Motion Correction and Parameter Estimation in DCE-MRI Sequences: Application to Colorectal Cancer

Thesis submitted by: Manav Bhushan, St.Cross College
Michaelmas Term 2013

Supervisors: M. Jenkinson, J. A. Schnabel and J. M. Brady

Abstract

Cancer is one of the leading causes of premature deaths across the world today, and there is an urgent need for imaging techniques that can help in early diagnosis and treatment planning for cancer patients. In the last four decades, magnetic resonance imaging (MRI) has emerged as one of the leading modalities for non-invasive imaging of tumours. By using dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI), this modality can be used to acquire information about perfusion and vascularity of tumours, which can help in predicting response to treatment.

There are many factors that complicate the analysis of DCE-MRI data, and make clinical predictions based on it unreliable. During data acquisition, there are many sources of uncertainties and errors, especially patient motion, which result in the same image position being representative of many different anatomical locations across time. Apart from motion, there are also other inherent uncertainties and noise associated with the measurement of DCE-MRI parameters, which contribute to the model-fitting error observed when trying to apply pharmacokinetic (PK) models to the data.

In this thesis, a probabilistic, model-based registration and parameter estimation (MoRPE) framework for motion correction and PK-parameter estimation in DCE-MRI sequences is presented. The MoRPE framework is first compared with conventional motion correction methods on simulated data, and then applied to data from a clinical trial involving twenty colorectal cancer patients. On clinical data, the ability of MoRPE to discriminate between responders and non-responders to combined chemo- and radiotherapy is tested, and found to be superior to other methods. The effect of incorporating different arterial input functions within MoRPE is also assessed.

Following this, a quantitative analysis of the uncertainties associated with the different PK parameters is performed using a variational Bayes mathematical framework. This analysis provides a quantitative estimate of the extent to which motion correction affects the uncertainties associated with different parameters. Finally, the importance of estimating spatial heterogeneity of PK parameters within tumours is assessed. The efficacy of different measures of spatial heterogeneity, in predicting response to therapy based on the pre-therapy scan alone are compared, and the prognostic value of a new derived PK parameter – the 'acceleration constant' – is investigated. The integration of uncertainty estimates of different DCE-MRI parameters into the calculation of their heterogeneity measures is also shown to improve the prediction of response to therapy.

Acknowledgements

First and foremost, I would like to thank my supervisers – Mark Jenkinson, Julia Schnabel and Sir Mike Brady – for their patience, guidance, and unwavering support, without which this thesis would never have been possible. I have been extremely fortunate to have not one, but three wonderful mentors, who have taken a keen interest in my work throughout the entire duration of the DPhil. But I owe a special thanks to Mark, my primary superviser, who has been exceedingly generous with his time and words of advice when I needed them the most.

I am also very grateful to the Oxford Cancer Imaging Center (OCIC), the EPSRC and Cancer Research UK, for providing the funding that made this DPhil possible, and to Julia for making sure that I never lacked any material or logistical support. I am also grateful to our colleagues and collaborators in the Churchll hospital – Dr. Fergus Gleeson, Dr. Mark Anderson, Jamie Franklin, Steve Turnbull – for providing the data as well as valuable advice.

And finally, I am very grateful to all my wonderful friends and colleagues in Oxford, and my family in India, without whom this journey would never have been possible.

Contents

Chapter 1

Introduction

In order to treat any disease, the first task that needs to be completed is to understand the nature of the disease and how it spreads in the body. Although we have come a long way in our understanding of cancer and in the development of techniques to treat it, there is still a lot that remains to be understood. The ultimate goal in this project is to use the information available to us through dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI), in order to make clinically relevant observations and predictions about colorectal tumours.

1.1 The Clinical Question

According to the Cancer Research UK newsletter for May 2013, cancer causes more than one in every four deaths in the UK, with 157,250 deaths caused in 2010. Worldwide, 12.7 million cancer cases and 7.6 million cancer deaths are estimated to have occurred in 2008 alone [Jemal et al., 2011]. Colorectal cancer has been found to be the third most commonly diagnosed form of cancer in men and women [Siegel et al., 2013]. The standard treatment for most types of cancer (especially colorectal cancer) is still surgery, chemotherapy, radiotherapy, or some combination of the three.

In the treatment protocol followed at many hospitals across the world, incuding the Oxford University hospitals, colorectal cancer patients are first sent for 6-12 cycles of neoadjuvant chemoradiotherapy (CRT), and then for total mesorectal excision (TME), which is the surgical removal of the bowel around the tumour site. After the surgery is performed, the tumour tissue is assessed histopathologically to determine the extent and nature of the tumour. In 15-20% of the patients, the tumour is found to contain no viable cancer cells upon excision, indicating that the surgery was probably unnecessary [Wynn et al., 2010].

In another 30% of the patients, the tumour is found to have been resistant to CRT and not responded at all. In this class of patients, at least a few cycles of CRT were avoidable, and the patient could have been sent straight to surgery. Currently, there is no reliable method for distinguishing between the 15-20% complete responders, the 30% non-responders and the remaining partial responders at an early stage during treatment, so that unnecessary pain, suffering and expenses could be avoided for colorectal cancer patients. Since both CRT and TME have short term and long term impacts on the patient's quality of life, it would be of great clinical significance if a non-invasive imaging technique could help in avoiding excessive treatment, without negatively impacting the end result.

1.2 The Importance of Imaging

In the course of the last century, many different techniques for non-invasive imaging have been developed, which enable clinicians to get an insight into the anatomy and physiology of a patient without making any incisions. Starting with X-rays, there have been huge advances in the use of computed tomography (CT), ultrasound imaging, positron emission tomography (PET) and magnetic resonance imaging (MRI). However,

with the exception of ultrasound imaging and MRI, all the others involve ionizing radiation or the administration of a radioactive tracer to the patient, and therefore entail undesirable side-effects.

Although PET has the ability to show metabolic and physiological activity, the spatial resolution of a PET scan is limited to 3-5 mm due to the physical properties of PET image formation. Ultrasound imaging also suffers from low resolution, low signal-to-noise ratios, and an inability to show soft tissue contrast in regions that are far from the imaging probe. In this context, MRI has emerged as one of the leading modalities for oncological imaging, due to its ability to show soft tissue contrast throughout the patient body, and the lack of any ionizing radiation or other known side-effects.

An anatomical MRI scan provides us with scalar intensities at every location in the image volume grid. Each location in the image grid corresponds to some anatomical location in the body, and the scalar intensity observed in the MRI scan is a complex function of the concentration of hydrogen nuclei, predominantly in water or fat molecules, and their magnetic properties at that location (as explained in Section 2.2). A DCE-MRI scan is a sequence of MRI scans taken at regular intervals after the injection of a contrast agent (CA) into the body. Thus, instead of a scalar value, the DCE-MRI scan provides us with a time-intensity curve at every location, which can be used to derive the contrast take-up curve (CTC).

The form of the CTC can provide very valuable information about the perfusion and vascularity in different parts of the tissue [Zahra et al., 2007]. Since the treatment of colorectal cancer is dependent upon the delivery of chemotherapeutic agents to the disease site, information about perfusion and vascularity is absolutely crucial for accurately predicting the response of different patients to CRT. However, extracting accurate and clinically meaningful information about the tumour vasculature from a 4D DCE-MRI scan is a difficult process due to the size and complexity of the data acquired.

In order to facilitate a clinical interpretation, biologically meaningful parameters need to be extracted from this data.

1.3 Challenges in Using DCE-MRI

Patient motion, noise during the scan, and the complex relationship between the MRI signal and the concentration of CA are the main challenges faced in using DCE-MRI data for oncological purposes. Once these issues have been dealt with, the CTC acquired at each voxel in the image space needs to be linked to biological properties of the tissue in a meaningful way.

A lot of work has been done in trying to relate the data from DCE-MRI scans to biological processes in the body, and several pharmacokinetic (PK) models have been developed to parameterize the CTCs acquired at each location [Tofts, 1997], [Hayton et al., 1997], [Naish et al., 2009]. These PK models fit a functional curve derived from a biological model, to the concentrations of CA observed at each voxel. This allows the CTC at each voxel to be characterized by a small number of PK parameters that are directly or indirectly linked to the physiological properties of the tissue.

The biggest hurdle in the PK analysis of DCE-MRI data, is the fitting error caused by a combination of motion and noise, which makes the PK parameters inaccurate and hard to interpret. Some of the ways that people have dealt with this problem is to use Bayesian estimation of the 'true' CTCs [Schmid et al., 2009a], or an image-registration technique for motion correction [Hayton et al., 1997], [Buonaccorsi et al., 2005] and [Zöllner et al., 2009]. However, most conventional motion correction methods rely on similar intensities or statistical compositions of images acquired at different time-points in a DCE-MRI scan. In practice, these conditions do not hold in the case of a scan with low temporal resolution, where the image contrast changes drastically between

consecutive time-points due to the uptake of the CA. Thus, there is a need for motion correction methods that take the uptake of CA into account.

There is also a need for estimating the errors and uncertainties present in PK parameter estimates, as shown in [Garpebring et al., 2012] and [Buckley, 2002]. This need is even more acute, when the objective is to make clinical predictions on the basis of DCE-MRI scans alone, since errors in PK analysis can easily get propagated and even enhanced in the process of making clinical predictions. Several studies have shown that PK parameters are correlated with histological properties of a tumour [George et al., 2001], [Hawighorst et al., 1998]. However, it has also been shown that the spatial distribution of the PK parameters, i.e. tumour heterogeneity, has a very important role to play in predicting patient prognosis [Yang and Knopp, 2011], [Rose et al., 2009a]. Thus, estimation of uncertainties and heterogeneity in PK parameter estimates are two important aspects that need to be tackled in order to make accurate predictions about colorectal cancer patients' response to CRT.

1.4 Scientific Contributions and Overview

The basic aim of this thesis, is to use the information extracted from DCE-MRI scans in order to make accurate and clinically valuable predictions about colorectal cancer patients' response to therapy. In pursuit of this objective, we first provide a broad overview of the basic concepts involved in DCE-MRI image analysis for oncological purposes in Chapter 2.

Subsequently, we present our own probabilistic mathematical framework for a PK model-based registration and parameter estimation (MoRPE) method in Chapter 3. The main aim of this framework is to overcome the shortcomings of conventional image registration methods, and correct for motion during a DCE-MRI scan using a physio-

logical image formation model based on the uptake of CA. The probabilistic framework of MoRPE is then compared against other conventional image similarity measures for its ability to recover accurate estimates of PK parameters from data that has been artificially corrupted with motion.

The MoRPE framework is applied to data from a colorectal cancer study in Chapter 4, and the difference between the PK parameters before and after five weeks of CRT is used to predict response to therapy. Here again, the MoRPE framework is compared against other motion correction and PK parameter estimation methods for its ability to accurately predict the response of colorectal cancer patients to CRT. The importance of using different arterial input functions (AIFs) within the framework of MoRPE is also assessed.

In Chapter 5, we present a quantitative analysis of uncertainties involved in PK analysis of DCE-MRI data using a Variational Bayes mathematical framework. The purpose of this analysis is to distinguish between different sources of uncertainty in PK parameter estimates, and obtain information on how PK parameter estimation could be further improved in the future. In Chapter 6 we discuss the importance of estimating heterogeneity in PK parameter maps, and use different measures of heterogeneity to predict outcome to therapy on the basis of the pre-therapy scan alone. Finally, in Chapter 7 we present a short summary of the conclusions and avenues for future work arising from this thesis.

Chapter 2

Background Material

In order to motivate our approach, we will first briefly discuss the clinical question, and the measurable characteristics of a tumour in Section 2.1. The basic principles underlying MRI will be described in Section 2.2, and in Section 2.3 we will build an understanding of how a DCE-MRI scan combines the advantages of a basic MRI scan with the biological properties of a contrast agent (CA) to extract physiological information, instead of simply capturing anatomical information. The problem of relating DCE-MRI data with physiological properties of the tissue will be discussed in Section 2.4.

Although a DCE-MRI scan has the ability to provide physiological information, it also poses many problems for clinicians seeking to analyze and interpret the data contained in the scans. Since a DCE-MRI involves the acquisition of several MRI scans over a period of 5-10 minutes, the patient motion occurring during a DCE-MRI scan is a significant problem. Fortunately, the field of medical image analysis is highly advanced, and provides many methods for dealing with the errors introduced due to motion after the images have been acquired [Maintz and Viergever, 1998a]. Since motion correction for DCE-MRI sequences is a key focus of this thesis, we will review the existing methods

for alignment (registration) of images acquired at different time points in Section 2.5. In Section 2.6, we will examine different theoretical frameworks that enable joint motion correction and parameter estimation on DCE-MRI sequences on the basis of different PK models. The process of linking the PK modeling to clinical predictions will be discussed in Section 2.7.

2.1 The Clinical Problem

As discussed in Chapter 1, the primary aim of this thesis is to develop a comprehensive framework for the analysis of DCE-MRI scans, which can be used to reliably predict the response of colorectal tumours to CRT at an early stage during treatment. In order to develop such methods, it is important to briefly discuss the general characteristics of solid tumours.

2.1.1 Biological Characteristics

The process of cell-division in the human body is governed by a complex set of biological mechanisms that are encoded in the DNA present in each cell. The growth of a tumour represents the breakdown of the set of rules that govern cell-division, and results in the uncontrolled proliferation of a certain set of cells [Weinberg, 2013] at the cost of others. This results in the death of other cells that are vital for the functioning of the body, and the accumulation of dead cells, which can usually be seen as a necrotic core in the central part of most solid tumours [Foryś and Mokwa-Borkowska, 2005], including colorectal tumours.

The vascular structure in and around a solid tumour is often characterized by the growth of tortuous, leaky blood vessels that are inefficient in delivering drugs or nutrients to the disease site. Since one of the primary nutrients carried by haemoglobin

in the blood stream is oxygen, hence abnormal vasculature in the tumour microenvironment often results in a lack of oxygen or 'hypoxia' [Hockel et al., 1996]. Chronic hypoxia represents an increase in the distance that oxygen has to diffuse, while acute hypoxia is related to dynamic changes in perfusion [Simonsen et al., 2012]. Hypoxia is extremely important in the context of tumours because its presence is thought to increase the resistance of a tumour to both radiation [Bertout et al., 2008], as well as chemotherapy with certain cytotoxic agents [Orlowski et al., 2012].

Apart from the presence or absence of oxygen in the tumour microenvironment, one of the other factors that has a big impact on the growth of a tumour, is the formation and growth of new blood vessels (angiogenesis) in the immediate neighbourhood of the tumour. This process is mediated by many cellular pathways, of which the vascular endothelial growth factor (VEGF) pathway is a key one [Shibuya, 2008]. The VEGF family of proteins and receptors are known to play an important role in the process that leads to angiogenesis, and the over-expression of VEGF has been found to be indicative of disease progression and poor prognosis in the case of colon cancer [Lee et al., 2000], as well as other types of cancer [Reinmuth et al., 2003], [Berns et al., 2003]. Interestingly, VEGF has been found to play a causal role not just in the formation of new blood vessels, but also in increasing their permeability [Dvorak et al., 1995], [Brown et al., 1997].

2.1.2 Measurable Quantities

The blood flow, oxygenation and vascularization of tumour tissue are all indicators of the physiological nature of the tumour, and therefore the prognosis of the patient. The partial pressure of oxygen, or expression of VEGF, etc can all be measured from histological samples collected after resection of a tumour, or even from biopsy samples collected before or during treatment. Another quantity that can be measured from histological samples, and which is used frequently as a measure of tumour vascularization,

is the microvessel density (MVD) [Weidner, 1995].

The most basic feature of a solid tumour that can be assessed through non-invasive imaging (usually CT or MRI) is a change in the size of the tumour, caused by therapy. The two most popular sets of guidelines that have been used to assess response by measuring a change in the size of the tumour, are the WHO guidelines, and the more recently developed Response Evaluation Criteria in Solid Tumours (RECIST) guidelines [Park, 2003]. Historically, tumour size was determined by calculating the product of the longest diameter and the second longest diameter perpendicular to it [James et al., 1999]. However, the RECIST criterion was evolved in 1998 to simplify the measurements and thus make observations from different clinicians more consistent with each other [Therasse et al., 2000].

Measuring the tumour dimensions is advantageous because it can be performed non-invasively using imaging, however, it is not always a reliable measure of response. The spatial measurement is very simplified, and does not take morphological appearance into account (which could be indicative of certain types of pathology). Hypoxia, angiogenesis and tumour grade are all physiological indicators that can provide information about the disease progression and about the patients response to therapy. The tumour grade is usually a measurement of the number of viable tumour cells left in the tumour after therapy, and there exist many different systems of grading tumours that are followed for different types of cancer. In the case of colorectal cancer, the competing grading systems are quite similar, and they measure response on the basis of the number of tumour cells visible in microscopy images of histological samples from the tumour area collected after therapy [Bateman et al., 2009].

Tumour grade, VEGF expression or MVD may be more reliable as measures of response than the tumour size alone, but information about these factors can only be obtained through a histology conducted on samples collected through an invasive clin-

ical procedure. DCE-MRI provides one possible method of obtaining such information through non-invasive imaging techniques. In Section 2.7 we will discuss the links between data gathered through DCE-MRI, and biological characteristics determined from histological data.

2.2 Basics of MRI

Magnetic Resonance Imaging (MRI) has been used widely in the field of medical image analysis as a method for non-invasively imaging different parts of the body. It has proven to be especially useful in the field of oncology due to its ability to show contrast between soft tissues, and its lack of ionizing radiation. A brief description of MR imaging is presented below. A more detailed explanation of the manner of excitation and relaxation of the hydrogen nuclei can be found in [Hashemi et al., 2010] or [Liang and Lauterbur, 2000]. MRI scans can be used to image any element with paramagnetic properties, but in most clinical protocols, it is used to image hydrogen nuclei, due to the abundance of water in the human body. Thus, in the following discussion, we will discuss the basic principles of an MRI scan with respect to hydrogen nuclei:

1. Hydrogen nuclei have a property known as 'nuclear spin', i.e. they behave like tiny rotating magnets. In the absence of any external magnetic field, the net magnetization of the hydrogen nuclei present in the tissue is zero.

2. In an MRI scanner, a strong magnetic field, referred to as B_0, is applied in one direction (usually along the z-axis). This induces the magnetic moments of the hydrogen nuclei, predominantly from water molecules, to precess about the magnetic field B_0, so that the net magnetization vector of the tissue being imaged (M_0) is in the same direction as B_0. However, the nuclear spins of the hydrogen nuclei are still out of phase with each other.

3. A Radio-Frequency (RF) pulse is repeatedly applied to the tissue, after short time intervals (known as the Repetition Time – TR) at the Larmor frequency of hydrogen nuclei, which is the natural frequency at which the hydrogen nuclei precess about their axes under the effect of an external magnetic field. The RF pulse, followed by relaxation and/or spoiling, causes the magnetic moments of the hydrogen nuclei in the tissue to start rotating in phase with each other. As a result of the RF pulse, a rotating net transverse magnetization is added, which can be seen as the tipping of the net magnetization M_0 away from the z-axis in the rotating frame of reference.

4. After the application of the RF pulse, the spins of the hydrogen nuclei gradually dephase, and the net magnetization (M_0) slowly returns to the same direction as B_o. The dephasing is a result of two relaxation processes – (1) spin-lattice (T_1) relaxation and (2) spin-spin (T_2) relaxation – which will be discussed later.

The change in the angle of M_0 (in the rotating reference frame) caused by the application of the RF pulse, is known as the flip-angle α. The flip-angle is dependent upon the magnitude and duration the RF pulse. The time taken (after applying the RF pulse) for the magnetic field along B_0 to regain 63% of its original strength, is known as the T_1-relaxation time, or the spin-lattice relaxation time. The T_1-relaxation time is a function of the efficiency with which energy is exchanged between the precessing hydrogen nuclei and the surrounding lattice. Assuming that the hydrogen nuclei are fully relaxed before the application of the RF pulse, the magnetic field along the z-axis after the application of a 90 degree RF pulse, can be expressed as:

$$M_z(t) = M_0 \left(1 - \exp\left(-t/T_1\right)\right) \tag{2.1}$$

Here t is the time that has elapsed since the application of the last RF pulse, T_1 refers to the known T_1-relaxation time, which is constant for a specific tissue at a specific magnetic field strength. The magnitude of the rotating magnetic field in the transverse direction, i.e in the xy-plane, is at its maximum immediately after application of the 90 degree RF pulse, and it decays according to the formula:

$$M_{xy}(t) = M_0 \exp\left(-t/T_2^*\right) \tag{2.2}$$

Here t is the same as above, α is the flip-angle and T_2^* refers to the spin-spin relaxation time, or the T_2-relaxation time, which is again tissue-specific like the T_1-relaxation time. The decay of the magnetic field strength in the transverse plane is a consequence of the dephasing of the rotating magnetic-field vectors of the individual nuclei. Apart from the flip-angle (α) and the repetition time (TR), another acquisition parameter that can be manipulated to suit the image contrast, is the time to echo (TE) – i.e. the time gap between the application of the RF pulse and the measurement of the net magnetic field in the transverse direction. The combined effect of all these acquisition parameters on the measured signal strength S at a given location, is given by:

$$S = M_0 \left(1 - \exp\left(-TR/T_1\right)\right) \exp\left(-TE/T_2^*\right) \tag{2.3}$$

Here M_0 is the net magnetization immediately after application of the 90 degree RF pulse. For a smaller flip-angle α, the net magnetization vector (immediately after application of the RF pulse) along the longitudinal direction can be expressed as $M_z = M_0 \cos(\alpha)$, and along the transverse direction, in the rotating frame of reference, can be expressed as $M_{xy} = M_0 \sin(\alpha)$.

In order to spatially encode the RF signals received from the hydrogen nuclei, gradient coils are used in each of the three directions. These gradient coils cause linear

variations in the Larmor frequency, the precessional frequency and phase of the rotating magnetic dipoles created by the hydrogen nuclei, and thus allow the signals being recorded by the receiver coils to be spatially encoded.

For the clinical data used in this thesis, a spoiled gradient echo (SPGR) sequence, with flip-angle $\alpha < 90^\circ$ was employed. In this type of sequence the transverse magnetization is destroyed after taking the measurement at time TE, with the use of a spoiler gradient pulse, in order to ensure that there is no residual transverse magnetization when the new RF pulse is applied. In this case, the signal is given by [Springer et al., 2010]:

$$S = M_0 \exp\left(-TE/T_2^*\right) \sin\left(\alpha\right) \frac{1 - \exp\left(-TR/T_1\right)}{1 - \cos\left(\alpha\right) \exp\left(-TR/T_1\right)} \tag{2.4}$$

In the following section, we will briefly discuss how an MRI scanner and a contrast agent (CA) can together be used to acquire physiological information about tissues.

2.3 From Anatomical to Functional Imaging

Water is one of the main constituents of the human body, constituting over 70% of the human body mass. As one would expect, the concentration of water is quite different in the different organs and types of organ parenchyma present in the body. The difference in the concentration of hydrogen nuclei, as well as other paramagnetic molecules results in different magnetic properties in different tissues, and allows us to see a contrast between their MRI intensities. Although this ability to visually distinguish between different soft tissues makes MRI a very useful tool for imaging the anatomy of a patient's body, it is still insufficient for determining the physiology of different tissues.

Thankfully this modality can be used to image physiology (e.g. vascularity and perfusion) by administering a paramagnetic contrast agent (CA) to the patient and

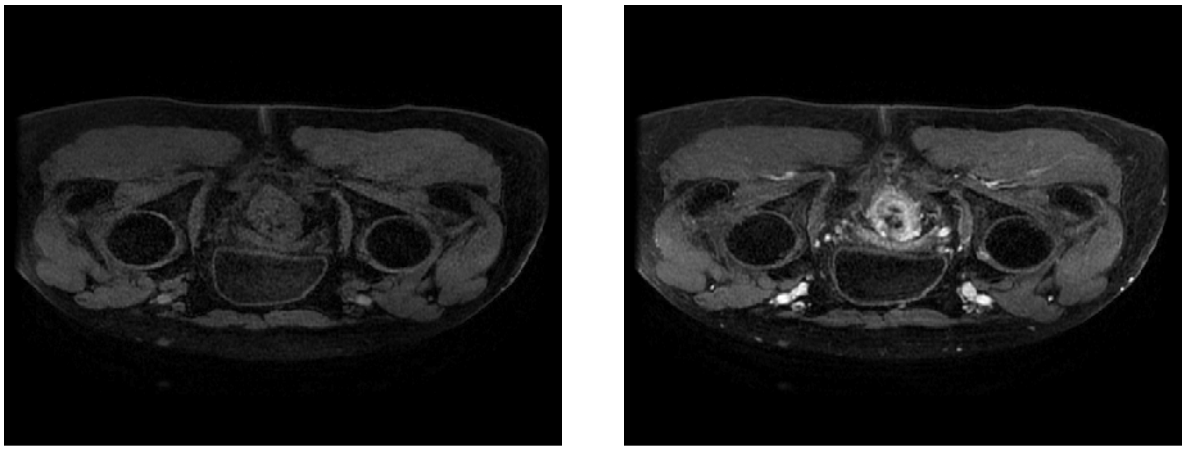

Figure 2.1: Central slices from the image volume acquired before injection of the CA (left) and the image taken 2 minutes after injection of CA (right) during a DCE-MRI scan of a colorectal tumour. The brighter portions of the right frame indicate the presence of the contrast agent.

then repeatedly acquiring over a period of time to build up a DCE-MRI scan. A DCE-MRI scan allows us to record the pattern of take-up and wash-out of the CA from different tissues, and thus provides information about the perfusion and vascularity of different areas. The key ingredient of a DCE-MRI scan is the CA, which is usually a molecule that can be administered safely to the body, and has the ability to change the paramagnetic properties of the hydrogen nuclei in its vicinity. The effect of the CA is usually exhibited as a change in the MRI intensity caused by the shortening of the T_1-relaxation time in the areas where it is present. As Figure 2.1 shows, in a T_1-weighted DCE-MRI scan, the tissue areas that retain a higher concentration of the CA appear brighter in the images acquired after the injection of the contrast agent.

