ESUR score for prostate MRI using in-bore MRI-guided biopsies as the reference standard. *European Radiology*, 23(11):3185–3190, 2013.

[141] P. Schuenke, J. Windschuh, V. Roeloffs, M. E. Ladd, P. Bachert, and M. Zaiss. Simultaneous mapping of water shift and B1(WASABI)—Application to field-Inhomogeneity correction of CEST MRI data. *Magnetic Resonance in Medicine*, 77(2):571–580, 2017.

[142] M. Scialpi, V. Rondoni, M. C. Aisa, E. Martorana, A. D'Andrea, C. M. Malaspina, A. Orlandi, G. Galassi, E. Orlandi, P. Scialpi, M. Dragone, D. Palladino, A. Simeone, M. Amenta, and G. Bianchi. Is contrast enhancement needed for diagnostic prostate mri? *Translational Andrology and Urology*, 6(3), 2017.

[143] J. V. Sehy, J. J. Ackerman, and J. J. Neil. Evidence that both fast and slow water ADC components arise from intracellular space. *Magnetic Resonance in Medicine*, 48(5):765–770, nov 2002.

[144] K. M. Selnæs, R. Vettukattil, H. Bertilsson, A. J. Wright, A. Heerschap, A. Angelsen, M.-B. Tessem, and T. F. Bathen. Tissue Microstructure Is Linked to MRI Parameters and Metabolite Levels in Prostate Cancer. *Frontiers in oncology*, 6:146, 2016.

[145] C. B. Shaw, J. H. Jensen, R. L. Deardorff, M. V. Spampinato, and J. A. Helpern. Comparison of Diffusion Metrics Obtained at 1.5T and 3T in Human Brain With Diffusional Kurtosis Imaging. *Journal of Magnetic Resonance Imaging*, 45(3), jul 2016.

[146] H. Shinmoto, K. Oshio, C. Tamura, S. Soga, T. Okamura, K. Yamada, T. Kaji, and R. V. Mulkern. Diffusion-weighted imaging of prostate cancer using a statistical model based on the gamma distribution. *Journal of Magnetic Resonance Imaging*, 42(1):56–62, jul 2015.

[147] H. Shinmoto, K. Oshio, A. Tanimoto, N. Higuchi, S. Okuda, S. Kuribayashi, and R. V. Mulkern. Biexponential apparent diffusion coefficients in prostate cancer. *Magnetic Resonance Imaging*, 27(3):355–359, 2009.

[148] L. A. M. Simmons, A. Kanthabalan, M. Arya, T. Briggs, D. Barratt, S. C. Charman, A. Freeman, J. Gelister, D. Hawkes, Y. Hu, C. Jameson, N. McCartan,

C. M. Moore, S. Punwani, N. Ramachandran, J. van der Meulen, M. Emberton, and H. U. Ahmed. The PICTURE study: diagnostic accuracy of multiparametric MRI in men requiring a repeat prostate biopsy. *British Journal Of Cancer*, 116:1159, mar 2017.

[149] C. R. Smittenaar, K. A. Petersen, K. Stewart, and N. Moitt. Cancer incidence and mortality projections in the UK until 2035. *British Journal of Cancer*, 115(9):1147–1155, oct 2016.

[150] G. J. Stanisz and R. M. Henkelman. Diffusional Anisotropy of T2 Components in Bovine Optic Nerve. 40(3).

[151] P. Steiger and H. C. Thoeny. Prostate MRI based on PI-RADS version 2: how we review and report. *Cancer Imaging*, 16(1):9, dec 2016.

[152] T. H. Storås, K.-I. Gjesdal, Ø. B. Gadmar, J. T. Geitung, and N.-E. Kløw. Prostate magnetic resonance imaging: multiexponential T2 decay in prostate tissue. *Journal of magnetic resonance imaging : JMRI*, 28(5):1166–72, nov 2008.