The big advantage of a DCE-MRI scan over a normal MRI scan, is the fact that it allows us to get some measure of the physiology of each location in the tissue, in addition to its MRI intensity. As Figure 2.2 shows, two anatomical locations (one inside, and one outside the tumour) may appear to have exactly the same intensity in the pre-contrast MRI volume, but they can be distinguished by their pattern of CA enhancement. In

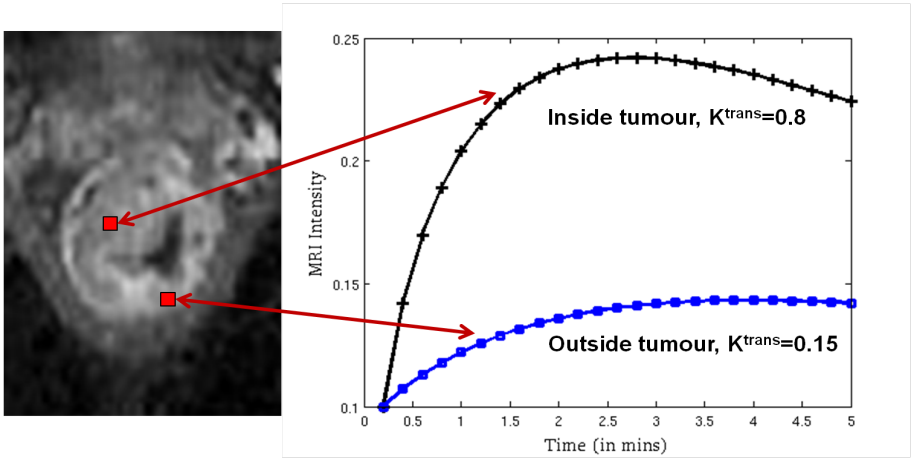

Figure 2.2: Two anatomical locations in the colorectum (left), which are indistinguishable in the pre-contrast volume, but clearly distinguishable by their time-intensity curves (shown on the right). The location with the higher value of K^{trans} is within the tumour, while the one with the lower value of K^{trans} is outside the tumour.

a similar manner, we can also observe different patterns of enhancement within the tumour, and therefore measure physiological heterogeneity within the tumour by way of a DCE-MRI scan.

A basic representation of the concentrations of CA in different tissues can be obtained through a simple contrast-enhanced MRI – where one MRI scan is acquired before injection of the contrast agent, and one after. However, this kind of scan does not allow the radiologist to obtain information about the vasculature, since it does not allow a calculation of how quickly or slowly the CA was taken up from the blood stream, and how quickly it was drained from the tissues. This crucial information, regarding the take-up and wash-out of the CA, can only be acquired in a *dynamic* contrast-enhanced MRI where the passage of contrast-agent is imaged over a period of time (Figure 2.3).

The most commonly used CA is Gadolinium Diethylenetriamine Pentaacetic Acid (Gd-DTPA), which consists of the paramagnetic trivalent ion of the Lanthanide element Gadolinium (Gd^{3+}) attached to a chelate (DTPA), as shown in Figure 2.4. Although Gadolinium can be toxic if administered by itself, the presence of the chelate makes the

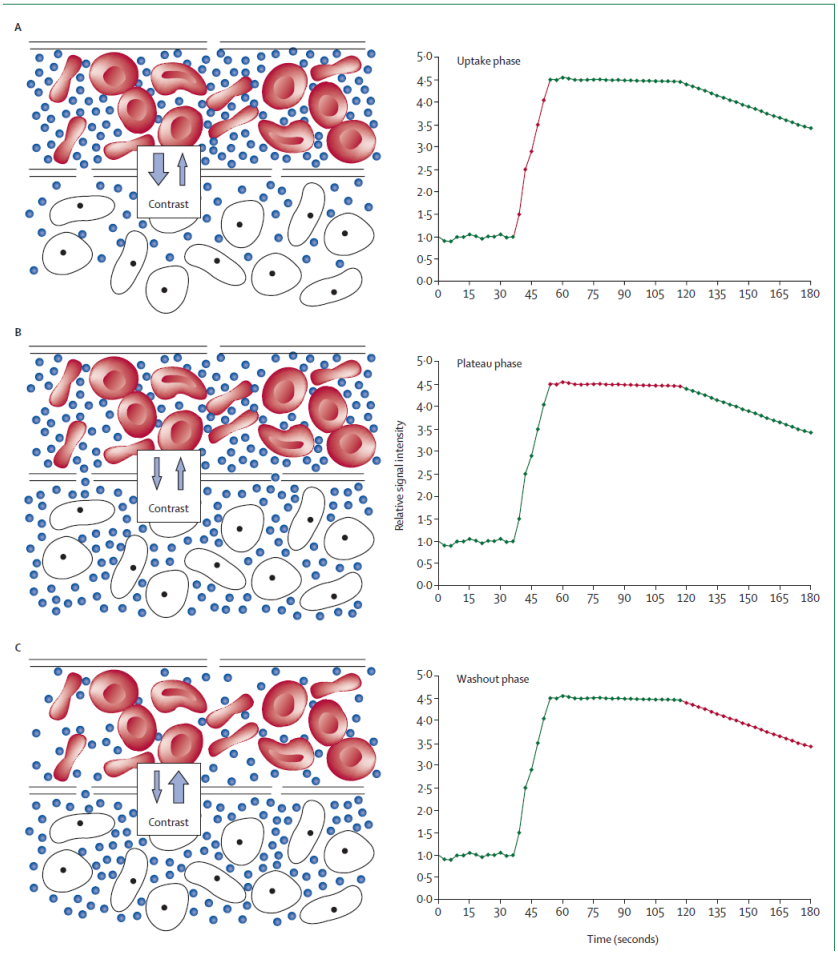

Figure 2.3: An illustration showing the exchange of contrast agent between the blood plasma (top in each panel) and the extra-cellular extra-vascular space (bottom in each panel) during different phases of the DCE-MRI scan (top: uptake; middle: plateau; bottom: washout). Figure taken from [Zahra et al., 2007].

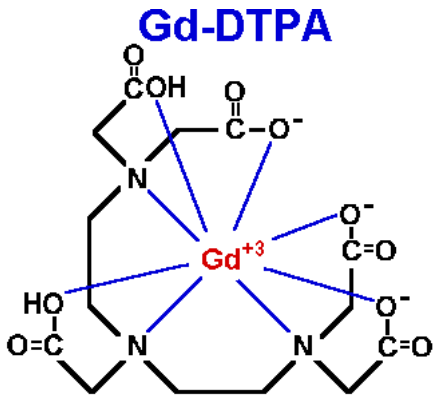

Figure 2.4: The Gadolinium-DTPA molecule.

molecule stable, soluble in water, and biologically safe (except for patients with acute or chronic severe renal insufficiency) without significantly diminishing its paramagnetic effects. Most MRI contrast agents, including Gd-DTPA, are designed to have a low molecular weight that allows them to move freely from the blood plasma in the arteries to the extra-cellular, extra-vascular space (EES) in the tissues [Weinmann et al., 1984]. Thus, the distribution of the CA in the body tissues, as seen in a DCE-MRI scan, can be used as an indicator of how nutrients and drugs such as chemotherapeutic agents may get absorbed in different tissues.

In order to make inferences about biological processes from DCE-MRI data, one option is to calculate semi-quantitative measures based on the time-intensity curves themselves. These measures – such as relative signal enhancement (RSI), time to peak, etc. – are easier to calculate than PK parameters, since they do not require conversions between the signal intensities and the concentration of CA. However, they suffer from many inadequacies, such as the fact that they are not directly representative of any biological process. In order to relate DCE-MRI data to physiological properties of the tissue, we need to accurately estimate the concentration of CA in different tissues as well as the underlying T_1 value of those tissues.

As discussed earlier, the concentration of CA in a particular area has a positive

correlation with the MRI intensity (in the T_1-weighted scan) of that area, but this relationship is not entirely straightforward. The concentration of CA alters the MR signal by changing the T_1 and T_2^* relaxation times of the tissue in which it is present [Tofts and Kermode, 1991]. If we use T_{10} and T_{20}^* to denote the original T_1 and T_2^* relaxation times of the tissue, then the post-contrast relaxation times can be expressed in the following manner:

$$\frac{1}{T_1} = \frac{1}{T_{10}} + r_1 C_T(t) \tag{2.5}$$

$$\frac{1}{T_2^*} = \frac{1}{T_{20}^*} + r_2 C_T(t) \tag{2.6}$$

Here T_1 and T_2^* are the new relaxation times after injection of the contrast agent, and $C_T(t)$ is the total concentration of the contrast agent at time t, and r_1 and r_2 are constants that depend upon the contrast agent being used, as well as the strength of the scanner's main magnetic field. If multiple images are acquired at different time-points after the injection of the contrast agent, then we can substitute T_1 and T_2^* in place of T_{10} and T_{20}^* in the following equation for the MRI signal in a SPGR sequence [Liang and Lauterbur, 2000]:

$$S = M_0 \exp\left(-TE/T_{20}^*\right) \sin\left(\alpha\right) \frac{1 - \exp\left(-TR/T_{10}\right)}{1 - \cos\left(\alpha\right) \exp\left(-TR/T_{10}\right)} \tag{2.7}$$

Using the above, we can derive a relationship between the observed MR intensity and the concentration of contrast agent at each voxel in the image [Donaldson et al., 2010b]. The relationship can be summarized as the following:

$$\frac{S(t)}{S(0)} = \exp\left(-r_2 C_T(t) TE\right) \frac{1 - \exp\left(-P - Q\right) - \cos\left(\alpha\right)\left(\exp\left(-P\right) - \exp\left(-P - 2Q\right)\right)}{1 - \exp\left(-P\right) - \cos\left(\alpha\right)\left(\exp\left(-P - Q\right) - \exp\left(-P - 2Q\right)\right)} \tag{2.8}$$

Here $P = TR/T_{10}$, $Q = r_1 C_T(t) TR$, $S(t)$ denotes the MR signal intensity at time t and $S(0)$ denotes the signal intensity at time 0, before CA is injected. In the above equations, all the terms, except for $C_T(t)$ and T_{10} are either known, or can be estimated directly from the data. Since we are ultimately interested in estimating the concentration of the contrast agent over time, and relating this to physiological characteristics of the tissue by way of PK models, all that remains to be done is to estimate T_{10}. There are many ways of estimating T_{10}, the most widespread of them being the variable flip-angle method, which is used in [Fram et al., 1987].

In order to estimate the inherent T_1-relaxation time of the tissue (also known as T_{10}) using the above-mentioned method, we can denote $\exp\left(-TR/T_1\right)$ as R, and rewrite Equation 2.4 as follows:

$$S\left(1 - R \cdot \cos\left(\alpha\right)\right) = M_0 \exp\left(-TE/T_2^*\right) \sin\left(\alpha\right) \cdot \left(1 - R\right) \tag{2.9}$$

In order to convert this expression into an equation of the form $y = mx + c$, we can further rewrite this as:

$$\frac{S}{\sin\left(\alpha\right)} = R \cdot \frac{S}{\tan\left(\alpha\right)} + M_0 \exp\left(-TE/T_2^*\right) \cdot \left(1 - R\right) \tag{2.10}$$

Now, if MRI volumes of the tissue of interest were to be acquired at different flip-angles (i.e. different values of α), then we would obtain a series of images where only the flip-angle and the MRI intensity (S) is varying, while everything else is held constant. From two or more such data points, the slope and intercept of the line expressed in Equation

2.10 can be calculated. If the slope of the line is found to be m, it can be surmised that $T_{10} = -TR/\log(m)$. The same method is followed in [Deoni et al., 2003], which additionally provides insights on how to make the method more efficient by optimizing the flip-angles at which the images are acquired.

2.4 Pharmacokinetic Modeling

Once T_{10} has been determined in the manner described above, Equation 2.8 can be used to calculate the concentration of CA over time, i.e. $C_T(t)$. Most Gadolinium-based DCE-MRI contrast agents are designed to be non-specific (i.e. they do not bind to a particular protein or pathway), and their low molecular weight and solubility allow them to be easily transferred between the blood vessels and the extracellular extra-vascular space (EES). For these reasons the variation in the concentration of the contrast agent over time is a useful indicator of tissue vascularity and perfusion.

The contrast take-up curve (CTC) acquired at each voxel in a DCE-MRI scan shows us the concentration of CA at each voxel as a function of time. The simplest way of using the information from the CTCs is to calculate semi-quantitative measures such as the slope of the CA uptake curve, the time to peak, or the area under the gadolinium curve (AUGC) at some fixed time interval after injection of the CA – usually 90s. Although these quantities contain some information on the rate and extent of enhancement, it is difficult to directly link them to the underlying physiological properties of the tissue [Zahra et al., 2007].

The most popular, and clinically meaningful way of extracting information from the CTCs has been to apply a physiology-based mathematical model to the data that is obtained at each voxel, and to then estimate the parameters that most accurately describe the observed data. In [Tofts and Kermode, 1991], Tofts and colleagues pre-

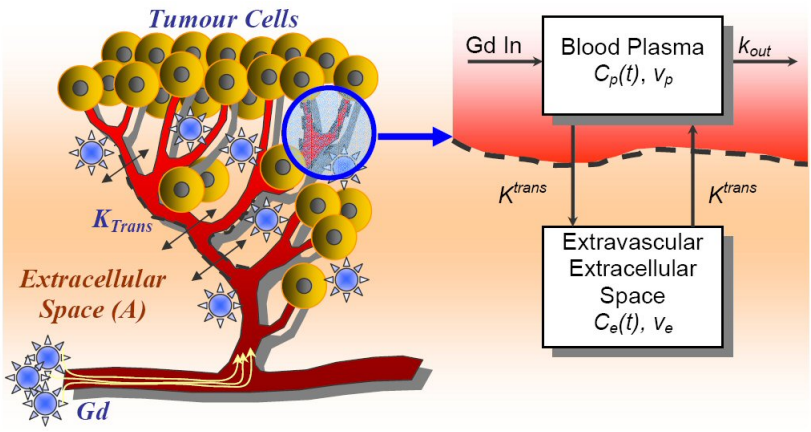

Figure 2.5: The two-compartment model (Figure taken from [Tanner, 2010]).

sented a model based on tumour physiology that predicted the CTC (and therefore the signal-intensity curve) at each voxel. This model was based on the assumption that the tumour tissue is divided into two compartments – blood plasma, and extra-cellular extra-vascular space (EES) – that are assumed to be well-mixed throughout the tumour tissue. The nature of the CTCs at each voxel is then a function of the proportion of space occupied by each compartment in that voxel, and the ease with that the CA can move between them.

Using the above model, and by applying the principle of conservation of mass, we arrive at the following differential equations [Tofts and Kermode, 1991]:

$$v_p \frac{dC_p}{dt} = M_{in} - (k_{out} + k_{pe}) v_p C_p(t) + k_{ep} v_e C_e(t) \tag{2.11}$$

$$v_e \frac{dC_e}{dt} = k_{pe} v_p C_p(t) - k_{ep} v_e C_e(t) \tag{2.12}$$

Here M_{in} refers to the rate of injection of the CA into the bloodstream (in mol/s), v_p and v_e are the volumes of the blood plasma and the EES respectively (expressed as proportions of the total volume) and k_{ep} and k_{pe} are the rate constants for CA going

from EES to blood plasma and blood plasma to EES respectively. $C_e(t)$ and $C_p(t)$ are the concentrations of the contrast agent in the EES and blood plasma at time t. In most implementatons of the model, $C_p(0)$ is assumed to be 0. One of the other fundamental assumptions of the Tofts model is that the system is in equilibrium, and therefore the amount of CA going into and out of blood plasma remains the same. Thus, he defined a new parameter K^{trans}, which can be expressed as:

$$K^{trans} = k_{pe}v_p = k_{ep}v_e \tag{2.13}$$

Most studies based on the standard Tofts model are primarily concerned with the values of the variables K^{trans}, k_{ep}, v_e and v_p. K^{trans}, also known as the transfer constant, is thought to be of great importance, although its physical interpretation changes on the basis of the assumptions of the PK model, as well as the nature of the tissue being imaged. In flow-limited situations, K^{trans} is though to represent the blood plasma flow per unit volume of tissue. In tissues that are permeability-limited, K^{trans} represents the permeability surface area product per unit volume of the tissue. The permeability or 'leakiness' of the blood vessels can be an important factor in determining how well the tumour will respond to chemotherapy or radiotherapy [George et al., 2001].

Similarly, v_e can be taken to be a measure of the amount of fluid in extra-cellular areas, while v_p tells us about the amount of vascularity in the tumour micro-environment. Thus, estimating the true value of these constants for each voxel can give us a lot of information about the grade and spread of the tumour, while the distribution of these values can shed light on the tumour heterogeneity [Rose et al., 2009b]. However, in order to determine these parameter values from the DCE-MRI data, we need to solve equations (2.11) and (2.12) to obtain an expression for $C_e(t)$. Tofts and colleagues make the following two assumptions while solving these two equations:

1. The CA is injected as a bolus (or a one-time injection), to obtain the following form, which enables us to say that $C_p(0) = \frac{D}{V_p}$, (where D is the dosage of the CA given to the patient), and neglect the term M_{in} in equation (2.11).

2. In the standard Tofts model, it is also assumed that the contribution to MRI signal intensity from the plasma volume is negligible, and thus the total concentration of CA at time t within a voxel can be written as

$$C_T(t) = v_e C_e(t) \tag{2.14}$$

(here $C_T(t)$ signifies the total amount of the CA in that voxel).

Based on these assumptions, and using Laplace transforms we can solve equations (2.11) and (2.12) to obtain the following form (detailed solution in [Tofts and Kermode, 1991]):

$$C_T(t) = K^{trans} \exp\left(\frac{-K^{trans}}{v_e} \cdot t\right) \otimes C_p(t) \tag{2.15}$$

Here $\otimes$ represents the convolution operation, and $C_p(t)$, which is the concentration of CA in the blood plasma, is commonly referred to as the arterial input function (AIF). Equation (2.15) describes the general form of the CTC when the physiological model parameters K^{trans} and v_e are known to us. Thus, in order to determine the model parameters from the DCE-MRI data, we need to fit a curve of the form given by equation (2.15) and calculate the parameters that provide the best curve-fit. In [Tofts et al., 1995a], Tofts and colleagues show that that their model provides reasonably good fits to actual data from DCE-MRI scans of breast tumours.

The Tofts two-compartment model has given rise to many other extensions and alterations that have resulted in other pharmacokinetic models. The simplest extension

to the Tofts model is the extended Tofts model, which does not assume the contribution of the blood pasma (v_p) to signal intensity to be negligible. Neglecting this assumption, Equation (2.14) is replaced by:

$$C_T(t) = v_e C_e(t) + v_p C_p(t) \tag{2.16}$$

Thus, for each voxel, there is one more parameter to estimate in the extended-Tofts model. Another model that was proposed by Hayton and colleagues in [Hayton et al., 1997] does not assume a one-time injection of contrast agent, and includes the CA injection term (M_{in}) while calculating the solution to the equations (2.11) and (2.12). In [Hayton et al., 1997] the authors calculate the analytic solution to the differential equations for three different cases: (1) when the injection of CA is instantaneous (as assumed in the standard Tofts model) (2) when it is injected gradually over a period of time using a slow drip, and (3) when there is a brief ramp-up followed by a constant flow. Thus, in the case of an instantaneous injection, the solution provided by Hayton and colleagues is:

$$C_T(t) = \frac{A}{a-b}\left(\exp\left(-bt\right) - \exp\left(-at\right)\right) \tag{2.17}$$

Here A, a and b are re-parameterization of the Tofts model parameters (K^{trans}, v_e, v_p). The advantage of the Hayton-Brady model over the Tofts model is that it incorporates the AIF within the model itself and does not involve any convolution operation (as in Equation (2.15)), or the independent determination of the AIF (that is a notoriously difficult task). However, the disadvantage of this model is that the relationship between model parameters a, b and A and the physiological characteristics of the tumour is more complex, and thus these model parameters are not as widely accepted in the clinical community as the parameters from the Tofts model. The

relationship between the parameters of the Hayton-Brady model and the Tofts model have been explained in detail in [Armitage et al., 2005].

Another PK model that has been applied successfully to DCE-MRI data, is the adiabatic approximation to tumour heterogeneity (AATH) model [St Lawrence and Lee, 1998], which includes two additional PK parameters to account for the blood plasma flow rate, and the gradual (i.e. not instantaneous) mixing of the two compartments. This PK model was shown to provide better fits to DCE-MRI data for patients with thoracic malignancies [Naish et al., 2009].

There are many differences between the different PK models, but the primary aim of all PK models is to characterize the take-up and washout of the CA from the tissue, which, in turn, is indicative of the nature of blood flow and perfusion in the region where it is being measured. However, the physiological properties of a tumour discussed in Section 2.1.1 are not directly measurable from a DCE-MRI scan, since there is a huge gap in the spatial resolution of the vascularization and the DCE-MRI data. While newly formed blood vessels operate at the scale of micro or nanometers, DCE-MRI data has a resolution in the scale of millimeters. Thus, although the PK models are developed keeping cellular structures and microvessels at the scale on nanometers in mind, the DCE-MRI data that is acquired captures only an aggregate effect of the physiological factors in play. In order to capture the spatial heterogeneity of the tumour, it is extremely important to do accurate voxel-wise PK analysis of DCE-MRI data.

Arguably, the greatest difficulty that is encountered while trying to do voxel-wise analysis in DCE-MRI sequences, is the fact that there is a considerable amount of non-linear motion in the region of interest between consecutive images in a scan. This difficulty is especially pronounced in the case of colorectal cancer, since it is more difficult to constrain motion in this area than other body parts, such as the head, neck or breasts. The motion during a scan makes it difficult to fit any PK model reliably to the CTCs de-

rived from the raw data. Due to these difficulties, many research groups choose to obtain only summary statistics for the entire region-of-interest and then draw clinical inferences based on these summary statistics [Donaldson et al., 2010a], [Hawighorst et al., 1998]. However, recent studies have shown that apart from the grade, size or other summary statistics of the tumour, the spatial pattern of the PK parameters within the tumour is also likely to be of great clinical importance [Schmid et al., 2009a], [Rose et al., 2007].

In order to study the heterogeneity of the tumour, it is absolutely necessary to do a voxel-wise pharmacokinetic analysis of the DCE-MRI sequence. Thus, before PK model parameters can be derived from the data, all the images from a scan need to be aligned into one common coordinate system so that each voxel in the image space represents the same position in the tissue over time. In order to ensure this, we need to *register* all the images from a DCE-MRI scan to one common template. In the next section, we will briefly review conventional image registration methods and the similarity measures that they employ. After discussing the general principles of image registration, we will review existing methods for motion correction and PK analysis in Section 2.6, and then examine existing methods for making clinical predictions based on DCE-MRI data in Section 2.7.

2.5 General Principles of Image Registration

The field of medical image registration was spawned as a result of the use of imaging techniques such as Ultrasound, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), etc. These imaging techniques have enabled clinicians to take a closer look at the underlying anatomy of the body non-invasively and use the information available to distinguish different tissues in the body on the basis of different properties. CT depicts different tissues on the basis of

variable X-ray attenuation, Ultrasound on the basis of variable reflection of ultrasound waves and PET displays the varying amounts of radioactive tracer uptakes in different parts of the body and thus can be used to image metabolism or other functions, depending on the tracer that is used. MRI detects the response of different tissues to a strong magnetic field and intermittent Radio-Frequency (RF) pulses. However, in order to combine the information gathered from these different sources, image registration is required to align the images acquired using different modalities.

Image registration is also needed in longitudinal studies, where a patient is imaged at different stages of treatment, and anatomical or physiological changes need to be assessed. As discussed earlier, the recent advent of DCE-MRI facilitates the use of MRI to image tissue vascularity and perfusion, which provide vital insights into the nature of the tumour and the tumour micro-environment [Jackson et al., 2007]. One of the challenges while analyzing DCE-MRI sequences is the need to co-register all the images obtained during a scan so that each voxel in image space represents the same anatomical location across all the different time-point images. Without this, the accuracy of the extracted model parameters would be severely compromised. In the following section we will first describe the basic principles of image registration in mathematical terms.

2.5.1 Problem Description

We can define an image I as a function $I : X \to \mathbb{R}$ from some 2D or 3D grid X into the space of allowed intensities values. For the sake of notational convenience, in this section we will describe most equations for the 2D case. These can easily be generalized to the 3D case when necessary. In this setting, a transformation T is a function $T : X \to Y$ that deforms the underlying grid X into another grid Y. Thus, strictly speaking, when we apply a transformation T to an image I we obtain $T(I) = I(T(X)) = I(Y)$.

However, for simplicity, we refer to the transformed version of I as $T(I)$ rather than as $I(T(X))$. If $T: X \to Y$ is the transformation (where X and Y are 2D grids), then at any point $(x, y) \in X$, we can describe the transformation as $T(x, y) = (x + u,\ y + v)$, where $u(x, y)$ is the local deformation along the x-axis (in the horizontal direction) and $v(x, y)$ is the local deformation along the y-axis (in the vertical direction).

For any two images I_s (source image) and I_t (target image), we are interested in estimating the transformation T such that I_t is maximally aligned to $T(I_s)$ in as realistic a way as possible. First let us put these two concepts – the similarity measure and the regularization term – into mathematical terms.

1. A similarity measure can be defined, for any two images I_s (source image) and I_t (target image), as a function $S : I \otimes I \to \mathbb{R}$ that gives us a measure of how well the two images are aligned. Ideally, this function should attain its maximum value when I_s and I_t are in perfect alignment. In practice, it is sometimes better to have an objective function S that is minimized at the best alignment. In this case S is called a 'cost-function'.

2. A regularization term $R : T \to \mathbb{R}$, which should have a high value for deformations that seem to be 'unrealistic' and should be lower for more realistic, or smoother deformations. Such a function should attain its minimum when no deformation is applied, and should ideally not penalize simple translations or rotations that may be caused only by a change in the viewing position.

It has been standard practice in image registration to combine the two objectives listed

above into one suitably weighted energy term:

$$E(T) = S(I_t, T(I_s)) + \alpha \cdot R(T) \tag{2.18}$$

Here S is a cost-function (i.e. it is sought to be minimized) and α is the constant of proportionality that decides how much weighting the regularization should be given. Assuming that S and R have the properties listed above, we are interested in estimating the 'best' transformation $\hat{T}$, where

$$\hat{T} = \arg\min_{T} E(T) \tag{2.19}$$

The basic objective then, is to find the transformation $\hat{T}$ from Equation (2.19) using a suitable optimization strategy. In the following section, we will discuss the two important components – the transformation model, and the similarity measure – of any standard registration method in some more detail. We will first examine different transformation models and regularization terms in Section 2.5.2.

2.5.2 Transformation Models

The transformation models that are used in the bulk of image registration techniques are usually motivated by physics or computer vision. This is because physics and computer vision are concerned with modeling how objects or images change in nature, and how this change can be modeled as a dynamic process. Thus, it can be applied to the transformation of medical images, which can be regarded as the evolution of a natural system over time. A survey of image registration in the field of computer vision is given in [Zitova and Flusser, 2003]. However, though most methods used in medical image analysis are motivated by these approaches, all of these may not be directly

applicable to medical images because of the specific constraints that apply to them.

The first significant difference is that in most cases, the methods used in computer vision are meant for 2D images as captured by a camera. In most medical imaging applications, we are dealing with 3D images, and thus the techniques used in computer vision can become computationally intractable. The other difference is that in medical imaging, we are interested in simulating the non-linear deformation of body tissues. Thus, the number of degrees of freedom in medical images is often far more than in the images that computer vision is concerned with. In the case of dynamic imaging, the sampling rate in most computer vision applications is of the order of 25 Hertz, and is thus incomparable to the sampling rate in DCE-MRI sequences ($\approx 1/15$ Hertz) A comprehensive survey of the techniques used in medical image registration is presented in [Maintz and Viergever, 1998b].

The most basic transformation models that have been used in image registration are the rigid and affine transformations. In medical image registration these methods have been used successfully in some applications - such as registration of brains, which are relatively rigid objects [Shan et al., 2011]. However, for most body parts, a simple rigid or even affine transformation is not enough to capture the whole range of non-linear deformations that can take place in body tissues. A simple extension that can be used to simulate tissue deformations is the poly-rigid or poly-affine model [Arsigny et al., 2005]. This model assumes that the volume of interest is divided into many smaller components, each of that is deformed in a rigid (or affine) manner. A suitable interpolation method is used to calculate the new position of each point under the influence of multiple rigid body transformations.

One of the first methods taken from computer vision and applied successfully to medical imaging applications was the concept of optic flow. Using optic flow as an intermediate step, there have been successful attempts at describing the deformation

of images as a diffusion process in [Thirion, 1998]. This has given birth to many other methods that penalize deformations by defining the regularization term R in a manner that assumes that the image is getting deformed in a process akin to the diffusion of a fluid [Bro-Nielsen and Gramkow, 1996]. Similarly, other methods describe the transformation acting on the image as the flow of a viscous fluid [D'Agostino et al., 2002]. Other popular techniques that have been used are elastic registration, which assumes that each tissue is an elastic body having a certain Young's modulus. Once again, these methods penalize motion in an elastic framework, i.e. quadratic in the magnitude of the relative deformation, and thus constrain the motion using the elasticity equations [Christensen, 1999].

There have also been attempts to introduce biomechanical models into registration methods, so that the deformations are constrained to behave like body tissues. One of the approaches for doing this is to model the organ being imaged as being constructed from a finite number of building blocks, each of that has certain biomechanical properties. These methods are known as Finite Element Models (FEMs) and more details about these methods can be found in [Ferrant et al., 1999]. However, the boundary conditions for the differential equations that are encountered in FEMs are often problematic. Further, the parameters for each element in the model depend on the biomechanical properties of that tissue, and are largely unknown or difficult to measure.

Another popular transformation model that has the ability to simulate highly non-rigid transformations in a realistic way is the spline model. In a very general sense, splines are smooth polynomial functions that are defined piece-wise, and have an extra degree of smoothness at the points where different pieces connect with each other (knots). The most commonly used splines in registration are thin-plate splines and cubic B-Splines [Schnabel et al., 2003], [Rueckert et al., 2002].

The basic purpose of choosing any particular transformation model is to accurately model the deformations that are actually taking place in the tissues or organ being imaged. Thus, one of the important properties that any transformation should have (especially in studies where the same patient and same body part is imaged), is that there should be a bijective correspondence between the source image I_s and the target image I_t. However, apart from rigid and affine transformations, all other models that have been mentioned above do not ensure that the transformation T is bijective. Thus, there have been attempts in the more recent literature to make this an explicit constraint, and to ensure that that the transformation is bijective.

One class of transformations that satisfy the objective stated above – i.e. ensuring that T is smooth and models the actual deformation in body tissues – is known as the class of *diffeomorphisms*, which are one-to-one (bijective) transformations that must satisfy a certain smoothness constraint. After the publication of [Beg et al., 2005a], there have been many papers extending the approach followed by Beg and colleagues and trying to employ velocity fields in order to parameterize deformations and make them diffeomorphic, such as [Ashburner, 2007] and [Avants et al., 2008a]. Since a diffeomorphism is a smooth bijection, it is an ideal deformation framework to be used in applications where there is no expected change in the topology of the source and target images.

The choice of the regularization term depends entirely on the application for that the registration is being used. The regularization scheme used in our implementation is discussed in detail in Section 3.4. In the following section, we will discuss the second crucial component of registration energy functions, i.e. the similarity term.

2.5.3 Similarity Measures

In [Roche et al., 1999] and [Zollei et al., 2007] it has been shown that most of the

conventional image similarity measures can be derived in a probabilistic setting depending on the image formation model that is assumed. In this section, we will present these derivations in one unified framework without going into the details of each individual measure. In Chapter 3, we will go on to derive a new probabilistic similarity measure assuming a physiological image formation model that can be used to align 3D images from a DCE-MRI scan. There are a myriad of similarity measures that are currently used in different image registration problems depending on the particular application. The derivation of any particular similarity measure as opposed to another depends entirely on our assumptions about the following aspects:

1. The process of formation, and the noise model that we assume for the imaging process.

2. The nature of the relationship assumed between the intensities in the two images that are to be registered.

The simplest case is when we expect two images to be the same, barring the effects of noise and motion. In this scenario, if we use X to denote the source image and Y to denote the target image (both written as N-dimensional vectors, where N is the number of voxels in each image) then we can say that:

$$Y = T(X) + \epsilon \tag{2.20}$$

Here ϵ is some noise process. For MRI images, although the noise follows a Rician distribution, it is known to be approximately Gaussian for a moderate signal-to-noise ratio (SNR), and especially in high-intensity areas [Coupe et al., 2008]. Thus, it is standard practice to assume that ϵ is an independent and identically distributed Gaussian

noise process with mean 0 and standard deviation σ [Schmid et al., 2009a]. For the remainder of this thesis, the relationship between an image and its underlying grid will be kept implicit, and we will use X to denote the source image, and Y to denote the target image.

We are now interested in estimating the most likely deformation T that would have transformed X into Y. There are several approaches that have been used to derive similarity measures in a probabilistic setting, starting from some image formation model. In [Roche et al., 1999] and [Roche et al., 2000] the authors show how most of the existing similarity measures can be derived in a probabilistic framework by deriving an expression for the maximum a-posteriori (MAP) transformation that best 'explains' the data (i.e. X and $T(X)$).

If we assume that the intensities in the source and target image must be the same, barring the effect of additive, uncorrelated white Gaussian noise, then the approach followed in [Roche et al., 1999] shows that the most likely transformation T that deforms X into Y must be the one which minimizes the sum-of-squared differences (SSD) between the intensities of the two images.

Normalized corss-correlation (NCC), is a similarity measure that has been used successfully in many single-modal and multi-modal applications that require image registration or segmentation [Du et al., 2010]. Since NCC is just a statistical measure of the linearity of the relationship between two random variables, it can also be applied to quantities other than the intensities, such as geometrical features, local phase, etc as has been done in [Van den Elsen et al., 1995]. It is also very easy to compute, and thus computationally very efficient and this has allowed researchers to adapt it to devise a more local measure by calculating NCC between smaller image patches instead of the whole image at a time. In this way the registration can be driven by a cost function that is more locally adaptive. This has been implemented in [Avants et al., 2008a] in a

diffeomorphic transformation framework.

In many cases, the assumption of a linear or affine relationship between intensities in the two images does not hold, and only a statistical relationship can be assumed – for instance in the case of CT and MRI images. The use of statistical point-based similarity measures became popular after the publication of [Woods et al., 1993] and [Hill et al., 1994], and the measure of joint entropy has been used as a similarity measure for non-rigid registration in a reasonably successful manner in [Studholme et al., 1999]. However, one drawback of using joint entropy as a similarity measure, is that in some situations, it forces a transformation that reduces the overlap between corresponding structures in two images.

The minimization of the Kullback-Leibler (KL) divergence between expected and observed joint distributions is also a similarity measure that has been used successfully in medical image registration, as shown in [Chung et al., 2002] and [Chan et al., 2003]. However, the accuracy of the registration achieved by using this method depends largely upon the size of the training data that was used and how similar the images in the training data are to the images that are sought to be registered and in practice large, aligned data sets are very rarely available.

Another similarity measure that has been used very successfully in medical image registration is mutual information (MI). Although MI is not derived directly from any particular image formation model, it has been used widely as a similarity measure for the registration of multi-modal images when their intensities share a statistical relationship with each other, as shown in [Viola and Wells III, 1997] and [Collignon et al., 1995]. A survey of most of the registration techniques that make use of this similarity measure is presented in [Pluim et al., 2003]. However, a more commonly used variant of mutual information is Normalized Mutual Information (NMI), which is obtained by dividing MI by the entropy of the target image. In this case, the best value of the similarity

measure becomes 1 and the worst value becomes 0. NMI has been used successfully as a similarity measure in a B-Splines deformation framework in [Rueckert et al., 1998] and subsequently in [Rueckert et al., 2002].

In Section 2.4, we discussed different methods for PK modeling of DCE-MRI data, and in this section, we have examined the general principles and methods involved in image registration. In the next section we will review the existing approaches for simultaneous motion correction and PK analysis of DCE-MRI data.

2.6 Existing Methods for Motion Correction and PK Analysis

The earliest attempt to correct for motion in DCE-MRI scans using a PK model-based non-rigid registration technique was in [Hayton et al., 1997]. This registration technique was formalized in an optic flow deformation framework for 2D images, using the Laplacian of the deformation field as the regularizer. The cost function was taken to be the curve-fitting error obtained after fitting the Hayton-Brady model shown in Equation 2.17, to the CTC obtained from each voxel. In this paper, the estimation of the model parameters and the process of image registration are two separate processes, and both are essentially based on curve-fitting. More recently, Buonaccorsi and colleagues have implemented a model-based registration technique [Buonaccorsi et al., 2005] that has a very similar framework to [Hayton et al., 1997], but that uses the Tofts model, or the extended Kety model to drive the registration cost-function. Both [Hayton et al., 1997] and [Buonaccorsi et al., 2005] have motion correction algorithms that consist of the following basic steps (although they differ in the details of how to perform each step):

- Step 1: Fit the chosen PK model to the original motion-corrupted data from the DCE-MRI scan.
- Step 2: Deform the images corresponding to each time-point in a manner that improves the model fit.
- Step 3: Refit the model to the new set of deformed images.
- Step 4: Repeat steps 2 and 3 until convergence, or for a fixed number of iterations.

In [Buonaccorsi et al., 2005], a simplex optimization scheme is used to perform Step 1, whereas [Hayton et al., 1997] uses the Levenberg-Marquardt optimization scheme. One of the major limitations of [Buonaccorsi et al., 2005] is that it only considers translations (in 3D) in its deformation model in order to make the whole process computationally tractable. However, in [Buonaccorsi et al., 2005] the authors demonstrate the superiority of their model-based registration technique over a simple motion correction technique based on registering each image to the mean intensity image.

There have also been attempts to co-register the images using standard single-modal or multi-modal similarity functions such as Mutual Information [Zöllner et al., 2009] and then assessing the improvement in how well the data fits some suitable PK-model. The study in [Zöllner et al., 2009] uses a standard B-Splines deformation framework and Mutual Information as a cost function to register individual frames from a DCE-MRI acquisition of kidneys. The method is compared with a symmetric registration algorithm using a non-parameteric deformation model and an intensity-based similarity measure, but there were no quantitative results to show which one performed better. Along with registration of the images, the different compartments of the kidney are segmented using k-means clustering on the voxel time-courses. The metric used for the k-means clustering was the cosine distance function, and the clustering was performed ten times to negate the effect of the random initialization. It was found that the

method is quite successful in distinguishing the different compartments of the kidney (cortex, medulla and pelvis), and that the registration causes a significant improvement in the segmentation using the time courses. However, as we mentioned in Section 2.5.3, Mutual Information assumes a statistical relationship between the two images being registered, and this relationship is bound to change due to the large intensity variations caused by the presence of a contrast agent.

Instead of performing motion correction, it is also possible to extract accurate PK parameters by fitting models that take motion and other effects into account during the optimization process itself. One such approach has been the use of Bayesian hierarchical models to incorporate fixed effects (effect of treatment) and random effects (variations across individual patients and across voxels within a patient) into one unified framework, and then use this to make clinical predictions [Schmid et al., 2009a]. This paper uses a bi-exponential function described in [Fritz-Hansen et al., 1996] as the Arterial Input Function, and the extended-Kety model as the PK model. In [Schmid et al., 2009a], the authors designate each model parameter to be a linear combination of the fixed and random effects so that it is possible to isolate the effect of one or more of these factors when examining the change in the values of the model parameters. In this way the authors are able to distinguish responders from non-responders after taking differences between patients and the heterogeneity of the tumour into account. The parameters are calculated in a fully Bayesian manner by sampling from the full posterior distributions using MCMC sampling. Thus, the process of parameter estimation is very time-consuming, and takes about 2 days.

The discussion in this section has been focussed on different methods for extracting accurate PK parameters from DCE-MRI data by taking the effects of motion into account. The next step is to convert the PK parameters into clinically relevant information. As we discussed in Section 2.4, PK parameters are not directly relatable to tumour

characteristics such as microvessel density, oxygenation, etc due to the difference in spatial scales. However, there are many studies that demonstrate the relationship between PK parameters and clinical outcomes. Some of these studies will be discussed in the following section.

2.7 Linking Image Analysis to Clinical Predictions

The ultimate aim of analyzing DCE-MRI scans and estimating semi-quantitative and quantitative measures based on the scans is to use the information contained in them for the process of predicting clinical outcomes, and for designing treatment protocols. The efficacy of DCE-MRI in predicting biological or clinical outcomes can be tested in two ways:

1. By comparing the DCE-MRI parameters with histological data collected after resection of the tumour.

2. By examining the correlation between DCE-MRI parameters and other clinical outcomes such as disease-free survival.

Various studies have been conducted to show that semi-quantitative measures from DCE-MRI data correlate with histopathological data, and the methodology followed in these studies is to compare the parameters estimated from DCE-MRI data with the histological studies performed after resection of the tumour. In the case of rectal cancer, [Zhang et al., 2008] found that the time to peak enhancement was negatively correlated with microvessel density (MVD), as well and with the expression of the vascular endothelial growth factor (VEGF). In effect, this demonstrated that tissues showing signs of angiogenesis also showed early enhancement due to contrast agents in a DCE-MRI scan. While looking at other semi-quantitative measures calculated from

DCE-MRI data, it was found in [Loncaster et al., 2002] that the relative signal intensity (RSI) and the amplitude, both correlate well with the median partial-pressure of oxygen (pO2) as measured from the histology data. Similarly, in [Cooper et al., 2000], the maximum signal intensity was found to be strongly correlated with pO2, while the rate of enhancement was found to be weakly correlated with pO2.

As we discussed in Section 2.1.1, the growth of a tumour requires the proliferation of new blood vessels, which in turn is dependent on the up-regulation of VEGF. However, as far as PK model-based parameters are concerned, there are conflicting reports about the correlations betweeen K^{trans} and histological markers such as MVD or VEGF. [Atkin et al., 2006] showed that there was no statistically significant correlation between K^{trans} and VEGF markers, and a small negative correlation between K^{trans} and MVD. Since K^{trans} is a measure of blood flow and perfusion, and MVD is a measure of new blood vessels, this negative correlation is surprising and paradoxical.

In another study on colorectal cancer, it has been shown that K^{trans} had a significant ($p = 0.01$) positive correlation with the pre-treatment serum VEGF [George et al., 2001]. In [Hawighorst et al., 1998] as well, it was also shown that the amplitude of the signal-intensity curve (averaged over the tumour ROI), as well as the transfer coefficient K^{trans} are positively correlated with MVD for uterine cervical tumours. On the other hand [Hawighorst et al., 1998] also found that histological markers for VEGF did not correlate with amplitude or K^{trans}.

Although MVD and VEGF markers are indicative of the physiological state of a tumour, and provide information about its growth, they are not sufficient as measures of prognosis or future survival of the patient. The tumour grade, and of course data on patient survival, are more accurate benchmarks for prognosis and the nature of the disease. In [Hawighorst et al., 1998] it was found that median K^{trans} was a better indicator of disease-free survival than histological markers such as HMVD or VEGF expression.

This indicated the importance of building confidence in non-invasive imaging biomarkers, which could perhaps be more reliable than histological samples collected through surgery. As far as the relationship between PK parameters and tumour grades is concerned, [Patankar et al., 2005] showed that there is good agreement between K^{trans} and histological grades in gliomas. Subsequently, another study [Donaldson et al., 2010a], has also shown that the semi-quantitative value of the enhancing fraction (i.e. percentage of voxels within the tumour that show contrast-enhancement) of tumours is negatively correlated with disease-free survival.

An extensive review conducted in [Zahra et al., 2007] shows that that DCE-MRI can be used for characterizing the mass and stage of tumours, as well as for predicting response to therapy in a variety of tumour types. This paper also lays out many of the problems that are encountered while using DCE-MRI scans for making clinical predictions. The standard approach has been to use the properties of CTCs to estimate the physiological characteristics of the tumour and its neighboring areas. Although the principle area where this modality has been tested and has proven to be of clinical value is in breast cancer [Turnbull, 2009], recent studies such as [George et al., 2001] have also shown its value for predicting response to therapy in colorectal cancer.

In the course of treatment of most types of cancer (especially colorectal cancer), DCE-MRI data can be collected at several stages during treatment – before, during, or after CRT. Thus, it is important to decide which set of data are the best predictors of response. The correlation between radiological response (assessed according to WHO guidelines based on the length and thickness of the tumour) and the pre-therapy mean K^{trans} was found to be statistically significant in [George et al., 2001]. The study in [George et al., 2001] also showed that the difference in DCE-MRI parameter maps measured before and after CRT was correlated with radiological response. In [Gollub et al., 2012], the authors investigated the correlation between post-therapy

mean K^{trans} and pathological complete response (pCR), and found that the post-therapy mean K^{trans} was a better predictor than either pre-therapy measurements, or the difference between pre- and post-therapy. A study on rectal cancer patients also showed that the decrease in the slope of the contrast-enhancement curve before and after chemotherapy was a better predictor of overall response than the decrease in tumour volume [Dinter et al., 2009].

Recently, there have been some promising developments in the area of heterogeneity measurement and usage in prediction [Yang and Knopp, 2011], [Jackson et al., 2007]. These studies indicate that apart from semi-quantitative summary values, and mean or median K^{trans} measured before or during therapy, the spatial distribution of PK parameters can be of great prognostic value. In [Rose et al., 2007], the authors apply the concept of Renyi fractals to the parameter maps obtained after fitting a PK-model to the data and use this to quantify the amount of heterogeneity present in the tumour. This quantification of heterogeneity is then used to characterize tumours as high or low grade, and it is found that heterogeneity is more sensitive than summary statistics such as the median and standard deviation in distinguishing between different tumour grades. The studies on intra-tumoural heterogeneity reiterate the importance of performing accurate voxel-wise PK analysis of DCE-MRI data in order to make meaningful predictions about response to therapy.

Chapter 3

The MoRPE Framework

As demonstrated in Chapter 2, there have been many attempts to co-register images from a DCE-MRI sequence and to improve the estimation of the Contrast Take-up Curve (CTC) at each voxel. There have also been attempts to fit physiology-driven pharmacokinetic models to the CTCs and thus derive some clinically meaningful model-parameters [Schmid et al., 2009a]. These two above-mentioned objectives are combined into one unified framework in [Buonaccorsi et al., 2005] and [Hayton et al., 1997], as described in Chapter 1. However, while [Buonaccorsi et al., 2005] allows only simple translations, [Hayton et al., 1997] has been implemented for 2D images, and uses a PK model in which the parameters cannot be related directly to physiological properties of the tissue.

In this chapter, we will describe the details of our novel probabilistic, model-based registration and parameter estimation framework (MoRPE) for DCE-MRI sequences. To this end, we will first derive our probabilistic similarity measure, from an image

formation model based on the uptake of CA in the body.

3.1 The Mathematical Formulation

For the purpose of this section, let us assume that there are $i = 1, ..., t$ time points at which the MRI volumes are acquired during the dynamic scan, and that the DCE-MRI data that we are presented with is given as $Y = [Y_i]_{i=1,..,t}$. Let us further assume that the unknown ground-truth image is $X = [X_i]_{i=1,..,t}$ such that the image Y_i is formed by deformations of X_i (due to patient-motion during the scan), and an addition of noise by some noise process. Our objective is to determine the transformation T_i for each time-point i, which has been implicitly applied (by motion) to the original ground-truth image X_i in order to generate the observed image (raw data) Y_i. In mathematical terms, we assume that:

$$Y_i = T(X_i) + \epsilon_i \tag{3.1}$$

It has been shown previously that though noise in MRI scans is known to follow a Rician distribution, it is well approximated by a Gaussian model for moderate signal-to-noise ratios [Schmid et al., 2009b]. Thus, in our study, ϵ_i is assumed to be a Gaussian noise process with mean 0 and standard deviation σ_i. In this report, we have assumed that the standard deviation for the noise-process is constant for all time-points, i.e. $\forall i,\ \epsilon_i \sim N(0, \sigma)$.

The key assumption that we make in order to define our image formation model, is that the 'true' uncorrupted image X_i (for each time-point $i = 1, ..., t$) is a function of some physiological characteristics of the tissue. Thus, we assume that each voxel $j = 1, ..., N$ in the image, has some physiological characteristics denoted by the parameter θ_j. Further, the intensity of voxel j at time-point i can be written as $f_i(\theta_j)$, and the entire ground-truth image X_i can be expressed as $f_i(\theta)$, where $\theta = [\theta_j]_{j=1,..,N}$. Thus our

image formation model can be expressed as:

$$Y_i = T_i\left(f_i(\theta)\right) + \epsilon \;\forall i = 1, .., t \tag{3.2}$$

Here, the MRI intensity of a voxel in the target image is not assumed to be a function of the *intensity* of the same voxel in the source image, but rather, a function of the *physiological characteristics* of that voxel in the tissue. Since we want to correct for motion in the entire time-series Y, and also calculate the parameter vector θ on the basis of the entire dataset, we can rewrite Equation (3.2) as:

$$Y = T\left(f(\theta)\right) + \epsilon \tag{3.3}$$

Here, $f(\theta) = \left[f_1\left(\theta\right), f_2\left(\theta\right), ..., f_t\left(\theta\right)\right]$ forms a 4D array of the same size as the original dataset $Y = [Y_1, Y_2, .., Y_t]$. In this scenario, the likelihood term, or the probability that we observe a certain image Y can be given by:

$$\begin{aligned} P\left(Y \mid X_0, T, \theta, \sigma\right) &= P\left(\epsilon = T\left(f(\theta)\right) - Y\right) & (3.4)\\ &= \frac{1}{\left(2\pi\sigma^2\right)^{N/2}} \exp\left(\frac{-\left\|T\left(f(\theta)\right) - Y\right\|^2}{2\sigma^2}\right) & (3.5) \end{aligned}$$

Here X_0 is the pre-contrast image (which is known to us), σ is the unknown noise parameter, and the norm $\left\|T\left(f(\theta)\right) - Y\right\|$ is the root-mean-squared-error, calculated across all points in the 4D dataset. Now, unlike the cases discussed in Section (2.5.3), here we have two simultaneous objectives:

1. To correct for motion in the image sequence and accurately estimate the ground-truth image X_i as well as the deformation T_i for each time-point i.

2. To estimate the set of model parameters $\theta = \left[\theta_j\right]_{j=1,..,N}$ (where N is the total

number of voxels) that predicts the data $X = [X_i]_{i=1,..,t}$ (where t is the total number of time points) most accurately.

Thus, instead of the posterior probability of just the transformations T_i, we will look at the joint posterior probability of θ and T.

$$\begin{aligned} P(\theta, T \mid X_0, Y, \sigma) &= \frac{P(\theta,T,X_0,Y,\sigma)}{P(X_0,Y,\sigma)} \\ &= P\left(Y \mid X_0, T, \theta, \sigma\right) \frac{P(X_0,T,\theta,\sigma)}{P(X_0,Y,\sigma)} \end{aligned} \tag{3.6}$$

Since there is no *a-priori* reason for the intensities in the pre-contrast image (X_0), the patient-motion (T) during the scan, the physiological properties (θ) of the tissue and the amount of noise (σ) in the image to be related to each other, we assume them to be independent of each other, and write the prior term as:

$$P\left(X_0, T, \theta, \sigma\right) = P\left(X_0\right) \cdot P\left(T\right) \cdot P\left(\theta\right) \cdot P\left(\sigma\right) \tag{3.7}$$

Since X_0, Y and σ remain fixed during the iterations of the algorithm, we can combine Equations (3.6) and (3.7), to infer that:

$$P(\theta, T \mid X_0, Y, \sigma) \propto P\left(Y \mid X_0, T, \theta, \sigma\right) P(\theta) P(T) \tag{3.8}$$

The terms $P(\theta)$ and $P(T)$ reflect the prior probability distributions of θ and T. In this chapter, we assume a uniform distribution on k_{ep} and v_e (between their physiologically relevant bounds – [0, 1]), and constrain the transformations to be a smooth, reversible diffusion process (Section 3.4). However, the generic framework of MoRPE allows us to incorporate different priors for any of these terms when required.

To facilitate the accurate computations with floating point precisions, and to ensure that the optimization process remains robust to noise and to the choice of σ, we actually

seek to minimize the negative log of the joint posterior probability rather than maximizing the probability itself. Thus, we now aim to estimate the set of transformations $\hat{T}$, and the PK parameters $\hat{\theta}$ that will minimize $L = -\log\left(P(\theta, T \mid X_0, Y, \sigma)\right)$:

$$[\hat{\theta}, \hat{T}] = \arg\min_{\boldsymbol{\theta}, \boldsymbol{T}} -\log\left(P(\theta, T \mid X_0, Y, \sigma)\right) \tag{3.9}$$

In order to calculate the optimal values of the transformation and PK parameters ($[\hat{\theta}, \hat{T}]$), we need to use some kind of robust optimization scheme. The simplest kind of optimization method – gradient descent – requires only the first-order derivatives with respect to θ and T. In our implementation, we have employed the Levenberg-Marquardt optimization method, which also requires only the first-order derivatives to be calculated, but that converges much faster than gradient descent [Marquardt, 1963]. In our experiments, we found the second-derivatives of the cost-function to be numerically unstable, and hence the Newton method or other methods requiring the second-derivative were not used. Following from the above discussion, the basic steps of our algorithm are as follows:

1. Choose initial values for θ and T and calculate $L = -\log\left(P(\theta, T \mid X_0, Y, \sigma)\right)$

2. Calculate $\frac{\partial L}{\partial T}$, $\frac{\partial L}{\partial \theta}$ using the current estimates of θ and T, and update the values of θ and T (simultaneously) in order to minimize L.

3. Repeat Step 2 using the new estimates of θ, T and L until convergence or until the updates become smaller than a suitable threshold value.