[153] S. Suo, X. Chen, L. Wu, X. Zhang, Q. Yao, Y. Fan, H. Wang, and J. Xu. Non-Gaussian water diffusion kurtosis imaging of prostate cancer. *Magnetic Resonance Imaging*, 32(5):421–427, 2014.

[154] T. Tamada, T. Sone, Y. Jo, S. Toshimitsu, T. Yamashita, A. Yamamoto, D. Tanimoto, and K. Ito. Apparent diffusion coefficient values in peripheral and transition zones of the prostate: Comparison between normal and malignant prostatic tissues and correlation with histologic grade. *Journal of Magnetic Resonance Imaging*, 28(3):720–726, sep 2008.

[155] C. Tamura, H. Shinmoto, S. Soga, T. Okamura, H. Sato, T. Okuaki, Y. Pang, S. Kosuda, and T. Kaji. Diffusion kurtosis imaging study of prostate cancer: Preliminary findings. *Journal of Magnetic Resonance Imaging*, 40(3):723–729, sep 2014.

[156] J. Toivonen, H. Merisaari, M. Pesola, P. Taimen, P. J. Boström, T. Pahikkala, H. J. Aronen, and I. Jambor. Mathematical models for diffusion-weighted imaging of prostate cancer using b values up to 2000 s/mm ² : Correlation with Gleason score and repeatability of region of interest analysis. *Magnetic Resonance in Medicine*, 74(4):1116–1124, oct 2015.

[157] M. Truong, G. Hollenberg, E. Weinberg, E. M. Messing, H. Miyamoto, and T. P. Frye. Impact of gleason subtype on prostate cancer detection using multiparametric magnetic resonance imaging: correlation with final histopathology. *The Journal of urology*, 198(2):316–321, 2017.

[158] B. Turkbey, P. A. Pinto, H. Mani, M. Bernardo, Y. Pang, Y. L. McKinney, K. Khurana, G. C. Ravizzini, P. S. Albert, M. J. Merino, and P. L. Choyke. Prostate Cancer: Value of Multiparametric MR Imaging at 3 T for Detection—Histopathologifc Correlation. *Radiology*, 255(1):89–99, apr 2010.

[159] B. Turkbey, V. P. Shah, Y. Pang, M. Bernardo, S. Xu, J. Kruecker, J. Locklin, A. A. Baccala, A. R. Rastinehad, M. J. Merino, J. H. Shih, B. J. Wood, P. A. Pinto, and P. L. Choyke. Is Apparent Diffusion Coefficient Associated with Clinical Risk Scores for Prostate Cancers that Are Visible on 3-T MR Images? *Radiology*, 258(2):488–495, feb 2011.

[160] Y. Ueda, S. Takahashi, N. Ohno, K. Kyotani, H. Kawamitu, T. Miyati, N. Aoyama, Y. Ueno, K. Kitajima, F. Kawakami, T. Okuaki, R. Tsukamoto, E. Yanagita, and K. Sugimura. Triexponential function analysis of diffusion-weighted MRI for diagnosing prostate cancer. *Journal of Magnetic Resonance Imaging*, 43(1):138–148, jan 2016.

[161] C. F. Uribe, E. C. Jones, S. D. Chang, S. L. Goldenberg, S. A. Reinsberg, and P. Kozlowski. In vivo 3T and ex vivo 7T diffusion tensor imaging of prostate cancer: Correlation with histology. *Magnetic Resonance Imaging*, 33(5):577–583, 2015.

[162] M. A. van Schie, C. V. Dinh, P. J. van Houdt, F. J. Pos, S. W. Heijmink, L. G. Kerkmeijer, A. N. Kotte, R. Oyen, K. Haustermans, and U. A. van der Heide. Contouring of prostate tumors on multiparametric mri: Evaluation of clinical delineations in a multicenter radiotherapy trial. *Radiotherapy and Oncology*, 128(2), 2018.

[163] M. A. Viergever, J. A. Maintz, S. Klein, K. Murphy, M. Staring, and J. P. Pluim. A survey of medical image registration - under review. *Medical Image Analysis*, 33:140 – 144, 2016. 20th anniversary of the Medical Image Analysis journal (MedIA).