In Step 1, the initial value of θ is chosen by fitting a curve through the initial data for each voxel, and T is initialized as the identity transformation. The standard deviation σ, is representative of the noise present in the dataset being analyzed. Thus, it is calculated from the data, as the standard deviation of the error between the data and the

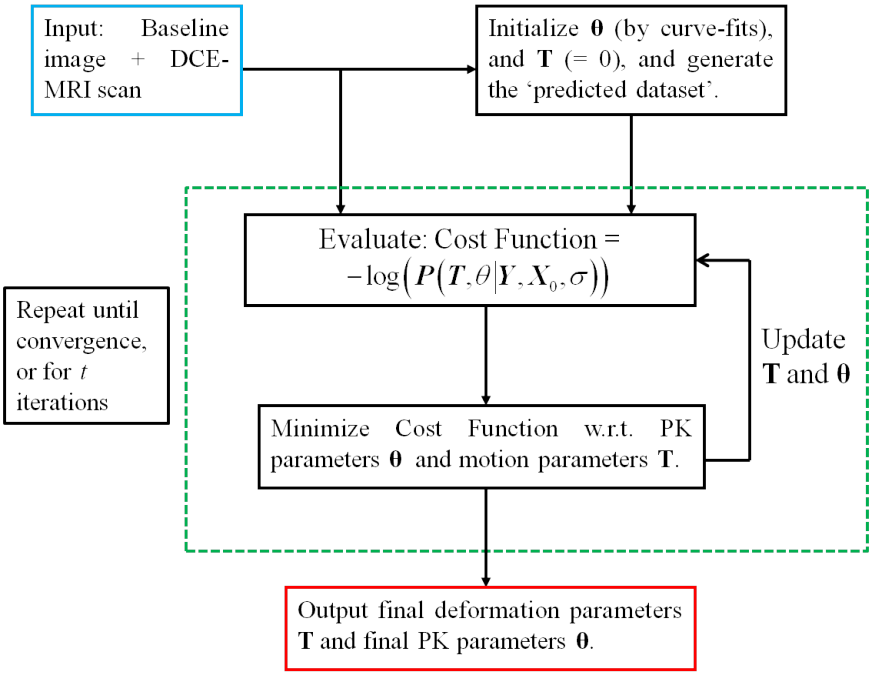

Figure 3.1: The schematic framework of the MoRPE algorithm.

inital curve-fits (before motion correction). However, since σ only scales the gradients with respect to the cost function, it can also be modified to suit the requirements of speed and precision. The main step in this procedure, as shown in Figure 3.1, is to calculate the analytic derivative of the joint posterior probability with respect to the transformation parameters and the PK-parameters. It is important to remember that the derivatives are calculated using *all* the images together in each iteration, and then used in a Levenberg-Marquardt optimization framework, as explained in Section (3.4).

We will now describe in detail the image formation model that we have used – i.e. the exact nature of f in Equation (3.3). We will also describe how we calculate derivatives of the posterior joint probability with respect to θ and T.

3.2 Choice of the Pharmacokinetic Model

In our current implementation, we have chosen the standard Tofts model as the pharmacokinetic model that describes the uptake of contrast agent. Thus we assume that the total concentration of CA in the tissue at time t, $C_T(t)$ is given by:

$$C_T(t) = k_{ep} \cdot v_e \cdot e^{-k_{ep} \cdot t} \otimes C_p(t) \tag{3.10}$$

Here, k_{ep} represents the transfer coefficient for transfer of CA from the EES space to the blood plasma, and v_e denotes the proportion of space occupied by the EES (with $K^{trans} = k_{ep} \cdot v_e$). Further, $C_p(t)$ represents the AIF, and $\otimes$ denotes a convolution. We have chosen this because it is the simplest and most widely accepted PK model that has physiologically relevant model parameters. However, it is important to remember that the PK model to be used is only one component of our algorithm, and it could be replaced with any other PK model if needed. For $C_p(t)$, we have chosen to test two different population-averaged AIFs that have been used previously in the literature: the Orton [Orton et al., 2008] and Weinmann [Weinmann et al., 1984] AIFs shown in Figure 3.2. The general form of the bi-exponential AIF, of which the Winmann AIF is one example, is given by:

$$C_p(t) = D\left(A_1 e^{-m_1 t} + A_2 e^{-m_2 t}\right) \tag{3.11}$$

Here D stands for the dosage (in mM l^{-1}) of the contrast agent being administered to the patient. In the original bi-exponential model, the first amplitude (A_1) and rate constant (m_1) reflect the leakage of blood plasma into EES, and the second amplitude (A_2) and rate-constant (m_2) reflect the transfer of blood plasma into the kidneys. These four parameters are determined by fitting a curve to population-averaged data, and for the Weinmann study [Weinmann et al., 1984], A_1 was found to be 3.99 kg/litre, m_1 was

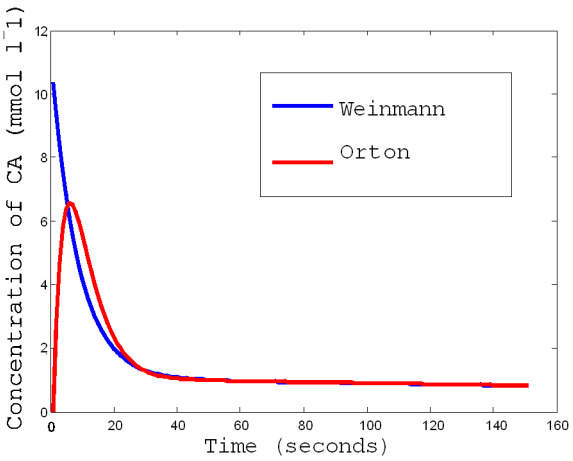

Figure 3.2: The Orton and Weinmann arterial input functions for humans.

0.144 min^{-1}, A_2 was 4.78 kg/litre and m_2 was 0.0111 min^{-1}.

In the Orton AIF model [Orton et al., 2008], there is an added factor that simulates the bolus arrival in a more realistic way. Thus, the form of the AIF becomes:

$$C_p(t) = A_1 t e^{-m_2 t} + A_2 \left(e^{-m_1 t} + e^{-m_2 t}\right) \tag{3.12}$$

For the Orton AIF (Model 2 in [Orton et al., 2008]), A_1 was found to be 323 mM min^{-1}, A_2 was found to be 1.07 mM, m_1 was found to be 0.172 min^{-1}, and m_2 was found to be 20.2 min^{-1}.

There are some disadvantages of using a simple bi-exponential AIF, such as the fact that it is strictly only applicable when the plasma fraction (v_p) is negligible. When this is not the case, the use of such an AIF tends to overestimate the transfer constant K^{trans} [Barboriak et al., 2008]. Another disadvantage is that this simplistic model cannot capture the second pass of the contrast agent. Although there are several other population-averaged AIFs (notably the Parker model [Parker et al., 2006]), the main advantage of using the bi-exponential AIFs, is that they both have analytic expressions for the convolution in Equation (3.10). This makes the framework computationally more efficient, but it is important to note that the AIF can easily be replaced by an-

other model when required.

In this thesis, we have chosen to neglect terms involving the vascular proportion v_p since earlier studies indicate that this is negligible for colorectal tumours [Gollub et al., 2012], [Bradley et al., 2008]. Consequently, using the arterial input function proposed by Weinmann and colleagues [Weinmann et al., 1984], the expression for $C_T(t)$ is given by:

$$C_T(t) = Dk_{ep}v_e\left\{3.99\left(\frac{e^{-k_{ep}(t-t_0)} - e^{-0.144(t-t_0)}}{0.144 - k_{ep}}\right) + 4.78\left(\frac{e^{-k_{ep}(t-t_0)} - e^{-0.0111(t-t_0)}}{0.0111 - k_{ep}}\right)\right\} \tag{3.13}$$

Similarly, if we use the Orton AIF (model 2 in [Orton et al., 2008]), and neglect the vascular fraction, then the expression becomes:

$$\begin{aligned} C_T(t) &= \frac{323Dk_{ep}v_e}{k_{ep} - 20.2}\left\{(t-t_0)e^{-20.2(t-t_0)} + \frac{e^{-20.2(t-t_0)} - e^{-k_{ep}(t-t_0)}}{20.2 - k_{ep}}\right\} \\ &\quad + 1.07Dk_{ep}v_e\left\{\frac{e^{-k_{ep}(t-t_0)} - e^{-0.172(t-t_0)}}{0.172 - k_{ep}} + \frac{e^{-k_{ep}(t-t_0)} - e^{-20.2(t-t_0)}}{20.2 - k_{ep}}\right\} \end{aligned} \tag{3.14}$$

Here D denotes the dosage (in mM l^{-1}) of the CA administered to the patient, and t_0 has been used to denote the 'onset time', or the time of arrival of the contrast-agent (in min) in the voxel. In both of the equations above it is implicitly assumed that $C_T(t) = 0$ for $t < t_0$.

It is important to remember that the PK model only provides us with a functional relationship between the physiological characteristics of the tissue and the concentration of CA present in the same tissue. The PK model does not give us a direct relationship between the model parameters and the intensity observed in the same voxel during a DCE-MRI scan. As shown in Chapter (1), the presence of CA alters the MRI signal by changing the T_1- and T_2-relaxation times of the tissue into which it is injected, in the manner described in Equations 2.5 and 2.6. Using these two equations, along with

Equation 2.4, the MRI signal enhancement observed at time t – i.e. X_t/X_0 – can be written as:

$$\frac{X_t}{X_0} = \exp\left(-r_2 C_T(t) TE\right) \frac{1 - \exp\left(-P - Q\right) - \cos\left(\alpha\right)\left(\exp\left(-P\right) - \exp\left(-P - 2Q\right)\right)}{1 - \exp\left(-P\right) - \cos\left(\alpha\right)\left(\exp\left(-P - Q\right) - \exp\left(-P - 2Q\right)\right)} \tag{3.15}$$

Here $P = TR/T_{10}$, $Q = r_1 C_T(t) TR$, X_t denotes the MR signal intensity at time t and X_0 denotes the signal intensity at time 0, before CA is injected. Since all terms apart from $C_T(t)$ remain constant for the duration of the DCE-MRI scan, we can write the signal at time t as a function of the concentration of CA at time t, i.e. $X_t = f\left(C_T(t)\right)$.

3.3 Derivatives of the Posterior Probability

In order to estimate the transformation T and the model parameters θ for that the negative log probability, $L = -\log\left(P(\theta, T \mid X_0, Y, \sigma)\right)$, attains its minimum (Equation 3.9), we need to calculate the derivatives of L with respect to the transformation parameters and with respect to the model parameters. In other words, we need to calculate $\frac{\partial}{\partial T}\left(-\log\left(P(\theta, T \mid X_0, Y, \sigma)\right)\right)$ and $\frac{\partial}{\partial \theta}\left(-\log\left(P(\theta, T \mid X_0, Y, \sigma)\right)\right)$. Clearly, the calculation of these terms depends upon the transformation model and the pharmacokinetic model that we choose. We will first describe how to calculate the derivative with respect to the transformation parameters, and then with respect to the model parameters.

3.3.1 Derivatives with respect to T

In order to explain the process of calculating $\frac{\partial}{\partial T}\left(-\log\left(P(\theta, T \mid X_0, Y, \sigma)\right)\right)$, we will illustrate the general principles for the simplest kind of non-parameteric transformation model, where each voxel has a deformation vector assigned to it. As noted in Section

2.5.2, for $3D$ image sequences, the deformation field T can be written as $[U, V, W]$, where U is the array denoting the displacements along the x-axis, V is the array denoting displacements along the y-axis, and W is the array denoting displacements along the z-axis. Thus, if I_o is the original image and $I_d = T(I_o)$ is the deformed image, then the original intensity at position (x, y, z) is denoted by $I_o(x, y, z)$ and the intensity of same position in the new image is given by $I_d(x, y, z) = I_o(x + u, y + v, z + w)$. In this scenario, the first-order Taylor-series expansion of the transformed image is given by:

$$T(I) = I + U \cdot \frac{\partial I}{\partial x} + V \cdot \frac{\partial I}{\partial y} + W \cdot \frac{\partial I}{\partial z} \tag{3.16}$$

Here, U, V, W, $\frac{\partial I}{\partial x}$, $\frac{\partial I}{\partial y}$ and $\frac{\partial I}{\partial z}$ are all arrays the same size as the image, and the product $(\cdot)$ denotes the element-wise product of two arrays. The first-order Taylor series approximation is reasonably good if the deformation is small, and also allows us to differentiate the transformation with respect to the parameters U, V and W in a simple manner. We use this approximation for only one iteration of the registration method, to ensure that the deformations are small. The exact mechanism for applying the deformations will be discussed in detail in Section (3.4). Thus, I is updated to be the transformed image at the end of each iteration. In this scenario, during one iteration, $\frac{\partial T(I)}{\partial U} = \frac{\partial I}{\partial x}$, $\frac{\partial T(I)}{\partial V} = \frac{\partial I}{\partial y}$, and $\frac{\partial T(I)}{\partial W} = \frac{\partial I}{\partial z}$.

The derivative of the cost function with respect to each component of the deformation field at a particular time point i is taken to be $\frac{\partial}{\partial U_i}(-\log P(\theta, T \mid X_0, Y, \sigma))$, and similarly for V and W. Now in order to calculate the gradients of the negative log joint posterior probability term with respect to U, V and W, we can use Equation (3.8) and mutual independence of the variables X_0, θ and σ to deduce that:

$$\frac{\partial}{\partial U}(-\log P(\theta, T \mid X_0, Y, \sigma)) \propto \frac{\partial}{\partial U}(-\log(P(Y \mid X_0, T, \theta, \sigma) P(\theta) P(T))) \tag{3.17}$$

Here $P(T)$ represents the prior probability of the deformation, or a measure of smoothness of the deformation field. This can be chosen to suit each particular application, and it will be discussed in detail in Section (3.4). In order to calculate the derivative of $-\log P\left(Y \mid X_0, T, \theta, \sigma\right)$ with respect to U, we have, from Equation (3.5):

$$\frac{\partial}{\partial U}\left(-\log P\left(Y \mid X_0, T, \theta, \sigma\right)\right) \tag{3.18}$$

$$= \frac{\partial}{\partial U}\left(-\log\left(\frac{1}{\left(2\pi\sigma^2\right)^{N/2}}\exp\left(\frac{-\left\|T\left(f(\theta)\right)-Y\right\|^2}{2\sigma^2}\right)\right)\right) \tag{3.19}$$

$$= \frac{\left(T\left(f(\theta)\right)-Y\right)}{\sigma^2}\cdot\frac{\partial T\left(f(\theta)\right)}{\partial U} \tag{3.20}$$

$$= \frac{\left(T\left(f(\theta)\right)-Y\right)}{\sigma^2}\cdot\frac{\partial f(\theta)}{\partial x} \tag{3.21}$$

Equations 3.18 – 3.21 explain how derivatives of the negative log likelihood term, with respect to the deformation, are calculated in general terms. As we shall see in Section (3.4), the choice of the deformation framework and the regularization method are dependent on the specific application, and can have an important impact on the final results.

3.3.2 Derivatives with respect to θ

The calculation of derivatives with respect to the model parameters is somewhat more complicated because of the complex relationship between the PK parameters, θ, and the final intensities in the image, $f(\theta)$. In our case, the model parameters are k_{ep}, v_e and $K^{trans} = k_{ep} \cdot v_e$. Thus, we need to calculate the derivative of the joint posterior probability with respect to these parameters. In order to illustrate the calculations of the derivatives, we will show them for the Tofts model with the Weinmann AIF, and use t' to denote the time after contrast uptake (instead of $t - t_0$). In this scenario,

Equation (3.13) can be re-written as:

$$C_T(t') = Dk_{ep}v_e \left\{ 3.99 \left(\frac{e^{-k_{ep}t'} - e^{-0.144t'}}{0.144 - k_{ep}} \right) + 4.78 \left(\frac{e^{-k_{ep}t'} - e^{-0.0111t'}}{0.0111 - k_{ep}} \right) \right\} \tag{3.22}$$

In this form, the shape of the contrast-uptake curve is governed only by k_{ep}, while the amplitude or height of the curve is governed both by k_{ep} and v_e. The derivatives of the concentration of CA with respect to k_{ep} can be derived as:

$$\begin{aligned} \frac{\partial C_T(t')}{\partial k_{ep}} &= \frac{1}{k_{ep}} C_T(t') + (3.99) Dk_{ep}v_e \left(\frac{t' \left(k_{ep} - 0.144\right) e^{-k_{ep}t'} + e^{-k_{ep}t'} - e^{-0.144t'}}{\left(0.144 - k_{ep}\right)^2} \right) \\ &\quad + (4.78) Dk_{ep}v_e \left(\frac{t' \left(k_{ep} - 0.0111\right) e^{-k_{ep}t'} + e^{-k_{ep}t'} - e^{-0.0111t'}}{\left(0.0111 - k_{ep}\right)^2} \right) \end{aligned} \tag{3.23}$$

The derivative of $C_T(t')$ with respect to v_e is simply given by:

$$\frac{\partial C_T(t')}{\partial v_e} = \frac{1}{v_e} C_T(t') \tag{3.24}$$

Now, as shown in Equation (3.15), the relationship between $C_T(t)$ and the final intensity is quite complex. But of all the terms in that equation, only $C_T(t)$ depends on the model parameters k_{ep} and v_e. The other terms can all be treated as constants. Thus, we can differentiate X_t from equation (3.15) to determine $\frac{\partial X_t}{\partial k_{ep}}$ and $\frac{\partial X_t}{\partial v_e}$ in terms of all the other constants that are known. In order to calculate the derivative of the log-likelihood with respect to the onset time t_0, we employ numerical derivatives. The derivative of X_t with respect to t_0 can be simply given as:

$$\frac{\partial X_t}{\partial t_0} = \frac{X_{t-t_0-1} - X_{t-t_0+1}}{2} \tag{3.25}$$

Here the time is measured in seconds, and therefore the units of $\frac{\partial X_t}{\partial t_0}$ are $\sec^{-1}$. This quantity can be calculated easily using Equation 3.15, since $X_t = f\left(C_T(t)\right)$. Once

we have calculated all the partial derivatives, we can proceed to optimize the objective function $-\log P(\theta, T \mid X_0, Y, \sigma)$ over the space of model and transformation parameters.

3.4 Choice of the Deformation Framework

As we had discussed in Section (2.5.2), there are numerous deformation models and regularization frameworks presently being used in medical imaging. The basic purpose of a deformation framework, is to apply some contraints upon the mapping between the source and target images. The nature of constraints that are to be applied depends on the biological application, and the *a-priori* knowledge about the motion being simulated.

From equation (3.8), it is evident that the negative log posterior probability can be written as:

$$-\log P(\theta, T \mid X_0, Y, \sigma) \propto -\log\left(P\left(Y \mid X_0, T, \theta, \sigma\right) P(\theta) P(T)\right) \tag{3.26}$$

$$= -\frac{\|T\left(f(\theta)\right) - Y\|^2}{\sigma^2} - \log\left(P(T)\right) \tag{3.27}$$

The form of Equation (3.27) is exactly equivalent to the general form of the cost function for a registration problem, as expressed in [Vercauteren et al., 2007], [Vercauteren et al., 2008], [Thirion, 1998] as well as in Equation (2.18) in the previous Chapter. If the basic objective is the registration of source image X to target image Y, then the problem can be described as the search for an optimal transformation T that minimizes the energy function $E(T)$, i.e:

$$\min_T E\left(T\right) = S\left(Y, T\left(X\right)\right) + \alpha R\left(T\right) \tag{3.28}$$

From our probabilistic formulation of the cost function, it is clear that the cost term

$S(Y, T(X))$ is equivalent to $\left(\frac{-\|T(f(\theta))-Y\|^2}{\sigma^2}\right)$ from Equation (3.18), whereas the regularization term $R(T)$ is equivalent to $(-\log(P(T)))$ from Equation (3.27).

As mentioned earlier, $P(T)$ represents the prior probability of the deformation, which is usually a measure of smoothness of the deformation field. In [Thirion, 1998], it was shown that it is very natural to visualize image matching as a diffusion process. It has also been shown earlier that constraining the deformation to be a diffusion process is equivalent to a Gaussian smoothing of the deformation field [Weickert, 1997], [Koenderink, 1984]. Thus, in our framework, we have chosen to use the Demons framework, with the diffusion-type regularization. In other words, the prior probability on the transformation in our model assumes that the deformation is a diffusion process that does not allow any folding in the tissue. However, our framework allows the user to incorporate other deformation models (priors on T) depending upon the application at hand.

It was shown in [Vercauteren et al., 2007], that a diffusion or fluid type of regularization could be realized by Gaussian smoothing of the deformation or update field, respectively. Whether the Gaussian smoothing results in a diffusion type, or fluid type of regularization depends upon whether only the update is smoothed at each iteration, or whether the whole deformation field is smoothed at each iteration. The Demons method decouples the process of calculating updates (derivatives of the similarity term) and the process of calculating the regularization term, and provides a computationally efficient mechanism for enforcing contraints upon a deformation. Due to the computational efficiency and simplicity in implementation, we chose to integrate our probabilistic similarity measure with different variants of the Demons algorithm to test our methods efficacy in motion correction and parameter estimation from DCE-MRI sequences.

The Demons algorithm, as presented in [Thirion, 1998], offers an efficient method for diffusion type regularization on a deformation field. According to the basic Demons

equation, for a source image X, and target image Y, the displacement vector $\boldsymbol{d}$ is given by:

$$\boldsymbol{d} = \frac{(X - Y)\,\nabla X}{(\nabla X)^2 + (X - Y)^2} \tag{3.29}$$

Here ∇X represents the vector of image gradients of X in each dimension, and $(X)^2$ represents the element-wise square of the matrix X (i.e. $(X)^2$ is a matrix of the same size as X). A more generalized form of the same equation was given in [Vercauteren et al., 2009]:

$$\boldsymbol{d} = \frac{(X - Y)\,\nabla X}{(\nabla X)^2 + \frac{\sigma_i^2}{\sigma_s^2}} \tag{3.30}$$

Here σ_i is an estimate of the local noise in the image, and σ_s if a factor that controls the step size. However, in both of the above cases, the deformations still form a vector space, and there is no restriction on the deformation field that make it bijective or invertible. In order to constrain the deformations to be invertible and smooth, we can optimize Equation 3.28 over the space of diffeomorphisms alone, instead of all deformations. Diffeomorphisms are technically defined as mappings that are infinitely differentiable at all points, and they do not form a vector space, but instead only a Lie group [Trouvé, 1998]. Thus, optimizing over the set of diffeomorphisms alone can be mathematically complex and computationally expensive.

One possible way of characterizing diffeomorphisms is to encode them using a time-varying velocity field, as shown in [Beg et al., 2005b]. Although this approach makes use of the entire group structure, it is computationally expensive. Instead, diffeomorphisms can be obtained by taking exponential mappings of stationary velocity fields, which form a Lie Algebra associated to the Lie group of diffeomorphisms, as shown in [Arsigny et al., 2006]. This has significant advantages in terms of mathematical sim-

plicity and computational time, since a simple scaling and squaring technique can be employed to approximate the exponential of the velocity field [Vercauteren et al., 2007]. In this scheme, the updates calculated through Equation 3.30 are treated as updates to the velocity field instead of the deformation field, and the deformation is calculated by composing the deformation field found from the previous iteration with the exponential of the update ($\exp(\boldsymbol{d})$).

Since the motion caused in the colorectal region during a DCE-MRI scan is primarily due to breathing and peristalsis, it is natural to assume that there is non-linear motion, but no change in topology when the anatomy is deformed. In other words, we would expect all deformations to be non-linear, but smooth and invertible. However, we wanted to test whether the normal Demons algorithm was sufficient, or whether the diffeomorphic Demons algorithm was needed to simulate deformations more accurately. For this reason, we tested our similarity measure with both deformation methods on simulated data. The steps of the regular Demons algorithm that we used are summarized in Algorithm 3.1.

Algorithm 3.1 The regular Demons MoRPE algorithm

Initialize $i = 0$ and $T_0 = I$ (Identity) transform.
Initialize θ_0 from curve-fits to raw data.
for i=1,...,n

1. Calculate $L = -\log\left(P(\theta_{i-1}, T_{i-1} \mid X_0, Y, \sigma)\right) \propto \|T_{i-1}\left(f(\theta_{i-1})\right) - Y\|^2 + \frac{\sigma_i^2}{\sigma_s^2}\|T_{i-1}\|^2$
2. Calculate $\frac{\partial L}{\partial \theta}$ and $\frac{\partial L}{\partial T}$ using the current estimate of L, θ and T.
3. Calculate $\theta_i = \theta_{i-1} - (H^{-1} + \lambda I) \cdot \delta\theta$, using $\frac{\partial L}{\partial \theta}$ and the Levenberg-Marquardt Method.
4. Calculate the update to the deformation field $u = \frac{(T_{i-1}(f(\theta_{i-1})) - Y)}{\left\|\frac{\partial L}{\partial T}\right\|^2 + \frac{\sigma_i^2}{\sigma_s^2}}\left[\frac{\partial L}{\partial T}\right]^T$.
5. Update $T_i = K_{diff} \otimes (T_{i-1} \circ u)$.

end loop

It is worth noting that $\frac{\partial L}{\partial \theta}$ calculated in Step 2 of the algorithm is $\left[\frac{\partial L}{\partial k_{ep}}, \frac{\partial L}{\partial v_e}, \frac{\partial L}{\partial t_0}\right]$, as shown in Section 3.3.2. For the calculation of the new estimate of θ in Step 3 of the algorithm, we use $\frac{\partial L}{\partial \theta}$ to calculate the pseudo-Hessian (H), which is defined as:

$$H = \begin{bmatrix} \frac{\partial L}{\partial k_{ep}} \cdot \frac{\partial L}{\partial k_{ep}} & \frac{\partial L}{\partial k_{ep}} \cdot \frac{\partial L}{\partial v_e} & \frac{\partial L}{\partial k_{ep}} \cdot \frac{\partial L}{\partial t_0} \\ \frac{\partial L}{\partial k_{ep}} \cdot \frac{\partial L}{\partial v_e} & \frac{\partial L}{\partial v_e} \cdot \frac{\partial L}{\partial v_e} & \frac{\partial L}{\partial v_e} \cdot \frac{\partial L}{\partial t_0} \\ \frac{\partial L}{\partial k_{ep}} \cdot \frac{\partial L}{\partial t_o} & \frac{\partial L}{\partial v_e} \cdot \frac{\partial L}{\partial t_o} & \frac{\partial L}{\partial t_0} \cdot \frac{\partial L}{\partial t_0} \end{bmatrix} \tag{3.31}$$

The pseudo-Hessian is then used in a Levenberg-Marquardt optimization scheme [Sun and Yuan, 2006], with $\lambda = 0.01$. It is also possible to calculate the actual Hessian, and use Newton's method to calculate the update. However, in our tests we found that the analytical calculations for the second-derivatives with respect to the PK parameters were not reliable in a large neighbourhood, and their estimates were numerically unstable. Since the Levenberg-Marquardt method, using the pseudo-Hessian, was found to be more robust and computationally efficient for the data available to us, we used this throughout our experiments.

In our implementation, we have used the diffusion-type regularization, which is achieved by applying a Gaussian kernel (K_{diff}) on the updated deformation field. In order to implement diffeomorphic Demons instead of the regular Demons algorithm, Step 5 of the algorithm is modified, and T_i is composed with the exponential of the update to the velocity field, instead of directly being composed with the update u calculated in Step 4. The MoRPE algorithm using diffeomorphic demons is summarized in Algorithm 3.2.

The exponential of the update to the velocity-field, $\exp(u)$, was calculated by the scaling and squaring technique as explained in [Vercauteren et al., 2007], with $N = 5$ time steps, i.e:

```
Calculate v_1 = 2^{-5} · u.
For i=2:5
    Perform recursive squaring v_i ← v_{i-1} ∘ v_{i-1}
end loop
```

In order to ensure that there is no folding in the tissue, the step sizes (v_i) that are composed together must be less than half a voxel. For the data that we were using, we found that five time steps were enough to ensure that $\max \|2^{-5} \cdot u\| \leq 0.5$. For both, the regular Demons and diffeomorphic Demons algorithms, we used the same Gaussian kernel K_{diff}, and the same values of σ_i and σ_s. It is worth noting that if we assume $\sigma_i = \|T(f(\theta)) - Y\|$ for every point, then the value chosen for σ_s controls the step-size, since $\|u\| \leq \sigma_s/2$ [Vercauteren et al., 2007]. Algorithms 3.1 and 3.2 only show the steps followed in one iteration of MoRPE. In our implementation, the entire procedure is performed in a multi-resolution framework, where the resolution is upsampled by a factor of two along each axis (a total factor of 8), and the estimates of T and θ from the previous iteration are upsampled and used as the starting estimates for the current resolution.

3.5 Tests on Simulated Data

The main objective of this study was to compare MoRPE with other motion correction algorithms using the same deformation framework, but different similarity measures. For this purpose, MoRPE was compared using the same framework with sum-of-squared differences (SSD), normalized mutual information (NMI as described in [Pluim et al., 2000]) and local normalized cross-correlation (NCC - as described in [Avants et al., 2008b]) as similarity measures that are based on the image intensities.

Algorithm 3.2 The diffeomorphic Demons MoRPE algorithm

Initialize $i = 0$ and $T_0 = I$ (Identity) transform.
Initialize θ_0 from curve-fits to raw data.
for i=1,...,n

1. Calculate $L = -\log\left(P(\theta_{i-1}, T_{i-1} \mid X_0, Y, \sigma)\right) \propto \|T_{i-1}\left(f(\theta_{i-1})\right) - Y\|^2 + \frac{\sigma_i^2}{\sigma_s^2}\|T_{i-1}\|^2$
2. Calculate $\frac{\partial L}{\partial \theta}$ and $\frac{\partial L}{\partial T}$ using the current estimate of L, θ and T.
3. Calculate $\theta_i = \theta_{i-1} - (H^{-1} + \lambda I) \cdot \delta\theta$, using $\frac{\partial L}{\partial \theta}$ and the Levenberg-Marquardt Method.
4. Calculate the update to the deformation field $u = \frac{(T_{i-1}(f(\theta_{i-1})) - Y)}{\left\|\frac{\partial L}{\partial T}\right\|^2 + \frac{\sigma_i^2}{\sigma_s^2}}\left[\frac{\partial L}{\partial T}\right]^T$.
5. Update $T_i = K_{diff} \otimes \left(T_{i-1} \circ \exp(u)\right)$.

end loop

In order to use these other similarity measures within our framework, we replaced the gradient with respect to the transformation T calculated in Step 2 of our algorithm with the gradient of the respective cost functions, and omitted the gradient with respect to the PK-parameters.

We generated eight simulated datasets by selecting the pre-contrast image volume from one DCE-MRI scan (Figure 3.3) and simulating contrast uptake using the Tofts model and four different artificial parameter maps for θ, using two different AIFs with

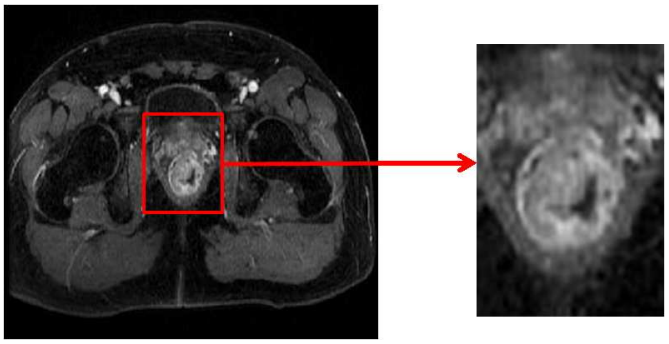

Figure 3.3: An axial slice of the volume obtained 2 mins after injection, with the selected ROI shown in the inset.

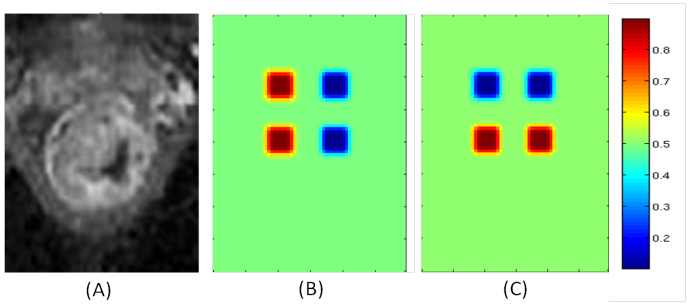

Figure 3.4: An axial slice of the pre-contrast volume (A), the simulated K^{trans} map (B), and the simulated v_e map (C) for one example simulated dataset.

each parameter map. The pre-contrast volume was of dimensions $100 \times 100 \times 52$, and each voxel was of size $1 \times 1 \times 2\text{mm}^3$. Each parameter map had four small ROIs ($\sim 10 \times 10 \times 5$ voxels), each ROI having one of the following combinations of k_{ep} and v_e: [0.1 0.1], [0.1 0.9], [0.9 0.1] and [0.9 0.9]. The other voxels had values of [0.5 0.5] for k_{ep} and v_e. An example of the parameter maps used is shown in Figure 3.4.

Each of the four parameter maps had a different spatial arrangement and shapes for the four ROIs. For each parameter map, we generated two simulated datasets – one using the Weinmann AIF and one using the Orton AIF. We used 15 different time-points (with a gap of 12 seconds between them) for each dataset. An artificial t_0 map was simulated by randomly assigning values to all voxels from a Gaussian distribution with mean 1 min and standard deviation 0.2 mins.

To simulate motion, we generated a deformation field independently for each time-point by taking a grid of dimensions $5 \times 5 \times 4$, applying a random displacement to each grid point (choosing from a uniform distribution between 0 and 5 mm along each coordinate axis) and then upsampling the grid to the image volume. We applied Gaussian smoothing (with $\sigma = 1.5$) to the final deformation field before applying it to the image with simulated contrast in order to generate motion that closely resembled the non-linear motion due to breathing and peristalsis that we have seen in real data. The parameters for the artificial deformations were chosen to be such that the maximum

deformation in the final deformation fields were of the order ± 3.5 voxels (± 5 mm) along each direction, and there were no negative Jacobians - that ensured that there was no folding in the image space. After applying the synthetic deformations, white Gaussian noise with zero mean and standard deviation σ was added independently to each time-point image. The standard deviation was chosen such that the coefficient of variation for each time-point image was between 5 and 10%.

3.5.1 Comparison between Deformation Fields

In quantitative terms, the target registration error (TRE) – i.e. the average absolute difference (in mm) between the original deformation vector and the one recovered by the algorithm – was 0.91 mm for the regular Demons algorithm, and 0.67 mm for for the diffeomorphic Demons algorithm, averaged over this entire dataset. It is worth noting that the TRE was calculated across all time-point images (15 in total).

One of the other quantities that we were concerned with was the Jacobian calculated at each position in the deformation field. As explained in Section 2.5.2, the Jacobian calculated at position (x, y, z), represents the proportional change in volume caused by the geometric transformation, in the local neighborhood of (x, y, z). A negative value for the Jacobian at any position indicates that the deformation is causing some folding in the tissue. There were no negative Jacobians in the original deformation field, and the one recovered by the diffeomorphic Demons algorithm. However, there were negative Jacobians in $\sim 0.2\%$ of the voxels in the deformations recovered by the regular Demons algorithm.

In the clinical data, the images are acquired every 12 seconds for 4-5 mins. Thus, there should be no folding in the tissue (negative Jacobians) in the deformations being applied due to motion during the scan, or the deformations being recovered by the motion correction algorithm. For this reason, coupled with the fact that it resulted in

a lower TRE, we decided to use MoRPE with the diffeomorphic demons algorithm in the further tests on simulated and clinical data.

3.5.2 Comparison between Similarity Measures

The primary purpose behind the tests on simulated data, was to compare MoRPE against other conventional similarity measures. Since SSD, MI and NCC are all pairwise cost functions as opposed to the groupwise cost function in MoRPE, three different methods were used to enable them to register all the images in the dataset:

1. Method A: All the images in that dataset were registered to the first time-point image of that dataset.

2. Method B: All the images were registered to the middle time-point image (i.e. to the image volume obtained halfway through the simulated scan).

3. Method C: Starting from the second time-point, each image was sequentially registered to the previous time-point image.

After applying each of these methods, the PK-parameters for each voxel were obtained by fitting curves to the registered dataset (using a Levenberg-Marquardt function in MATLAB 7.8, The MathWorks Inc., Natick, MA, 2000). The curve-fitting function for each dataset used the same AIF as the one used to generate that dataset. In order to test MoRPE against other similarity measures, each simulated dataset was used as an input for the algorithm, and then the output deformation fields and the output PK parameter maps were compared with the 'ground-truth' used to generate that dataset.

Table 3.1 shows the average values obtained for four different metrics after testing each method on the eight different datasets. In Table 3.1, MSE stands for mean squared error in intensity values between the registered images and the original simulated images before applying deformations, TRE for target registration error (i.e. the average

difference between the ground truth deformation field, and the deformation field estimated by the algorithm, in mm). The errors in k_{ep} and v_e denote the root mean squared difference between parameter values used to generate the dataset, and the parameter values estimated using each method. BR (before registration) denotes the value of each of the above-mentioned measures before any registration algorithm was run. In each row, the smallest values are indicated in bold.

For a dataset containing $100 \times 100 \times 52$ voxels, MoRPE takes 12 mins, NCC 4 mins, NMI 3.5 mins and SSD 1 min. It is worthwhile to note that in our method all the images are co-registered simultaneously and the PK-parameter values are updated in each iteration, with the final values being obtained as outputs at the completion of the algorithm. In the other methods, each image is registered independently, and the final PK-parameter estimates are obtained by curve-fitting after the registration process.

The effect of motion correction on the signal intensity curve for each voxel, is expected to be a closer fit between the model prediction and the signal intensity curve in the motion-corrected dataset. Figure 3.5 shows the plots for the signal intensity curve obtained before registration, and after motion correction using SSD, NCC and MoRPE, for two different locations in the simulated dataset. It is clear that the signal intensity curve obtained after performing motion-correction using MoRPE is closest to the signal intensity curve in the ground truth data (i.e. the model prediction curve in the simulated dataset before applying any artificial deformations).

From Table 3.1, it is clear that MoRPE performed the best, followed by NCC. Hence for further tests, we restricted ourselves to comparisons between MoRPE and NCC. In order to assess the statistical difference between the values obtained for NCC and MoRPE, we calculated the p-value between the results obtained for each of the four measures using a two sample t-test, across 1000 randomly sampled voxels from each of the 8 simulated datasets ($\sim 0.5\%$ of the total voxels for each dataset). This

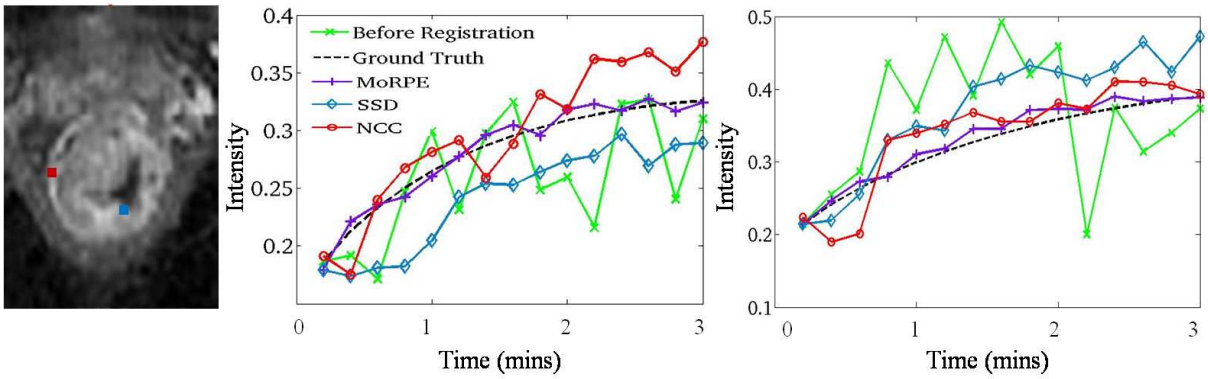

Figure 3.5: An axial slice of the first time-point image (left) and voxel plots for the blue voxel near the necrotic core (center), and the red voxel in the enhancing rim (right). The signal-intensity curve for the ground-truth image is shown as a dashed black line, the one observed before registration is shown in green, the one obtained after SSD registration is shown in light blue, after NCC registration in red, and after applying MoRPE in purple.

was done to avoid neighbouring voxels being included in the same test. The p-value for MSE was $\sim 2 \times 10^{-3}$, TRE $\sim 7 \times 10^{-4}$ and for the difference in k_{ep} and $v_e \sim 10^{-5}$. In order to get an idea of the distribution of errors of the PK-parameter estimation, we also looked at the standard deviations of the errors. The standard deviation of the errors in estimation of the PK-parameters, across all voxels was: 0.009 (MoRPE) and 0.011 (NCC using Method C) for k_{ep}; and 0.0027 (MoRPE) and 0.0043 (NCC using Method C) for v_e. This shows that the gap between the estimates and the ground truth values was fairly consistent.

3.6 Conclusions

In this Chapter, we presented the mathematical framework of a probabilistic model-based motion correction and parameter estimation method. The method was tested with different deformation models, and it was found that the diffeomorphic Demons method resulted in a lower TRE than the regular Demons algorithm in a test on sim-

		Method A			Method B			Method C			
Measure	BR	NMI	SSD	NCC	NMI	SSD	NCC	NMI	SSD	NCC	MoRPE
MSE	0.171	0.121	0.125	0.068	0.104	0.126	0.081	0.103	0.084	0.059	**0.057**
Error in k_{ep}	0.452	0.172	0.093	0.053	0.157	0.112	0.061	0.093	0.079	0.042	**0.036**
Error in v_e	0.121	0.059	0.056	0.024	0.048	0.075	0.032	0.042	0.027	0.025	**0.020**
Error in t_0	0.34	0.21	0.28	0.26	0.19	0.25	0.21	0.18	0.20	0.17	**0.09**
TRE	1.95	1.153	1.064	0.865	1.138	1.045	0.841	1.209	0.882	0.792	**0.713**

Table 3.1: Average results (over 8 datasets) for motion correction and parameter estimation on simulated data: BR stands for before registration, NMI for normalized mutual information, SSD for sum-of-squared-differences, NCC for normalized cross-correlation. MSE is the mean-square difference between the registered images and the ground-truth images and TRE is the target-registration error (i.e. average difference between the estimated and the ground truth deformation fields in mm). In Method A, all time point images are registered to the first time-point image. In method B, they are all registered to the middle time-point image, and in Method C, each image is registered to the previous time-point image. The lowest values in each row are shown in bold.

ulated data. The regular Demons algorithm also resulted in negative Jacobians (indicating folding of the tissue) in parts of the deformation field. Thus, we opted to use diffeomorphic Demons for the subsequent tests on simulated and real data.

The primary objective of the tests on simulated data was to compare the efficacy of MoRPE with the efficacy of other conventional similarity measures, in being able to recover the original PK parameter maps and deformations that are used to generate the simulated data. In the tests, it was found that MoRPE resulted in the lowest error between the recovered and original parameter maps. It also provided the lowest TRE, implying that it performed better than the other algorithms in being able to recover the deformations applied to the simulated data. The method that performed best in both these respects, after MoRPE, was the algorithm using NCC as the cost function. Thus, in our subsequent tests on clinical data in Chapter 4, we restricted ourselves to comparisons between MoRPE and NCC.

Although the tests on simulated data gave many useful indications about the performance of the method, they were obviously unable to provide information about the

applicability of the PK model and AIF that were used within MoRPE. This was because the PK model (Tofts) and the AIFs (Weinmann/Orton) used in the MoRPE algorithm or the curve-fitting routine for the other similarity measures, were the same as the PK model and AIF used to generate the data. In a clinical setting, there is no guarantee that the original, uncorrupted data would correspond to a particular PK model with a particular population averaged AIF. The only way to assess the usefulness of this method in a clinical setting is to determine whether the analysis of DCE-MRI data using MoRPE enables patient stratification or prediction of response in a more effective manner than using other algorithms. This will be investigated in the next chapter.

Chapter 4

Applying MoRPE to Colorectal Cancer

Currently, colorectal cancer is the third most common cancer [Jemal et al., 2009] and the fourth most common cause of cancer deaths worldwide, with an estimated 1.2 million cases and 609,000 deaths in 2008 alone [Karsa et al., 2010]. The most widely accepted clinical practices used to combat colorectal cancer are chemotherapy, radiation and surgery. Out of these, the first two are administered concurrently according to the clinical protocol followed at many hospitals, in the form of chemoradiotherapy (CRT). Roughly speaking, about 70% of the colorectal cancer patients respond to CRT, out of that 10-20% are cured completely. For the other 30% this procedure only results in unnecessary pain and suffering for the patients. Usually CRT is followed by surgery (unless the tumour is already thought to be too widespread, or completely cured by CRT) that sometimes results in complete responders undergoing unnecessary surgery.

In light of this, if the analysis of DCE-MRI sequences could enable us to distinguish the 30% non-responders and the 15% responders at an early stage during the treatment, it would help immensely in reducing unnecessary suffering as well as

expenses [George et al., 2001]. In two previous studies – [Bhushan et al., 2011] and [Bhushan et al., 2012], an analysis of DCE-MRI scans of colorectal cancer patients acquired before and after CRT has been presented, and the potential of MoRPE to discriminate between responders and non-responders has been demonstrated. In the study outlined in this chapter, further advances are made towards patient stratification on a larger cohort of patients.

The ultimate aim of analysing DCE-MRI data is to derive physiologically relevant parameters for each tumour that can help clinicians to devise the most appropriate treatment plan for each individual patient. Many groups across the world have attempted to use information derived from DCE-MRI scans to predict response to therapy [Padhani, 2002], [Zahra et al., 2007], [Zhang et al., 2008]. There are several different features in DCE-MRI images that can be used to predict response to CRT – such as relative signal intensity, slope of the time-intensity curve, area under the gadolinium curve, etc. However, these semi-quantitative measures vary depending upon the scanner and the contrast-agent that is used. In recent years, these measures have given way to the more standardized quantities that can be derived by applying pharmacokinetic models to the DCE-MRI data [Yang and Knopp, 2011].

As we discussed in the previous chapter, one of the leading causes for errors in DCE-MRI image analysis is the patient-motion occurring during the scan. A PK model-based procedure for motion correction and accurate parameter estimation (MoRPE) was outlined in Chapter 3 and tested on simulated data. The clinical value of using this method is the subject matter of this chapter. Once all the volumes acquired during one DCE-MRI scan have been co-registered to ensure that each anatomical location occupies the same position in all the images, then the next step is to use the physiological information captured by the PK parameter maps to predict a patient's response to therapy. Usually, the indicators that are used to determine whether a

patient was a (non-)responder are either long term follow-up data (at 1 or 3 years after treatment), or histological data obtained after resection of the tumour.

When looking at long term patient survival in studies where multiple forms of treatment – chemo/radiotherapy *and* surgery – have been used, it is often difficult to determine what the dominant cause of survival has been. Since the objective in this study is to quantify the effect of CRT on each patient within five weeks of treatment, and to modify the treatment plan accordingly, it was considered appropriate to use the modified rectal cancer regression grade (m-RCRG) [Bateman et al., 2009] determined from the histological data (obtained after TME at eight weeks) as the benchmark with that the prediction based on the DCE-MRI scan is compared.

According to the m-RCRG grading system, grade 1 tumours are those which contain no tumour epithelium, with malignant epithelium comprising $<5\%$ of the area of abnormality; grade 2 tumours contain a combination of tumour epithelium and fibrosis, with malignant epithelium comprising 5-50% of the area of abnormality; grade 3 tumours are those which appear not significantly different from tumours that have received no CRT at all, with malignant epithelium occupying over 50% of the area of abnormality.

Under the clinical protocol followed at the Oxford University hospitals, all patients were sent for surgery eight weeks after starting CRT, and the histological data collected at surgery was used to determine whether the patient had been a responder or non-responder to CRT. In this study, the predictions are made on the basis of DCE-MRI data collected before therapy, and five weeks after starting therapy – i.e. three weeks before the scheduled surgery – but the same technique could potentially be used to make predictions at an earlier stage as well. Since 15-20% of the patients are likely to be complete responders, if these patients could be identified even five weeks after the commencement of CRT, then several unnecessary surgeries could be avoided. Similarly, for patients who are not responding at all, three weeks of unnecessary chemo/radiotherapy

could be avoided if they were correctly identified.

4.1 Materials and Methods

In Chapter 3, we described the mathematical framework of MoRPE, and discussed tests on simulated data. We saw that MoRPE performs better than other motion correction and PK estimation methods in being able to recover the ground-truth deformations and PK parameter maps. However, the only way to determine whether MoRPE has any clinical benefits, is to test its efficacy in being able to predict outcome to therapy and its ability to help in patient stratification. For this purpose, we applied MoRPE retrospectively to clinical data collected from twenty patients suffering with colorectal cancer.

4.1.1 Data Acquisition

Twenty patients with histologically proven adenocarcinoma underwent DCE-MRI scans before, and after five weeks of CRT (total radiation dose of 45 Gy with concurrent Fluorouracil). All patients gave informed consent and ethical approval was obtained for the entire study. For acquiring and processing the clinical data, every effort was made to follow the international consensus developed by the Quantitative Imaging Biomarkers Alliance [Ashton, 2010]. All the data were acquired on a 1.5T General Electric scanner with four variable flip angle images (at 3, 9, 12 and 15 degrees) acquired before the dynamic sequence. The dynamic data were acquired using a T_1-weighted, spoiled gradient-echo, fat-suppressed sequence (LAVA) with TR=4.5ms, TE=2.2ms and flip angle $\alpha = 12$ degrees. The contrast agent (MultiHance) was injected via a peripheral vein after four initial acquisitions and the MRI images were then acquired every 12 seconds for the next 5 minutes. The volume acquired immediately before

arrival of CA (i.e. the fourth volume) was used as the pre-contrast volume (X_0). For most patients, 27 volumes were acquired during the DCE-MRI scan, with each volume being of dimensions $512 \times 512 \times 52$, and each voxel having the dimension $1 \times 1 \times 2\,\mathrm{mm}^3$. The T_{10} map was estimated from the variable flip-angle images using the procedure described in [Fram et al., 1987], and this map was used in the subsequent analysis.

All the tumours were resected eight weeks after the first DCE-MRI scan, and the modified rectal cancer regression grade (m-RCRG) was calculated by histological assessment [Bateman et al., 2009]. According to the m-RCRG, eleven of the twenty patients were found to be partial responders (m-RCRG= 2), while nine of the twenty were found to be non-responders (m-RCRG= 3). Since all the patients had a fairly advanced tumour by the time that they were recruited into the study, unfortunately none of them turned out to be complete responders. The DCE-MRI data from the pre- and post-therapy scans (at zero and five weeks) was analyzed retrospectively in order to classify patients as responders or non-responders (keeping the analysis blind to the histological gradings), and then compared to the m-RCRG, which was used as the gold standard. All patients with a complete dataset, i.e. pre- and post-therapy scans along with properly acquired T_{10} maps, were included in the study.

4.1.2 Test Design

Once the DCE-MRI data for each patient had been collected before and after therapy, the retrospective analysis of the data was carried out in the following manner:

1. The tumour was segmented before, and five weeks after CRT on the high-resolution T_2-weighted scan acquired before the dynamic scans, by a trained radiologist, and MoRPE was applied to a rectangular ROI containing the tumour and a 5 cm margin around it, for the pre and post-therapy scans for each patient, and the final PK parameter maps obtained as outputs from the MoRPE algorithm were

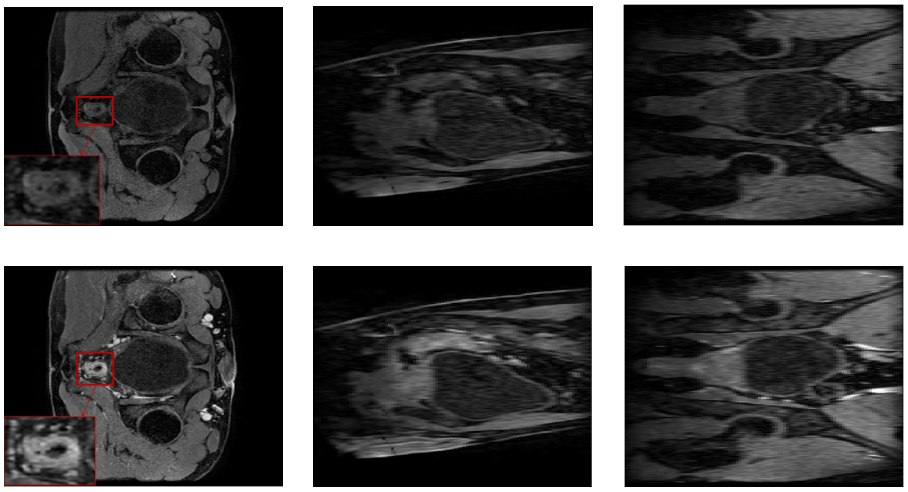

Figure 4.1: Images for $t = 30s$ (top row) and $t = 120s$ (bottom row) showing the central slices in the axial (left), sagittal (centre) and coronal (right) planes.

saved.

2. The histograms corresponding to each PK parameter map (within the segmented tumour) were calculated after the separate analysis of the pre- and post-therapy scans for each patient.

3. The patients were retrospectively classified into two groups (responders and non-responders) on the basis of a comparison between the pre- and post-therapy K^{trans}, k_{ep} or v_e histograms.