[164] J. Wang, C.-J. Wu, M.-L. Bao, J. Zhang, X.-N. Wang, and Y.-D. Zhang. Machine learning-based analysis of mr radiomics can help to improve the diagnostic performance of pi-rads v2 in clinically relevant prostate cancer. *European Radiology*, 27(10):4082–4090, Oct 2017.

[165] S. Webb, C. A. Munro, R. Midha, and G. J. Stanisz. Is multicomponentT2 a good measure of myelin content in peripheral nerve? *Magnetic Resonance in Medicine*, 49(4):638–645, apr 2003.

[166] A. Wefer, H. Hricak, D. Vigneron, F. Coakley, Y. Lu, J. Wefer, U. Mueller-Lisse, P. Carroll, and J. Kurhanewicz. Sextant localization of prostate cancer: comparison of sextant biopsy, MRI and MR Spectroscopic Imaging with step section histology. *The Journal of Urology*, 164(2):400–404, 2000.

[167] J. C. Weinreb, J. O. Barentsz, P. L. Choyke, F. Cornud, M. A. Haider, K. J. Macura, D. Margolis, M. D. Schnall, F. Shtern, C. M. Tempany, H. C. Thoeny, and S. Verma. PI-RADS Prostate Imaging â Reporting and Data System: 2015, Version 2. *European Urology*, 69(1):16–40, 2016.

[168] K. P. Whittall, A. L. MacKay, D. A. Graeb, R. A. Nugent, D. K. B. Li, and D. W. Paty. In vivo measurement ofT2 distributions and water contents in normal human brain. *Magnetic Resonance in Medicine*, 37(1):34–43, jan 1997.

[169] K. P. Whittall, A. L. MacKay, and D. K. Li. Are mono-exponential fits to a few echoes sufficient to determine T2 relaxation for in vivo human brain? *Magnetic Resonance in Medicine*, 41(6):1255–1257, jun 1999.

[170] S. Woo, C. H. Suh, S. Y. Kim, J. Y. Cho, and S. H. Kim. Diagnostic performance of prostate imaging reporting and data system version 2 for detection of prostate cancer: a systematic review and diagnostic meta-analysis. *European urology*, 72(2):177–188, 2017.

[171] E. X. Wu and M. M. Cheung. MR diffusion kurtosis imaging for neural tissue characterization. *NMR in biomedicine*, 23(7):836–48, aug 2010.

[172] J. Xu, P. A. Humphrey, A. S. Kibel, A. Z. Snyder, V. R. Narra, J. J. H. Ackerman, and S.-K. Song. Magnetic resonance diffusion characteristics of histologically defined prostate cancer in humans. *Magnetic resonance in medicine*, 61(4):842–50, apr 2009.

[173] D. A. Yablonskiy, G. L. Bretthorst, and J. J. Ackerman. Statistical model for diffusion attenuated MR signal. *Magnetic resonance in medicine*, 50(4):664–9, oct 2003.

[174] X. Yang, C. Liu, Z. Wang, J. Yang, H. L. Min, L. Wang, and K.-T. T. Cheng. Co-trained convolutional neural networks for automated detection of prostate cancer in multi-parametric MRI. *Medical image analysis*, 42:212–227, aug 2017.

[175] K. Yoshimitsu, K. Kiyoshima, H. Irie, T. Tajima, Y. Asayama, M. Hirakawa, K. Ishigami, S. Naito, and H. Honda. Usefulness of apparent diffusion coefficient map in diagnosing prostate carcinoma: Correlation with stepwise histopathology. *Journal of Magnetic Resonance Imaging*, 27(1):132–139, jan 2008.

[176] L. Zhang, M. Tang, S. Chen, X. Lei, X. Zhang, and Y. Huan. A meta-analysis of use of prostate imaging reporting and data system version 2 (pi-rads v2) with multiparametric mr imaging for the detection of prostate cancer. *European radiology*, 27(12):5204–5214, 2017.