4. The patients were also classified into two groups on the basis of the change in the size of the tumour due to CRT. This was measured as the difference between the size of the tumour before therapy, and the size of the tumour after therapy, in terms of a percentage (of the pre-therapy size), as determined from the segmentation on the high-resolution T_2-weighted scan.

5. The classification done in the two ways mentioned above was compared with the modified rectal cancer regression grade (m-RCRG) calculated by trained pathol-

ogists from the tumour that was resected eight weeks after the commencement of CRT, and the efficacy of the method was calculated based upon this comparison.

The mathematical framework of MoRPE, and the manner in which it is applied to DCE-MRI data has already been discussed in the previous chapter. The exact manner in which Steps 2, 3, 4 and 5 from above were performed, will be discussed in detail in Section 4.3. There were three primary objectives of our tests on clinical data. Firstly, we wanted to determine whether MoRPE was able to perform more accurate patient stratification as compared to other motion correction algorithms (i.e. NCC). Secondly, we wanted to determine which PK parameters are most useful for the purpose of distinguishing between responders and non-responders. Lastly, we wanted to assess the impact of using motion correction and different AIFs within the framework of MoRPE itself.

The importance of choosing the correct AIF within the PK model has been discussed previously in [Li et al., 2011] and [Wang et al., 2008]. In this study, we used two different AIFs (the Weinmann and Orton models) within each of the three methodologies: (1) fitting curves to the original data; (2) fitting curves to data registered using NCC; (3) applying MoRPE. Each of these three methods relies upon the Tofts model, and the chosen AIF is integral to the Tofts model. Thus, the importance of each AIF could be judged by using the Tofts model with each AIF separately, and then computing results.

4.2 Qualitative Observations

Before making predictions about the response to therapy, we compared the observations made on simulated data in Chapter 3, to some qualitative observations obtained after running MoRPE on the clinical data.

In our tests on simulated data, the PK parameter maps and deformations used to

generate the data were all known. Hence the parameter maps and deformations recovered by each algorithm could easily be compared with the original maps and deformations to determine the average error corresponding to each metric, for each algorithm. For the clinical data, no such 'ground-truth' deformation maps or PK parameter maps were available.

There was no direct quantitative way of determining whether the PK parameter maps estimated after applying MoRPE were more accurate than the ones estimated by simply fitting curves to the data (without motion correction). However, one possible way to judge each method was to examine the residual error between the data and the model prediction after applying each method to the data.

Naturally, one would expect that a lower value for the residual for a certain dataset indicates a closer relationship between the recovered/observed data, and the PK model-based prediction of the data. Thus, one immediate way to judge the validity of each motion correction method was to see whether there was an improvement in the estimation of the contrast take-up curves for each voxel. By 'improvement', we mean whether the quality of the curve-fit was improved, or whether the residual error was reduced.

We saw that for the tumour region, the estimation of the CTCs was improved substantially at many positions. At most positions that were a large distance away from the tumour region, the CTCs remained approximately the same. In order to get a quantitative estimate of how MoRPE reduces the residual error between the model prediction and the motion-corrected data, we calculated the residuals across the entire tumour volume after applying each algorithm.

Figure (4.2) shows the residuals for voxels across the tumour volume obtained after motion correction using NCC (middle row) and after running MoRPE (bottom row). Here, the residual for the NCC-based registration algorithm is calculated as the difference between the dataset obtained after motion correction, and the model prediction

obtained after fitting curves to the *registered* dataset. Similarly, the residual for MoRPE is calculated as the difference between the registered dataset, and the model prediction based on the PK parameters obtained after applying MoRPE.

It is clear from this Figure that using MoRPE results in lower residuals across most voxels. One could argue that this is simply due to the fact that MoRPE is made to reduce the residuals by design – i.e. by virtue of the mathematical formulation of the cost-function – whereas this is not the case for NCC. However, previous studies such as [Tofts et al., 1995b], [Tofts et al., 1999], etc have shown that DCE-MRI data does indeed follow the Tofts model, and therefore a good motion correction algorithm would be expected to reduce the difference between the motion-corrected data and the best-fit model prediction. As discussed in the results on simulated data, it seems that the most likely reason that conventional similarity measures are unable to reduce the residuals as much, is because the changes in intensities, as well as their statistical compositions in images acquired at consecutive time-points are too large for the conventional methods to compensate for as accurately as a PK model-based method.

4.3 Tests on Prediction of Response

As we discussed earlier, the real litmus test for demonstrating the advantages (over other methods) of using MoRPE, is to investigate whether its use improves the ability to distinguish between responders and non-responders to CRT.

The basic steps that were followed in order to divide the group of twenty patients into responders and non-responders have been outlined in Section 4.1.2. However, each of those steps consists of several smaller components. Since it is very difficult to segment the tumour directly in the dynamic scan, as prescribed in Step 1, the radiologists first segmented the tumour on the high-resolution T_2-weighted volume acquired before each

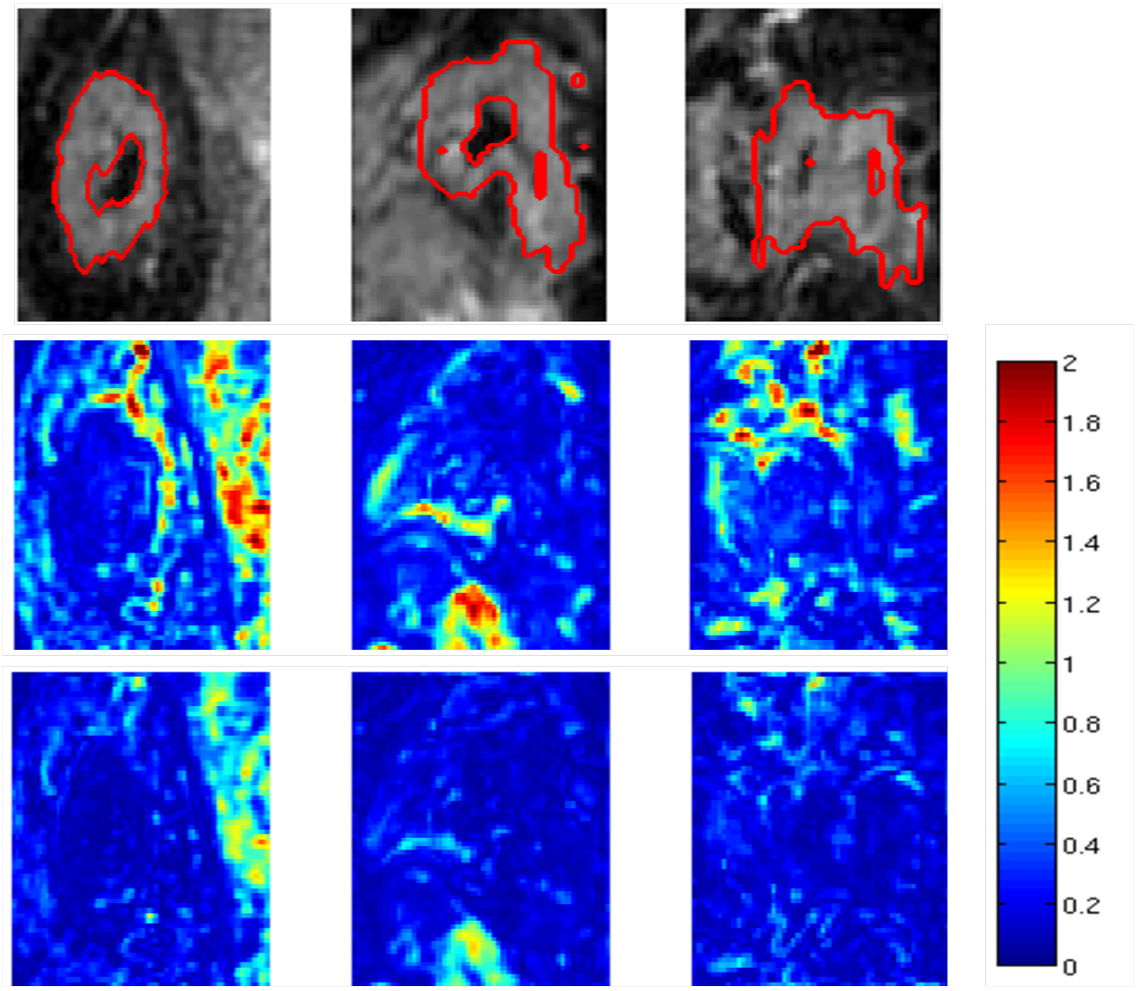

Figure 4.2: Axial (left), coronal (middle) and sagittal (right) slices through the tumour ROI with the tumour segmentation shown in red (top row). The corresponding slices of the residuals obtained after motion correction using NCC are shown in the middle row, and the residuals after running MoRPE are shown in the bottom row. The 'residual' refers to the absolute differences between the model prediction and the motion corrected data (scaled between 0 and 1), summed over all 25 time-points.

dynamic scan. The T_2-weighted volume was then rigidly registered to the first volume of the pre-therapy scan using FLIRT (part of FSL) with Mutual Information as the similarity measure [Jenkinson et al., 2012]. Thereafter, the mask delineated on the T_2-weighted volume was used to analyze the PK parameter maps obtained after running MoRPE (or NCC-based motion correction) on the dynamic data.

In order to run MoRPE on the dynamic scan, we also required the T_{10} map as an input. This was calculated from the variable flip-angle images acquired immediately before the dynamic scan, using the procedure described in Section 2.3. The flip-angle images (and therefore the T_{10} map) had exactly the same orientation and field-of-view as the dynamic scan, but they required resampling to compensate for the difference in resolutions. Once the T_{10} map had been acquired, MoRPE was run on a rectangular ROI containing the segmented tumour and a 5 cm margin around all edges. The PK parameter maps obtained as outputs from MoRPE were then saved for subsequent analysis.

After the motion correction and PK analysis, cumulative distribution functions (CDFs) for K^{trans}, k_{ep} and v_e were calculated over the segmented tumour region, from each PK-parameter map. The CDFs for the pre- and post-therapy K^{trans} maps were then compared by calculating the Kolmogorov-Smirnoff (KS) distance – i.e. the maximum absolute difference of the CDFs – between them, and this measure was used to classify each patient as a responder/non-responder. KS distance was chosen as a statistical measure since it is non-parameteric and insensitive to sampling, and thus does not make any assumptions about the distribution of K^{trans} within each tumour. This was desirable since there was no prior knowledge about the distribution that K^{trans} followed, and since it minimized the need for user interaction.

Since the current gold-standard for assessing response to therapy from MRI scans is a change in the tumour volume cause by CRT, we also calculated the change in

tumour volume for each patient, as a percentage of the pre-therapy size of the tumour. The size of the tumour was calculated from the segmentations performed by trained radiologists on the high-resolution T_2-weighted MRI scan acquired immediately before each dynamic scan. For the sake of consistency, in order to avoid partial-volume effects, only those voxels that fell wholly within the tumour boundaries were included while calculating the size of the tumour.

Once the changes in size of the tumour, and the KS-distances corresponding to each PK parameter, between the pre- and post-therapy scans for each patient had been calculated, the differences between the values obtained for the responder and non-responder patient groups were statistically assessed using two measures: (1) p-values calculated from a two-sample t-test between the distribution of changes in tumour size, and KS-distances for the two patient groups; (2) leave-one-out validation (LOOV), wherein normal distributions were fitted to the changes in tumour size and the KS-distances for each patient group, and then each individual sample (KS-distance/change in tumour size) was classified using all other samples as training data, to one group or the other on the basis of its relative proximity (in terms of standard-deviations).

4.4 Results

4.4.1 Changes in Tumour Volume

As discussed in the previous section, the current gold standard for assessing response to therapy, is the change in the volume of the tumour, as measured in the MRI scans acquired before and after therapy. Thus, we first tested the efficacy the change in tumour volume, expressed in terms of a percentage of the pre-therapy tumour volume, as a measure for distinguishing between responders and non-responders. Thus, the changes in volume for each patient were computed, and then retrospectively compared

with the m-RCRG grade determined after surgery.

In order to statistically evaluate the efficacy of this measure, we calculated the p-value (from a two-sample t-test) between the values of 'change in tumour volume' calculated for the partial responder and non-responder groups. The p-value was determined to be 0.22, whixh is clearly not sufficiently reliable as a measure for distinguishing between the two patient groups. Statistical evaluation of this measure was also performed by leave-one-out validation (LOOV), where normal distributions were fitted to the changes in tumour size for each patient group, and then each individual patient was classified using all other samples as training data, to the partial responder or non-responder group on the basis of its relative proximity (in terms of standard-deviations). Using this method, only 55% of the patients were correctly classified, while 45% of the patients were classified incorrectly.

4.4.2 Changes in PK Parameter Distributions

The reasons for the relative failure of the change in tumour size in correctly classifying patients will be discussed in Section 4.5. However, the main aim of this study was to assess the efficacy of the change in PK parameter maps, as measured by the KS-distance between the pre- and post-therapy distributions, in distinguishing between partial responders and non-responders.

Figure 4.3 shows the K^{trans} maps obtained for one example non-responder patient (m-RCRG = 3), analyzed using MoRPE (with the Orton AIF), and the NCC-based motion correction and curve fitting. Figure 4.4 shows the corresponding Figure for one example partial responder patient (m-RCRG = 2). The cumulative distribution functions (CDF) for the two patients, for the pre- and post-therapy K^{trans} maps are shown in Figure 4.5. In this example, it can be seen that MoRPE discriminates more clearly between the responder and non-responder even by a visual assessment of the

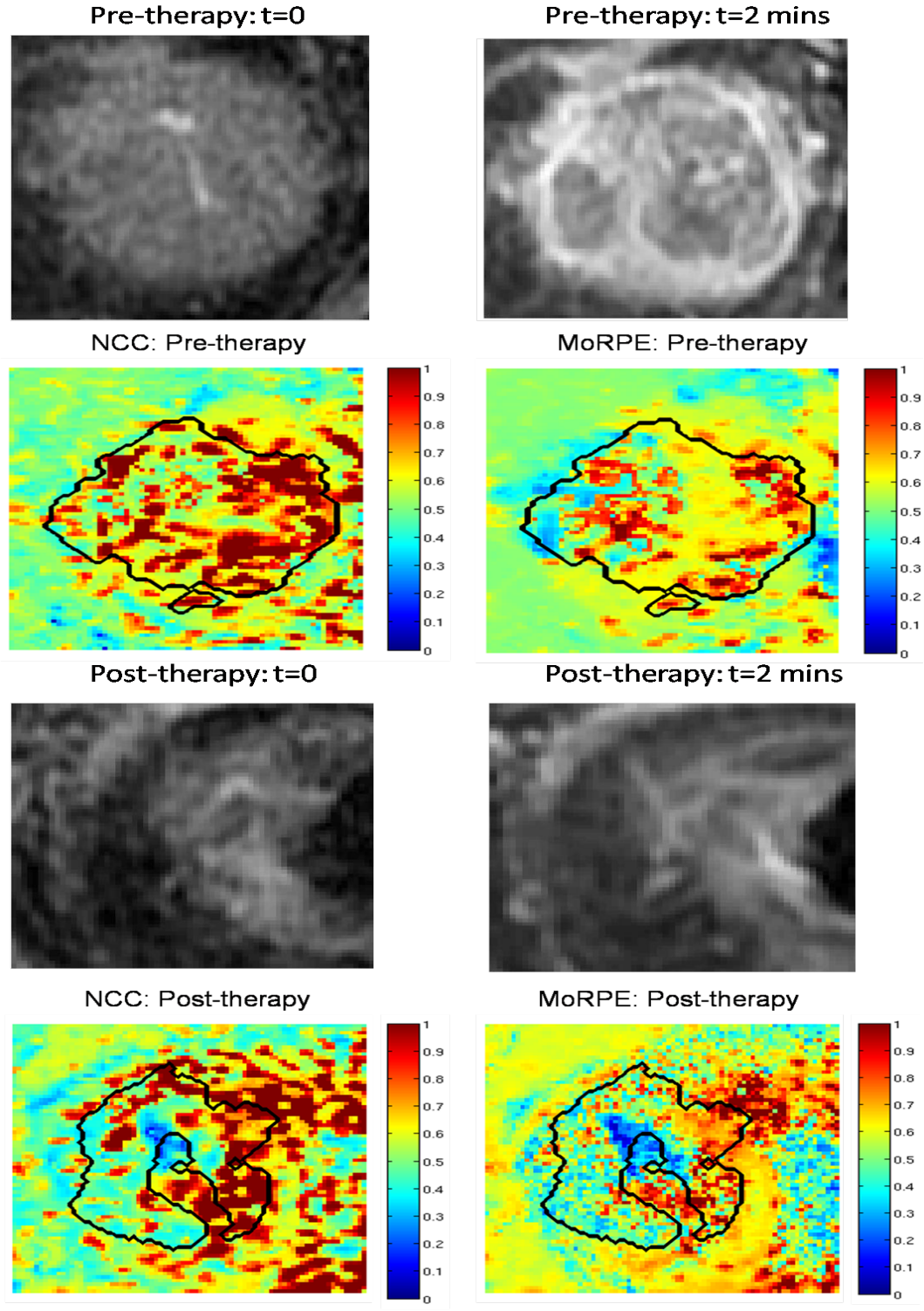

Figure 4.3: Axial slices of the pre-contrast volume (left column, 1st row) and the post-contrast volume (right column, 1st row) acquired *before* therapy for a *non-responder*, and the K^{trans} maps obtained for the same using NCC (left column, 2nd row) and using MoRPE (right column, 2nd row). Axial slices of the pre-contrast volume (left column, 3rd row) and the post-contrast volume (right column, 3rd row) acquired *after* therapy for the same patient, and the K^{trans} maps obtained using NCC (left column, 4th row) and using MoRPE (right column, 4th row).

CDFs. This difference between the pre- and post-therapy distributions can be quantified by measuring the KS-distance. Figure 4.5 shows that the KS-distance for the responder is substantially greater than the KS-distance for the non-responder when using MoRPE. This difference between the KS-distances is much smaller when the NCC-based motion correction is used.

As stated earlier, the objectives of the tests on the clinical data were: (a) to estimate the importance of motion correction within MoRPE, and the advantage of using MoRPE over NCC-based registration; (b) to determine that PK parameters are most useful for the purpose of distinguishing between responders and non-responders; (c) to assess the impact of using different AIFs within MoRPE. In order to answer these questions, we used three methods to analyse the data:

1. Curve-fits – PK model curves were fitted to the data, without applying any motion correction.

2. NCC – PK model curves were fitted to the data after applying motion correction using NCC as a similarity measure.

3. MoRPE – The MoRPE algorithm was applied to the data.

Table 4.1 presents a summary of each method's efficacy in distinguishing between responders and non-responders. It shows the p-values and the percentage of patients mis-classified during leave-one-out-validation (LOOV), while using each of the three methods, with the two different AIFs, to distinguish between partial responders and non-responders on the basis of the KS-distance between the pre- and post-therapy distributions of each PK parameter. Four things become apparent from Table 4.1:

1. For each method (with or without motion correction), and with each AIF, the discriminatory power of K^{trans} as a PK parameter, is greater than that of k_{ep}, and the discriminatory power of k_{ep} is greater than that of v_e.

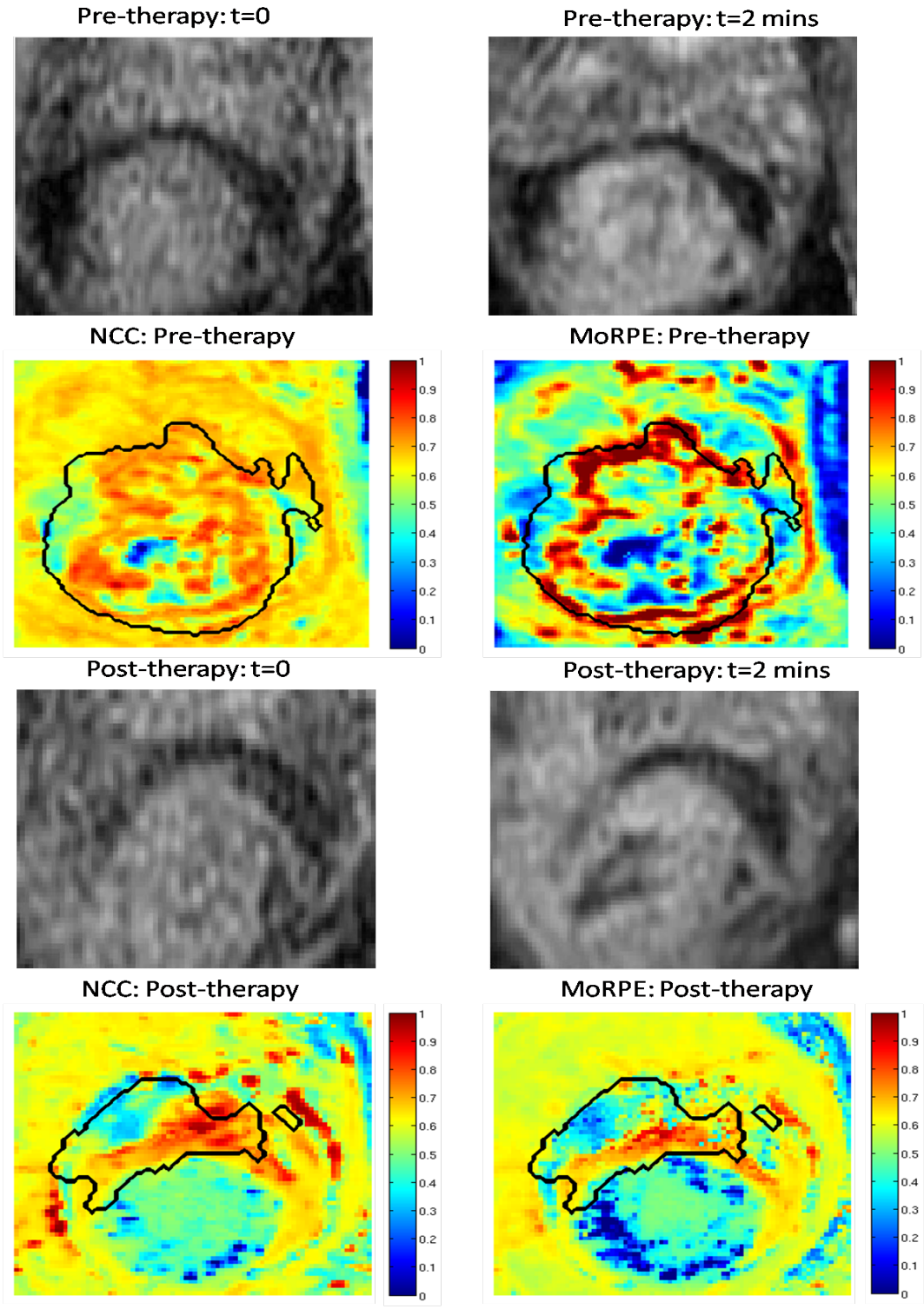

Figure 4.4: Axial slices of the pre-contrast volume (left column, 1st row) and the post-contrast volume (right column, 1st row) acquired *before* therapy for a *partial responder*, and the K^{trans} maps obtained for the same using NCC (left column, 2nd row) and using MoRPE (right column, 2nd row). Axial slices of the pre-contrast volume (left column, 3rd row) and the post-contrast volume (right column, 3rd row) acquired *after* therapy for the same patient, and the K^{trans} maps obtained using NCC (left column, 4th row) and using MoRPE (right column, 4th row).

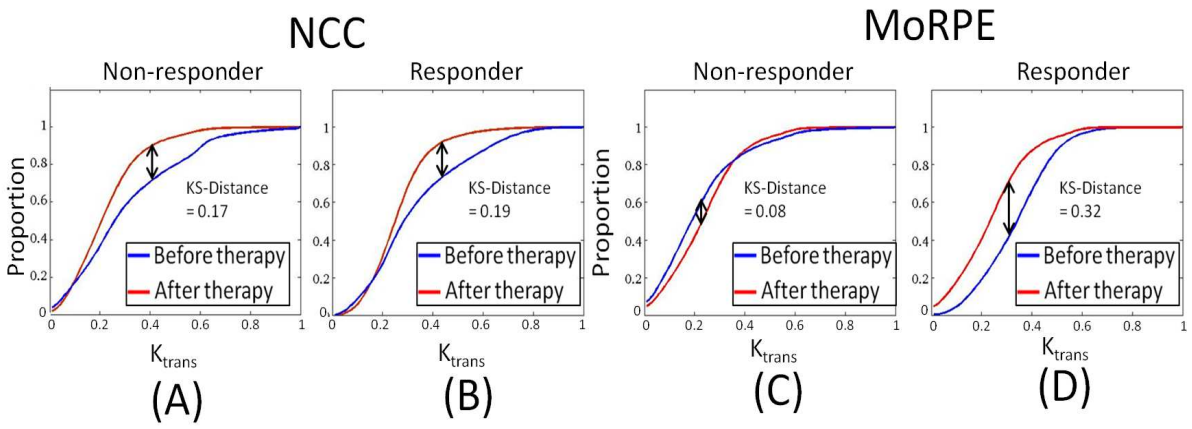

Figure 4.5: The cumulative distribution functions obtained for the pre-therapy (in blue) and post-therapy (in red) K^{trans} distributions estimated for a non-responder using NCC (A) and MoRPE (C), and for a responder using NCC (B) and MoRPE (D).

2. For each PK parameter, and each AIF used within the PK model, the discrimination between responders and non-responders is better when some form of motion correction (NCC or MoRPE) is used.

3. For each PK parameter, and each AIF used within the PK model, MoRPE is able to distinguish between responders and non-responders more easily than NCC.

4. For each method (curve-fits or NCC or MoRPE), the discrimination is better when the Orton AIF is used within the PK model, than when the Weinmann AIF is used.

Each of the four observations stated above are supported by p-values as well as leave-one-out-validation values. Due to the first observation outlined above, we have restricted ourselves to examining the change in distributions of K^{trans} for the remainder of this section. From the results, it was clear that the K^{trans} distributions obtained with MoRPE (using the Orton AIF) were the most effective in distinguishing between responders and non-responders.

While Table 4.1 gives a complete summary of how each method and each PK param-

PK parameter	Statistical Measure	Curve-fit with Weinmann	Curve-fit with Orton	NCC with Weinmann	NCC with Orton	MoRPE with Weinmann	MoRPE with Orton
v_e	p-value	0.27	0.19	0.18	0.11	0.10	0.08
	LOOV	53%	46%	43%	37%	36%	34%
k_{ep}	p-value	0.2	0.12	0.17	0.06	0.07	0.04
	LOOV	45%	38%	41%	32%	33%	29%
K^{trans}	p-value	3.6×10^{-3}	3.2×10^{-4}	3.1×10^{-4}	2.7×10^{-4}	2×10^{-4}	4×10^{-7}
	LOOV	23%	19%	18%	14%	12%	0%

Table 4.1: Statistical measures for the difference between the partial responder ($n = 11$) and non-responder ($n = 9$) groups for the KS-distance measured between the pre- and post-therapy distributions of K^{trans} and k_{ep} obtained by employing each of the three methods with the two different AIFs. The p-values are calculated using a two-sample t-test, and LOOV represents the percentage of samples mis-classified during leave-one-out validation.

eter performed in the task of distinguishing between responders and non-responders, two of the most striking observations can be viewed separately in Figures 4.6 and 4.7. On the left, in Figure 4.6, we can see a plot of the KS-distances between the pre- and post-therapy K^{trans} distributions for each patient obtained by doing curve-fits (without any motion correction) using the Tofts model with the Orton AIF. The plot on the right in Figure 4.6 shows the corresponding values obtained when MoRPE (with the Orton AIF) was used to determine the K^{trans} distributions. The Figure clearly shows that there was a distinctly greater difference between KS-distances measured for responders and non-responders, when the analysis was done using MoRPE. This demonstrates the benefits of using PK model based motion correction for calculating the K^{trans} maps for each clinical dataset.

The impact of using different AIFs within MoRPE, on the KS-distances calculated for each patient, is shown in Figure 4.7. In this Figure , the KS-distances calculated using MoRPE with the Weinmann AIF are shown on the left, and KS-distances calculated using MoRPE with the Orton AIF are shown on the right. From Figure 4.7, it is apparent that the separation between responders and non-responders is more apparent

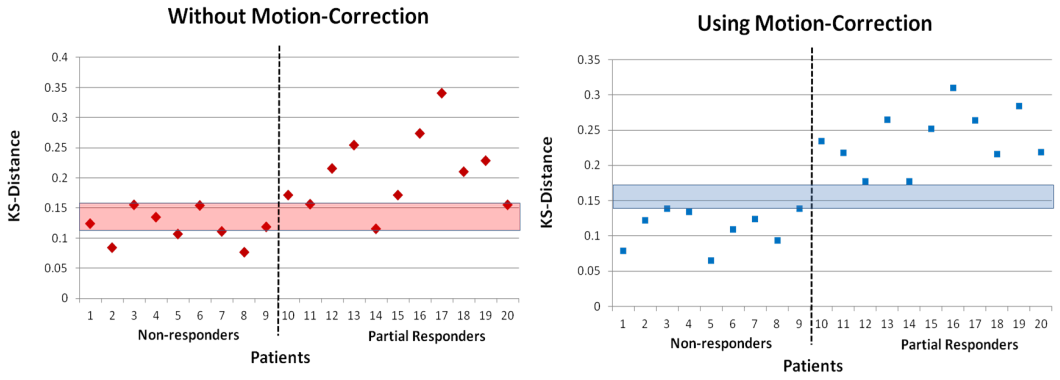

Figure 4.6: KS-distance (on y-axis) between pre- and post-therapy K^{trans} distributions obtained for each patient using curve-fits with the Orton AIF (left), and the KS-distances obtained using MoRPE with Orton AIF (right). The blue band shows the positive gap between the non-responder with the highest KS-distance, and the partial responder with the lowest KS-distance when using motion correction. The red band shows that this gap is negative without using motion correction and so a clean separation between the groups is not possible if motion correction is not performed.

when the Orton AIF is used within MoRPE. This was also reflected by the smaller p-value and LOOV value calculated for MoRPE with the Orton AIF in Table 4.1.

The differences between the two plots in Figure 4.7 arise from the fact that the K^{trans} maps being estimated by MoRPE change appreciably when different AIFs are used within the mathematical framework. Figure 4.8 illustrates the difference between the K^{trans} maps obtained for one example patient ROI using the two different AIFs. The K^{trans} map estimated by using the Weinmann AIF within MoRPE is shown in the middle, whereas the K^{trans} map estimated by using the Orton AIF within MoRPE is shown on the right.

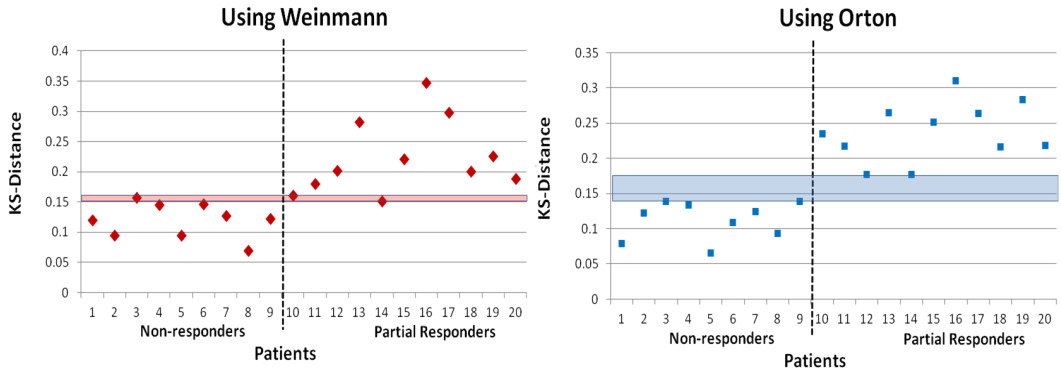

Figure 4.7: KS-distance (on y-axis) between pre- and post-therapy K^{trans} distributions obtained for each patient using MoRPE with Weinmann AIF (left) and with Orton AIF (right). The blue band shows the positive gap between the non-responder with the highest KS-distance, and the partial responder with the lowest KS-distance when using the Orton AIF. The red band shows that there is a non-responder with a higher KS-distance than a partial responder when using Weinmann, i.e. the gap is negative, and a clean separation between the two patient groups is no longer possible.

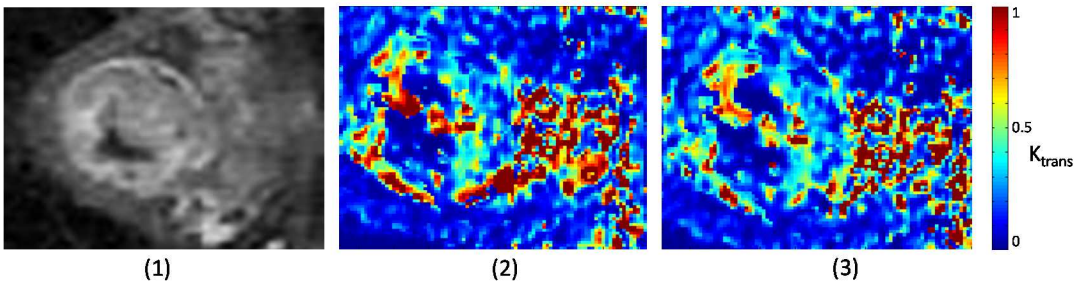

Figure 4.8: K^{trans} maps obtained using MoRPE for the ROI shown in (1): The K^{trans} map obtained with the Weinmann AIF is shown in (2) and with the Orton AIF is shown in (3).

4.5 Discussion

4.5.1 Tumour Volume as a Measure of Response

The tests described in Section 4.4.1 show that the change in tumour volume was relatively unsuccessful in helping clinicians to distinguish between partial responders and non-responders to CRT. In this study, we found that the change in tumour volume due to CRT, was very similar (30-40%) for most patients, since the segmentation of the tumour also included the necrotic core in the post-therapy scan. Thus, even if the cycles of CRT had been very effective, and succeeded in killing many of the viable cancer cells within the tumour, this was not necessarily reflected in the change in volume, since these parts may still have been included within the tumour boundaries.

As discussed earlier, all the patients recruited into this particular study had an advanced form of rectal adenocarcinoma. This was the reason that there were no complete responders in the study. This is also likely to be one of the reason for why most patients showed a similar reduction in volume of the tumour due to CRT. However, the principle reason for why there was a poor correspondence between the reduction in tumour volume, and the m-RCRG grading, is that the tumour volume alone does not provide any information on the physiological nature of the cancer cells that are present within the tumour.

Since the volume is assessed from the high-resolution T_2-weighted scan, it obviously does not take any of the dynamic information into account. Thus, it cannot be used by clinicians to determine whether the cancer cells within the tumour are viable or not, whether they are likely to multiply and grow, or whether the tumour also includes a necrotic core. In order to capture this information, it is necessary to turn to the PK parameters that are calculated from the CTCs obtained at each voxel, and which attempt to mathematically encode the physiological information contained in the CTCs.

4.5.2 Choice of PK Parameter

The results of the tests on clinical data show that motion correction, as well as the choice of the AIF used within a PK model, can have a big impact on our ability to distinguish between responders and non-responders. Section 4.4 also shows that as far as MoRPE (using the Tofts model) is concerned, the PK parameter $K^{trans}(= k_{ep} \cdot v_e)$ is most useful for the purpose of patient stratification, followed by k_{ep} and then by v_e. Although at first sight this difference may seem inexplicable, it is important to remember that each PK parameter captures slightly different physiological information. In our framework, the Tofts model is parameterized by k_{ep} and v_e in the following manner:

$$C_T(t) = k_{ep} \cdot v_e \exp\left(-k_{ep} \cdot t\right) \otimes C_p(t) \tag{4.1}$$

From the above equation, it is clear that v_e impacts only the amplitude of the curve, whereas the shape is governed by k_{ep} alone. Thus, if this parameterization is followed, then the optimization over [k_{ep}, v_e] results in k_{ep} reflecting only the shape of the curve, and v_e reflecting only the amplitude. However, K^{trans}, which is a product of both, captures changes in both – shape and amplitude of the curve. This is probably the reason that it shows better results than k_{ep} and v_e in being able to discriminate between responders and non-responders.

4.5.3 Choices for PK Model and AIFs

The pharmacokinetic analysis of DCE-MRI data of colorectal cancer patients described in this chapter, is based upon many assumptions. The central assumption behind the framework of MoRPE, is that the passage of the contrast-agent through the tumour is well described by the Tofts/Kety two-compartment model. Although there are many studies showing that this model does provide a good description of contrast-

enhancement [Tofts, 1997], there are also studies to show that other PK models may provide a better fit to the DCE-MRI data [Naish et al., 2009], [Srikanchana et al., 2004]. Thus, it is quite possible that the parameters of another model may provide even better results for discrimination between responders and non-responders than we have found in this study using K^{trans}.

Similarly, different choices for the AIF could also result in a closer correspondence between the PK model and the data [Parker et al., 2006]. In this study, we have studied the impact of two different population averaged AIFs on the ability of MoRPE to distinguish between responders and non-responders. However, in future, the applicability of other AIF models needs to be tested to determine whether the ability of this method to predict response can be improved further. Intuitively it is obvious that individually determined AIFs ought to improve the discrimination between responders and non-responders by providing a more accurate estimate of the PK parameters. However, there are several methods that can be used to determine the AIF, and there are different kinds of errors associated with each of them [Yang et al., 2004]. There are studies that also indicate the need for determining different AIFs for different regions of the body using a model-based approach, rather than from an artery or any reference tissue [Fluckiger et al., 2009]. In light of the above, it is fair to say that the subject of AIFs and their impact on prediction of response deserves a more detailed analysis than what has been possible in this thesis.

4.5.4 Methods of Prediction

The measure used to distinguish between responders and non-responders in this study, was the KS-distance between their pre- and post-therapy K^{trans}. The KS-distance between two sample distributions is simply the maximum difference calculated between the CDFs for those two samples. Although this measure does have many advantages – such

as symmetry and insensitivity to outliers, etc – there are other methods of calculating differences between probability distributions that may be better ([Rubner et al., 2000], [Hollander and Wolfe, 1973]) and hence should be tested for their ability to discriminate between responders and non-responders.

The fact that we are trying to estimate differences between two histograms calculated before and after therapy, itself merits closer examination. When a histogram is constructed from a 3D PK parameter map, all the information regarding the spatial arrangement of the PK parameter values is being lost. However, some recent studies propose alternative ways of estimating distributions of parameters, which preserve this information [Joshi and Brady, 2009], [Rajwade et al., 2006]. These non-parameteric methods of creating histograms of PK parameters present an opportunity for incorporating the spatial information contained in the original PK parameter maps, and could potential lead to an improvement in the prediction capabilities of MoRPE. This is another potential avenue for future exploration.

4.5.5 Gold Standard for Validation

In the case of the clinical data no ground-truth was available, and hence the accuracy of the output deformation fields and parameter maps of the algorithm could not be quantitatively evaluated in a direct manner. In the clinical trial data available to us, the most meaningful way of comparing different methods was to test their abilities to discriminate between partial responders and non-responders, and compare this with the results from histological analysis. It is important to remember that even though the histological measure used in this study (m-RCRG) was the only gold standard measure available, it is also subject to inter-observer variability [Bateman et al., 2009]. However, for the clinical data available to us, it was still the best measure against that other methods of discrimination could be compared.

It could be argued that follow-up data collected at 1- year or 3 years after surgery was a better measure of the patient's response to therapy [Zahra et al., 2007]. However, the key clinical question in our study was to determine if DCE-MRI scans could help to avoid unnecessary surgeries (in the case of responders), and to avoid unnecessary cycles of CRT (in the case of non-responders). The key question was to determine how well the tumour in the primary site was responding to CRT. The 1 or 3-year survival data is a cumulative result of many factors – such as distant metastases, the success of surgery, etc. Thus, it could not be used as a measure of the patient's response to CRT alone, or a gold-standard of whether surgery was necessary or not. The histology grading calculated after resection of the tumour shows whether or not there was any viable tumour left in the tissue that has been resected. Thus, a complete absence of viable tumour tissue indicates that the patient had responded well to CRT, and that the surgery had been unnecessary. This is precisely the information that was sought to be extracted from the DCE-MRI scan conducted three weeks before surgery. Hence the m-RCRG grading seemed to be the most appropriate measure against which predictions made by our method could be compared.

4.6 Conclusions

Within the framework of MoRPE, we were seeking to investigate the impact of (a) motion correction; (b) different PK parameters; and (c) the choice of the AIF; on our ability to distinguish between responders and non-responders to CRT. Although the tests on clinical data were conducted on a reasonably small cohort of patients, we have shown that using MoRPE enables a better discrimination between responders and non-responders than estimating PK-parameters through direct curve-fits, or after motion correction using NCC. The results show that K^{trans} is the PK parameter that is most

successful in discriminating between responders and non-responders, followed by k_{ep} and then by v_e. We have also shown that using the Orton AIF within the framework of MoRPE provides better results than using the Weinmann AIF.

It is likely that the reason that motion correction and the choice of the Orton AIF improve results for prediction of response, is that they improve the accuracy of the PK parameter maps that are being estimated. In practice, the use of motion correction and other analysis methods change not only the estimates of the PK parameters, but also the degree of confidence that can be placed upon those estimates. However, we need a method that can calculate uncertainties or precisions of estimates in addition to the estimates themselves, in order to further investigate this issue. Such a method will be presented in Chapter 5 and applied to the colorectal cancer data analyzed in this chapter.

Another aspect of the clinical data analysis that becomes clear from Section 4.4, is the fact that the K^{trans} maps that are obtained as an output from MoRPE are very heterogeneous, and contain a lot of spatial information that is lost when we compress the information contained in the 3D map into a histogram. Many existing studies point to the importance of heterogeneity within tumours and its impact on assessment of disease and treatment outcome [Yang and Knopp, 2011], [Jackson et al., 2007]. Calculation of tumour heterogeneity and its impact on prediction of response to therapy will be discussed in detail in Chapter 6.

Chapter 5

Estimating Uncertainties

One of the key issues that needs to be considered in the analysis of DCE-MRI sequences is the inherent uncertainty in the estimation of the PK parameters. The noise in the baseline signal, in the flip-angle sequences (that are needed to calculate the T_{10} map), in the dynamic sequence, and most importantly, the patient motion during the scan, can make it very difficult to estimate the PK parameters accurately. Thus, in order to correctly estimate PK maps, and hence to predict response to therapy correctly, it is important to take uncertainty into account.

5.1 Previous Work

The uncertainty in estimating the T_{10} map from the flip-angle images has been investigated previously in [Schabel and Morrell, 2009], and the same group of authors have also investigated uncertainty due to bias in the B_1 field in [Schabel and Parker, 2008]. The uncertainty in the onset time has been estimated using Hessians in [Orton et al., 2007], and the effect of choosing a sub-optimal flip-angle in a SPGR sequence on the precisions of the PK parameters has been investigated in [De Naeyer et al., 2011].

A more holistic analysis of uncertainty in the PK parameters has been performed

in [Buckley, 2002] and [Garpebring et al., 2012]. In [Buckley, 2002], the authors simulated data using a distributed pathways model for tracer kinetics, which takes heterogeneity and flow dispersion into account, in order to generate contrast-takeup curves. These curves were then fitted using three different PK models – the Tofts/Kety model (Model 1), the extended Tofts/Kety model (Model 2), and a third model described in [St Lawrence and Lee, 1998] (Model 3). This study found that for the same simulations, Model 1 overestimated K^{trans}; whereas Model 2 underestimated v_p (the vascular fraction, which is not included in Model 1), and again overestimated K^{trans}. Even in Model 3, which was mathematically most similar to the distributed pathways model used to simulate the data, the authors found similar discrepancies between the ground truth simulations and the outcome of the model fitting.

One of the interesting findings of the study in [St Lawrence and Lee, 1998], was that the uncertainty in k_{ep} ($= K^{trans}/v_e$) remained relatively low ($\sim 10\%$) even though the uncertainty in other parameters was much larger, and that the estimates of k_{ep} were relatively unaffected by T_{10}, thus indicating that it may not be very crucial to determine T_{10} for the purpose of getting accurate estimates of k_{ep}. This study also outlined many of the issues related to optimization while fitting models, and the lack of unique solutions at some parameter values. However, although it provides quantitative estimates of the error at different positions and for different parameter values, these quantitative estimates are still based on a simulated study for which the ground truth was known, and the method of analysis is not extendable to clinical studies where no ground-truth is available.

In contrast, the mathematical framework employed in [Garpebring et al., 2012] is more robust, and readily applicable to clinical data. The basic idea behind the model is to describe the data observed in the DCE-MRI scan as a function of three sets of parameters: (1) time, (2) $\mathbf{p}$ – the PK parameters of interest, such as K^{trans}, v_e and (3)

$\mathbf{q}$ – the auxilliary parameters, such as T_{10}, onset time, baseline signal, etc. Thus, the model for the signal S_i is described as:

$$S_i = f\left(t_i, \mathbf{p}, \mathbf{q}\right) + \epsilon_S \tag{5.1}$$

Here, ϵ_S is assumed to be independent (at each voxel) and identically distributed white Gaussian noise. The authors further assume that the variation in the primary parameters and the auxilliary parameters is small enough to approximate the signal using a Taylor-Series expansion. They finally derive an expression for the covariance matrix for the PK parameters $\mathbf{p}$, which can be used to quantify the uncertainty (in terms of the coefficient of variation) in each of the PK parameters, caused by noise in each of the auxilliary parameters. Since the calculations are based on the estimates of the Jacobian (matrix of first order partial derivatives) of the error function with respect to the parameters, this method can be applied to simulated, as well as to clinical data.

In [Garpebring et al., 2012], the authors first test the method on data generated using Monte-Carlo simulations, and the noise estimated by the method is found to resemble the noise levels in the simulated data closely. Subsequently, the method is applied to in-vivo DCE-MRI data, and the uncertainty in K^{trans}, v_e and v_p is calculated as a percentage and further broken up into its individual constituents – i.e. the uncertainties caused by the different auxilliary parameters. The study found that the onset time (t_0) was the biggest source of error, followed by uncertainty in the AIF measurement.

There are also several studies that have examined the uncertainty in non-rigid registration [Simpson et al., 2013], and also analysed the impact of uncertainty in registration parameters on surgical decisions [Risholm et al., 2010], [Risholm et al., 2013].

5.2 Our Approach

The mathematical framework used in this chapter is somewhat similar to the one used in [Garpebring et al., 2012], since both use a Taylor series approximation in a non-linear forward model. The mathematical framework used in our study is based on a Variational Bayes (VB) framework, and it will be discussed in detail in the next section. The key difference between the VB framework and the one used in [Garpebring et al., 2012], is that the VB framework employs a mean-field approximation – in this case a multivariate normal (MVN) distribution for the posterior probability.

In terms of the analysis and the results, another important difference is the parameterization followed in our study. We have not separated the parameters into two categories, since we felt that parameters treated as 'auxilliary' in [Garpebring et al., 2012], such as T_{10} and the onset time (t_0) could also be of clinical value, and we were interested in estimating and comparing uncertainties in all parameters, and also in measuring covariances between them.

Although uncertainty in registration of images has been analyzed [Risholm et al., 2013] and other studies, an aspect that has not been analysed very deeply is how motion affects uncertainties in the estimates of PK parameters. This will be the main focus in the current study. However, we are also interested in a broad and comprehensive analysis of all sources of uncertainty in pharmacokinetic modeling of DCE-MRI data.

In order to obtain a quantitative estimate of the uncertainties in each parameter involved in our DCE-MRI analysis, and to investigate the inter-dependencies between all these different parameters, we implemented our code within a VB framework described in [Chappell et al., 2009], and then analyzed the variances and covariances of all the relevant parameters and their combinations. We were primarily interested in investigating the following aspects of DCE-MRI analysis:

1. The impact of motion correction upon the uncertainties in the PK-parameter estimates.

2. Determining which parameters are strongly correlated, and that ones convey independent information.

3. Analysing the impact of incorporating uncertainties on the prediction of response to therapy.

5.3 The Variational Bayes Framework

In a DCE-MRI scan, the patient is injected with a low molecular weight contrast agent (CA), and MRI volumes are acquired at regular intervals after injection. The basic assumption of all PK models is that the concentration of the CA at each position is a function of a small set of physiological parameters (θ) of the tissue. As explained in Section 3.2, in this chapter as well, we have used the Tofts one-compartment PK model [Tofts et al., 1999]. According to the Tofts model, for a given Arterial Input Function (AIF), the concentration of CA at time t, $C(t)$, can be written as:

$$C(t) \quad = AIF(t - t_0) \otimes k_{ep} v_e e^{-k_{ep}(t-t_0)} \quad (for\ t > t_0) \tag{5.2}$$

Here $\otimes$ denotes a convolution, t_0 denotes the time at that the CA reaches the voxel under consideration, k_{ep} denotes the rate of transfer of CA from the extra-cellular extra-vascular space (EES) to the blood plasma, v_e denotes the proportion of space in that voxel occupied by the EES and $K^{trans} = k_{ep} \cdot v_e$. In order to restrict the number of variables, and to make the analysis more tractable, we have elected to use the Orton (population-averaged) AIF in this study [Orton et al., 2008].

We are interested in extracting the 'true' PK parameters ($\theta = [K^{trans}, v_e, t_0]$), which

best explain the data Y, and the uncertainties associated with these parameters. In order to accomplish this task, a Variational Bayes (VB) framework has been employed for the DCE-MRI analysis [Chappell et al., 2009]. The changing concentration of CA leads to the acquisition of MRI signal intensity curves, which are functions of θ, the MRI acquisition parameters (α), the motion applied during the scan (parameterized by T), and a noise process (ϵ). Thus, the observed data Y (expressed as a row vector) is given by:

$$Y = f(T, \theta, \alpha) + \epsilon \tag{5.3}$$

This generic formulation allows any PK model, f, and any noise model, ϵ, to be used. Here, we have assumed ϵ to be a Gaussian noise process with mean zero and precision ϕ, i.e. $\epsilon \sim N(0, \phi^{-1})$. Since the MRI acquisition parameters (such as the flip-angle, repetition time, echo time, etc) – α – remain fixed throughout, and since the motion parameters T also remain fixed during the part of the VB iteration that updates the physiological parameters, we can express $f(T, \theta, \alpha)$ as $f(\theta)$. Now, using Equation (5.3), due to the assumption of Gaussian noise, we can rewrite the likelihood term (for N data observations) from Equation (5.7) as:

$$\begin{aligned} P(Y \mid \theta, \phi) &= P(\epsilon = f(\theta) - Y) && (5.4) \\ &= \left(\frac{\phi}{2\pi}\right)^{N/2} \exp\left(-\frac{1}{2}\left((f(\theta) - Y) \cdot \phi \cdot (f(\theta) - Y)^T\right)\right) && (5.5) \end{aligned}$$

Since the noise is assumed to be normally distributed with zero mean and precision ϕ, the parameters of interest are $\{\theta, \phi\}$. The likelihood term can be linked with the posterior distribution of the parameters of interest using Bayes' rule in the following

manner:

$$P(Y \mid \theta, \phi) \cdot P(\theta, \phi) = P(Y, \theta, \phi) \tag{5.6}$$

$$= P(\theta, \phi \mid Y) \cdot P(Y) \tag{5.7}$$

Our objective is to maximize the probability of the PK parameters given the data, and to simultaneously calculate the uncertainty in each parameter. Thus, we are actually interested in the estimates of θ and ϕ that maximize their posterior probability $P(\theta, \phi \mid Y)$. From Equation 5.7, it is clear that:

$$P(\theta, \phi \mid Y) \propto P(Y \mid \theta, \phi) \cdot P(\theta, \phi) \tag{5.8}$$

Since we are using VB, the approximate posterior and prior distributions can be factorized separately over the model parameters θ and the noise parameters ϕ, in the following manner:

$$P(\theta, \phi \mid Y) = P(\theta \mid Y) \cdot P(\phi \mid Y) \tag{5.9}$$

$$P(\theta, \phi) = P(\theta) \cdot P(\phi) \tag{5.10}$$

In this study, the prior distribution on the PK-parameters, $P(\theta)$ is assumed to be a multi-variate normal (MVN) distribution, i.e:

$$P(\theta) \sim MVN(m_o, \Sigma_0) \tag{5.11}$$

Here $m_0 = [m_{k_{ep}}, m_{v_e}, m_{t_0}]$ represents the prior means of the PK parameters and Σ_0 represents their prior covariance matrix. Further, the prior distribution on the noise

parameter, $P(\phi)$ is chosen to be a Gamma distribution with parameters s_0 and c_0, i.e:

$$P(\phi) \sim Ga(s_o, c_0) \tag{5.12}$$

The factorized posteriors are chosen to be conjugate to the factorized priors, in the following manner:

$$P(\theta \mid Y) \sim MVN(m, \Sigma) \tag{5.13}$$

$$P(\phi \mid Y) \sim Ga(s, c) \tag{5.14}$$

The VB algorithm assumes a mean-field approximation to the posterior distribution, in this case using a MVN for the distribution over the PK parameters. By choosing this conjugate distribution, an iterative update procedure can be derived to find the mean and precision of the distribution. Now, combining the likelihood term and the prior distributions, we can write the negative log-posterior ($L = -\log P(\theta, \phi \mid Y) = -\log(P(\theta \mid Y) \cdot P(\phi \mid Y))$) as:

$$\begin{aligned} L &= \frac{1}{2}\left((f(\theta) - Y) \cdot \phi \cdot (f(\theta) - Y)^T\right) - \frac{N}{2}\log\phi + \frac{1}{s_0}\phi \\ &\quad + \frac{1}{2}(\theta - m)\,\Sigma\,(\theta - m)^T - (c_0 - 1)\log\phi + const \end{aligned} \tag{5.15}$$

In order to minimize L, the function $f(\theta)$ is approximated by a first-order Taylor-series expansion about the mean of the posterior (MVN) distribution to give: $f(\theta) = f(m) + J(\theta - m)$, where J denotes the Jacobian (matrix of partial derivatives), calculated at the current estimate of the mean, i.e. $J_i = \left.\frac{\partial f(\theta)}{\partial \theta_i}\right|_{\theta=m}$. The parameters that we are interested in estimating are those associated with the PK properties (θ) and the noise (ϕ). Using this Taylor-series approximation, and following the procedure described in [Chappell et al., 2009], the updates to the mean and covariance matrix for the PK

parameters can be computed as:

$$\Sigma_{new}^{-1} = scJ^T J + \Sigma_0^{-1} \tag{5.16}$$

$$m_{new}\Sigma_{new} = scJ\left(k + m_{old}J\right) + m_0\Sigma_0 \tag{5.17}$$

Here, $k = Y - f(\theta)$ represents the (column) vector of residuals, whereas s and c are the parameters governing the noise precision. While c remains fixed as a function of the number of time-points (N) in the dataset ($c = \frac{N}{2} + 1$), the other noise precision parameter s is updated in the following manner:

$$\frac{1}{s} = \frac{1}{s_0} + \frac{1}{2}k^T k + \frac{1}{2}\mathrm{Tr}\left(\Sigma \cdot J^T J\right) \tag{5.18}$$

The criteria for convergence of the algorithm, as well as issues relating to the optimization are discussed in detail in [Chappell et al., 2009]. However, in our implementation, in addition to the updates detailed above, we also register the raw data to the best-fit model prediction after each iteration of the VB algorithm. A non-linear diffeomorphic Demons deformation framework (as described in Section 3.4) is used for the registration, and new estimates of the motion parameters T and the data Y, are obtained after each iteration of the VB algorithm. Due to the addition of this crucial motion correction step, the convergence criteria discussed in [Chappell et al., 2009] are no longer strictly valid. However, since the registration step only serves to reduce the error between the model prediction and the data, the convergence is usually not affected by the addition of this step. In practice, all the empirical tests show that if the optimization converges without motion correction, then it also converges with motion correction.

To give a summary of the entire procedure of how the VB and the motion correction algorithms are combined, here is an outline of the entire algorithm:

Algorithm 5.1 The VB and motion correction algorithm

Initialize $T_0 = I$ (Identity) transform, Y_0= original DCE-MRI data, $[m_0, \Sigma_0]$ =hyper-parameters of $P(\theta)$ (prior distribution of the PK-parameters) and $[s_0, c_0]$ = hyper-parameters of $P(\phi)$ (prior distribution of the noise parameters).

for i=1,...,t

1. Minimize $L = -\log\left(P\left(\theta \mid Y_{i-1}\right) \cdot P\left(\phi \mid Y_{i-1}\right)\right)$ with respect to hyper-parameters and arrive at new estimates $[m_i, \Sigma_i]$ and $[s_i, c_i]$
2. Update $P\left(\theta \mid Y_{i-1}\right) \sim MVN\left(m_i, \Sigma_i\right)$ and $P\left(\phi \mid Y_{i-1}\right) \sim Ga\left(s_i, c_i\right)$
3. Update model prediction $f_i(\theta) = f_i(m_i)$
4. Calculate $T_i = \arg\min_T \left(f_i(\theta) - T\left(Y_{i-1}\right)\right)^2$
5. Update $Y_i = T_i(Y_{i-1})$
6. Update i := i+1

end loop

5.4 Initializing the Covariance Matrix and Noise Distribution

It is clear from the update equations for the covariance matrix, the PK parameter means, as well as the noise distribution parameters, that the prior distribution for the PK parameters, i.e. $MVN\left(m_o, \Sigma_0\right)$, and the prior distribution for the noise precision ϕ, i.e. $Ga\left(\phi; s_o, c_0\right)$ are key constituents in determing the final estimates of noise and final covariance matrix for the PK parameters. Thus, it is important for us to arrive at some form of principled procedure for determining these prior distributions.

The prior means for the PK parameters are relatively easy to fix, since we already have some knowledge about the physiologically plausible bounds for the values of k_{ep} and v_e. The values for T_{10} are determined in a completely independent manner (using the variable flip-angle sequences), hence we can use these values as the prior mean of T_{10}. Since the procedure for acquiring the dynamic MRI scans in this study dictates

that the contrast agent is injected after four initial acquisitions, the prior mean for the onset time (t_0) can be set to be the same as the time of injection. The mean intensity of the pre-contrast volume (X_0) can be set to be the actual intensity observed in the time-frame immediately before the arrival of the contrast agent, or the average of the four volumes acquired before injection of the contrast agent. The data being analysed in the VB algorithm consists only of the images acquired after injection of the CA, therefore, the four initial acquisitions can be treated as independent information, exactly as in the case of the flip-angle images.

Although this approach is not strictly fully Bayesian, since the priors are not independent of all data, it avoids problems associated with Empirical Bayes [Berger, 1985], since the data that is being used to inform the priors is not the same as the data being analysed in the VB algorithm. In future, this approach could be replaced by a fully Bayesian approach, though this would lead to a substantial increase in the computational time.

While the prior means for each of the PK parameters can be assigned in the manner outlined above, the prior variances for each of these parameters are considerably more difficult to determine, since they require some estimate of how wide these values are likely to be spread. Since the physiologically plausible values of k_{ep} and v_e are bounded between $[0, 1]$, the means were set to 0.5, and the prior standard deviation was fixed as $(1/3)*0.5$, in order to ensure that 98% of the values would lie in the interval $[0, 1]$. These experiments were also repeated with higher values for the prior standard deviation (0.5 and 1), but we found that the results remained almost unchanged.

In order to set a prior variance for T_{10}, we took a measure of the residuals in the curve-fitting process while ascertaining the T_{10} map. As explained in Section 2.3, the T_{10} map can be calculated from the variable flip-angle sequences using the following equation:

$$\frac{S}{\sin(\alpha)} = R \cdot \frac{S}{\tan(\alpha)} + M_0 \exp(-TE/T_2^*) \cdot (1 - R) \tag{5.19}$$

Here $R = \exp(-TR/T_{10})$. Since this is an equation of the form $y = mx + c$, we can calculate the values of m and c from two or more data points, and then use those to surmise that $T_{10} = -TR/\log(m)$. Once we have calculated the value of R and the intercept c, we can calculate the value of the term $M_0 \exp(-TE/T_2^*)$. Now, if we fix the value of $M_0 \exp(-TE/T_2^*)$, and then re-calculate the values of R (and hence T_{10}) corresponding to each individual flip-angle, we can thus arrive at a measure of variance for the T_{10} values corresponding to each voxel. In order to have a constant value of the prior variance across all patients, the average variance in T_{10} across all voxels in all scans was calculated, and then fixed at 5 milliseconds.

Similarly, for the intensity of the first volume, since there are four MRI acquisitions before the injection of the contrast-agent in the area of interest, we used the variance in the intensities recorded at each of these time-points. This was essentially an estimate of the noise in the images before the contrast agent was injected, and was found to be of the same order for all scans acquired in the same scanner under similar conditions. In order to maintain consistency in the analysis of different scans, this was fixed at 10^{-2}, which was the average of the variances observed in the first four acquisitions in each voxel across all scans after the MRI intensities had been scaled between 0 and 1.

As far as the onset time (t_0) is concerned, it was reasonable to assume a common value of variance across all voxels since t_0 is a function of the vascular distance between the area of interest and the point of injection of the CA. Before analysing the DCE-MRI scan, there is no way of knowing how the vasculature may vary from one voxel to the other, and thus there is no clear and principled way of estimating different values of t_0 for each voxel *before* running the VB algorithm. Thus, the variance for t_0 was determined from pilot studies, and fixed for the the scans that were being analysed.

For the clinical data, the variance was determined to be 0.2 mins, and for pre-clinical data, it was found to be 0.12 mins.

In addition to the variances for each individual parameter, the prior covariance matrix Σ_0 is also meant to contain the prior covariances between each pair of parameters. However, since there is no biological reason for assuming that any of the PK parameters, the first volume intensity or T_{10} are functionally related to each other, we set the initial estimates for all covariances to 0. The final estimates for covariances between different parameters are obtained through the VB updates calculated during the optimization. As we shall see later, these covariance estimates can be quite informative about the relationships between different PK parameters that were considered *a priori* to be unrelated. It is also worth noting that the initialization for the posterior is set to be the same as the priors. Thus, the optimization process starts with the initial estimate being the same as the priors.

5.5 The Impact of Motion Correction on Uncertainty

The first question that we investigated, was whether having motion correction within the VB framework had any impact on the uncertainties that are estimated for each PK parameter. The VB framework described in the previous section provides a principled framework to obtain a quantitative estimate of the extent to which motion is responsible for the uncertainties in the PK parameter estimates.

From Equation 5.16, it is clear that the uncertainties corresponding to each PK parameter calculated in the VB framework, are a product of two main factors (if we neglect the initializations and priors for the time being). The two factors are s, the current estimate of the noise parameter, and $J^T J$, which is the inner product of the Jacobians with respect to the PK parameters. Since c remains fixed, a higher value of

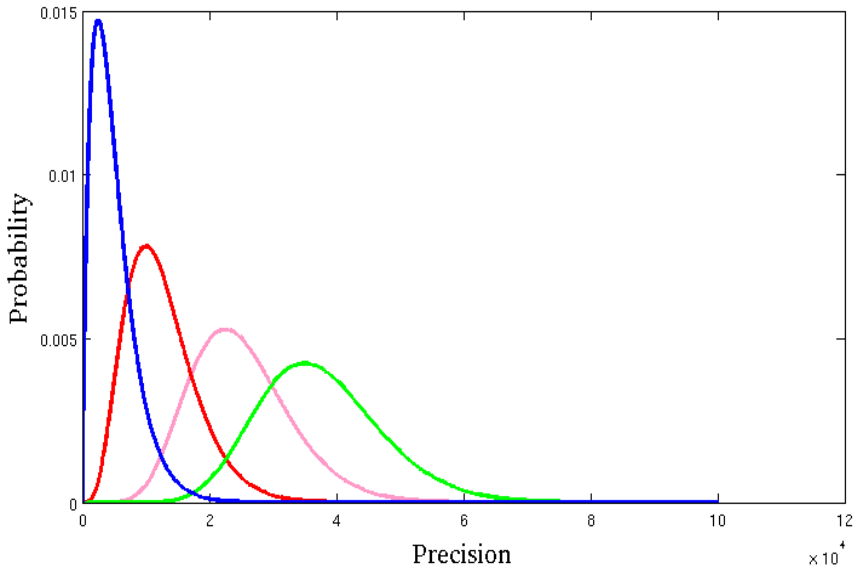

Figure 5.1: Gamma probability distributions on the noise precision ϕ. Each curve represents $Ga(\phi; s, c)$, with c held constant at 10, and s varying. The blue curve represents $s = 0.1$, the red curve represents $s = 0.2$, the pink curve $s = 0.5$ and the green curve $s = 1$.

s implies that the noise precision ϕ is likely to be higher, as demonstrated in Figure 5.1. The noise parameter s, in turn is driven by an additional factor $(k^T k)$, as shown in Equation 5.18. It is worth noting that the precisions (Σ) of the PK parameters are positively correlated with s, while s is negatively correlated with the residual error between the model-fit and the actual data $(k^T k)$. Thus, to get an intuitive idea of how motion correction is likely to affect the precisions or uncertainties of the PK parameter estimates, it is sufficient to note that since the residual is likely to be reduced by the process of motion correction (as shown in Chapter 4), the precisions for the PK parameters are likely to increase. The change in the residuals is not directly proportional to the change in precisions, of course, since the Jacobians with respect to the PK parameters and the final covariance matrix itself are also a part of the equations. However, this is an intuitive way for understanding how the two are related.

In the following sections, we will examine how the uncertainties associated with

each individual PK parameter are impacted by the process of motion correction. The analysis has been performed on two distinct datasets, in order to avoid any bias in the results. We will first present the results obtained from a pre-clinical study, followed by results obtained on the clinical RIT study, which has been described earlier in Chapter 4.

5.5.1 Tests on Simulated Data

The first objective of the tests on simulated data, was to determine whether the parameter maps recovered by the VB algorithm are an accurate estimation of the true parameter maps, after taking uncertainty into account. In other words, we wanted to examine whether the uncertainty estimated by the VB algorithm was able to account for the observed differences between the true and the estimated parameter maps. Another question that we sought to answer was how different the values determined by MoRPE and the VB algorithm were for the same datasets.

In order to answer this question, we replicated the same simulated experiment that has been described in Chapter 2, where different parameter maps were generated, contrast enhancement was simulated on the basis of those parameter maps, and then synthetic motion was added to each dataset. In addition to the synthetic k_{ep} and v_e maps, we also used an artificial t_0 map – in which the values were randomly assigned to all voxels from a Gaussian distribution with mean 1 min and standard deviation 0.2 mins. After simulating the contrast uptake and the motion, we applied the VB algorithm, with motion correction, and computed the standard deviations associated with each of the three parameters.

Table 5.1 shows that for all three parameters, the actual error between the ground-truth parameter values and those estimated by the VB algorithm or MoRPE algorithm are usually less than one standard deviation, and in all cases less than two standard

Measure / **PK Parameter**	Error Before Registration	Error after MoRPE	Error after VB	σ from VB	Difference between MoRPE and VB
k_{ep}	0.452	0.037	0.032	0.027	0.008
v_e	0.121	0.021	0.019	0.031	0.005
t_0 (min)	0.34	0.09	0.04	0.14	0.07

Table 5.1: Average results (over 8 datasets) for motion correction and parameter estimation on simulated data: The 'error' denotes the average value of the absolute difference between the parameter maps used to generate the data, and those estimated by the VB algorithm/MoRPE. The second column shows the errors calculated before registration, the third column shows the errors calculated after running MoRPE, the fourth column shows the errors between the original PK parameters and thse estimated by the VB algorithm, the fifth column presents the standard deviation determined for each parameter using the VB algorithm, and the last (sixth) column shows the average absolute differences between the MoRPE and VB estimates.

deviations. This shows that the uncertainties being estimated by the VB algorithm are in the same range as the actual errors of either algorithm on simulated data. In other words, the standard deviations estimated by the VB algorithm are a good estimate of how 'reliable' the estimates for the PK parameters are.

The results also show that the estimates of the VB algorithm and MoRPE are very similar to each other – i.e. the differences between the two estimates are much less than the errors associated with those two estimates. There are many differences between the optimization frameworks followed in the MoRPE and VB algorithms, but these results invoked confidence that the two algorithms are calculating nearly the same results when analyzing the same data.

Another significant and encouraging implication of the results in Table 5.1, was that the errors between the actual and estimated PK parameters are significantly reduced after applying motion correction using MoRPE or the VB algorithm.

5.5.2 Pre-clinical Data

In addition to the simulated data, we acquired pre-clinical data for a mouse with a xenograft tumour imaged on a 4.7T Agilent VNMRS MRI scanner. The mouse was anaesthetised with iso urane (1-3%) in air and its temperature maintained at 37 degrees Celsius using a rectal temperature feedback system. For fast imaging, a T_1-weighted 3D gradient echo sequence with RF and gradient spoiling was used. A $128 \times 64 \times 64$ voxel matrix covered the entire volume of the RF coil and gave a nominal isotropic resolution of 0.42 mm. For the dynamic scan, TR and TE were 1.11 and 0.585 ms respectively, the flip angle was 5 degrees, and MRI volumes were acquired every 6 seconds. A pre-contrast T_1 map was generated using a variable flip angle method with nominal flip angles of 2, 4, 6, 8, 10 and 12 degrees. The contrast agent Omniscan is injected in a bolus with a dose of 1.2 ml/kg. Omniscan has relaxation rates of $r_1 = 3.9$ and $r_2 = 4.3$.

The tumour xenografted onto the mouse was quite large in proportion to the size of the mouse. Furthermore, since the mouse used in this study was imaged five days after the tumour had been xenografted, the tumour had grown and occupied a substantial part of the mouse body at the time of imaging. Thus, in the Figures shown below, the tumour occupies most of the visible volume, and the background has been masked out with the help of a delineation made by a trained clinician.

For the estimation of the PK parameters and the motion-parameters, we used a population-averaged bi-exponential AIF for mice (analogous to the Weinmann AIF for humans) expressed as Model B in [McGrath et al., 2009]. All other components of the VB algorithm (with and without motion correction) remained the same as have been described in sections (5.3) and (5.4). In the following section, we will analyse the PK parameter maps obtained with and without using motion correction, and also scrutinize changes in the variances associated with each parameter estimated by the VB algorithm.

5.5.2.1 Results on Pre-clinical Data

Figure 5.2 shows the k_{ep} maps obtained using the VB algorithm with and without motion correction. The Figure indicates that when motion correction is not used, the values for many voxels remain close to 0.5, which is the prior mean, and the starting estimate for the posterior. After applying motion correction, the values in many of the same voxels are either increased or decreased. This indicates that in the absence of motion correction, the VB algorithm sometimes does not have enough information to calculate accurate estimates. This is especially the case at tissue boundaries, which are more likely to be affected by motion.

The variances as calculated by the VB algorithm for k_{ep} are as shown in Figure 5.3. It is clear from this Figure that the estimates of variance are reduced after motion correction in most voxels, and particularly in voxels where the estimate of k_{ep} itself has changed from 0.5 to some other value. Once again, the reduction in the variances of the k_{ep} estimates is most pronounced at the tissue boundaries. As seen in Equation 5.20, the variance on k_{ep} represents the uncertainty in the shape (i.e. the slope of the uptake and the washout) *and* amplitude of the curve, whereas the variance in v_e indicates uncertainty in the amplitude of the contrast take-up curve.

The maps obtained using the VB algorithm for v_e are shown in Figure 5.4. Once again, it can be seen that when motion correction is used, the v_e estimate for many voxels moves away from the prior mean (0.5). Figure 5.5, which shows the variances obtained for the estimates of v_e, confirms that the uncertainties (variances) are also reduced for the same voxels. It is worth noting that although the variances estimated for v_e are reduced in the tissue boundaries as well as the interior regions as a result of motion correction, they are still higher in the tissue boundaries as compared to the interior regions.

Figures 5.3 and 5.5 show that the variances obtained for v_e are smaller in magnitude

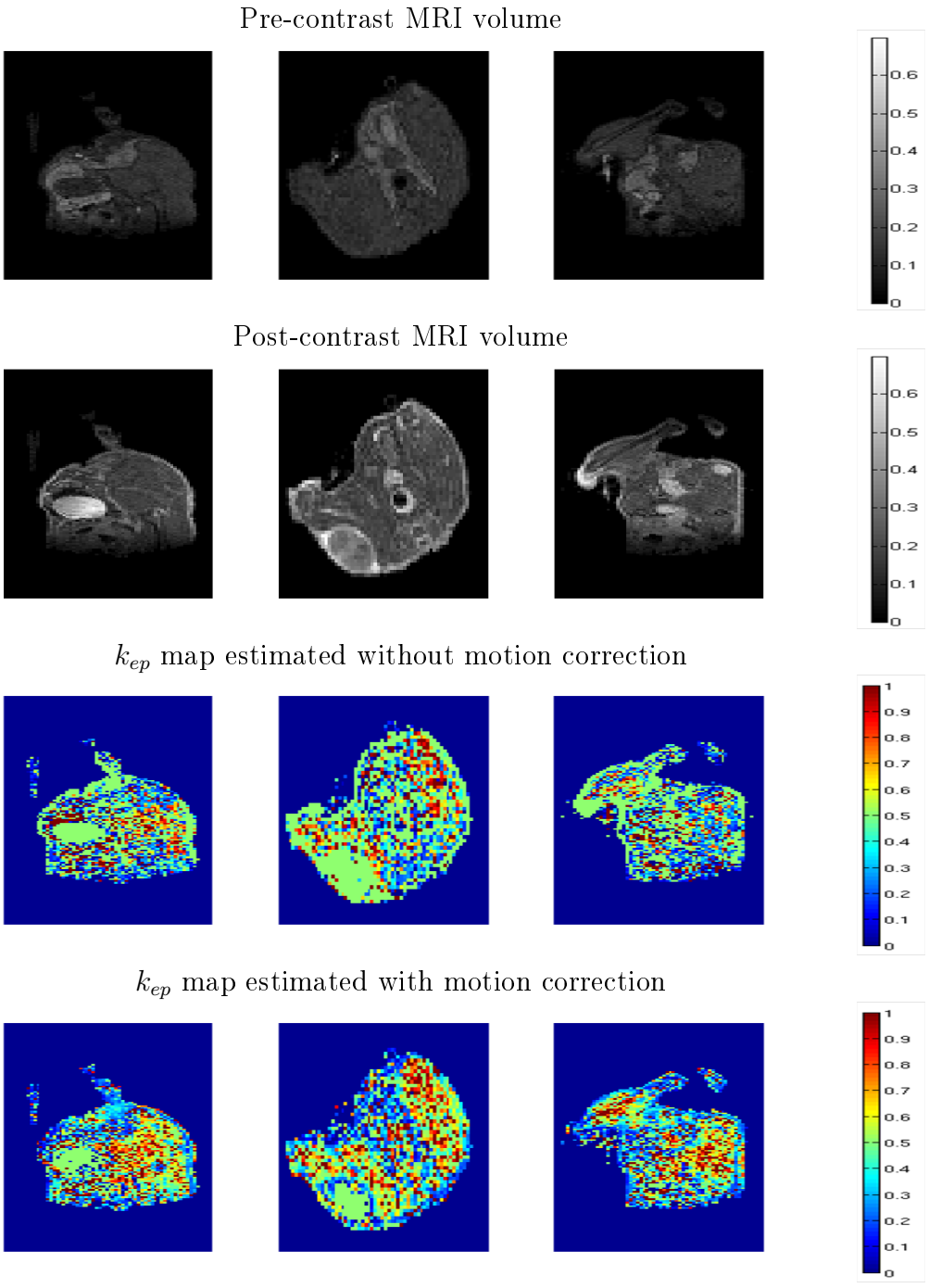

Figure 5.2: Axial (left column), coronal (middle column) and sagittal (right column) slices through the pre-contrast volume (top row), the post-contrast volume (2nd row), the k_{ep} maps for the pre-clinical data without motion correction (3rd row) and with motion correction (bottom row).

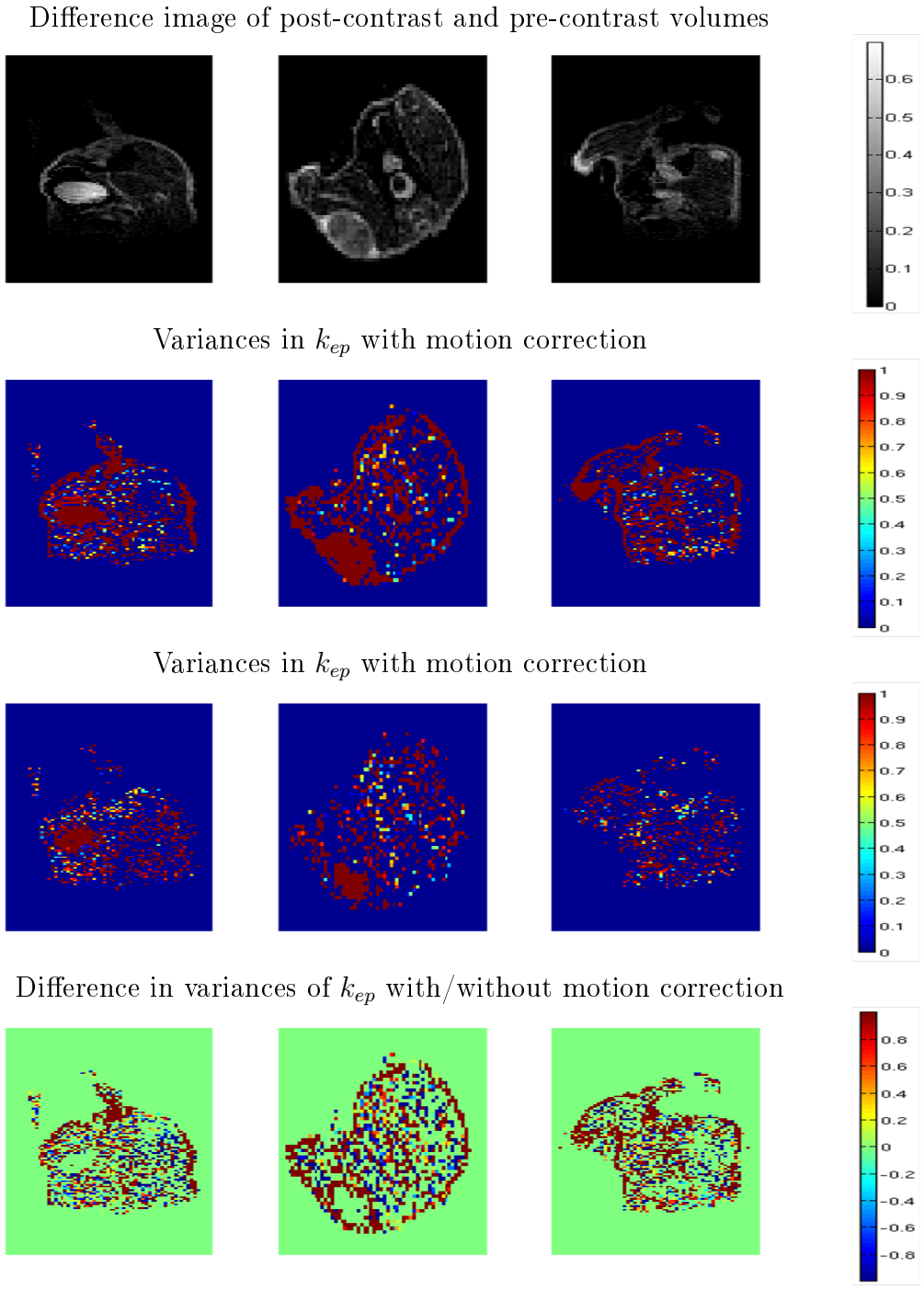

Figure 5.3: Axial (left column), coronal (middle column) and sagittal (right column) slices through the difference image of the post-contrast and pre-contrast volumes (top row) maps showing the variance in k_{ep} for the pre-clinical data without motion correction (2nd row) and with motion correction (3rd row), and the difference between them (bottom row).

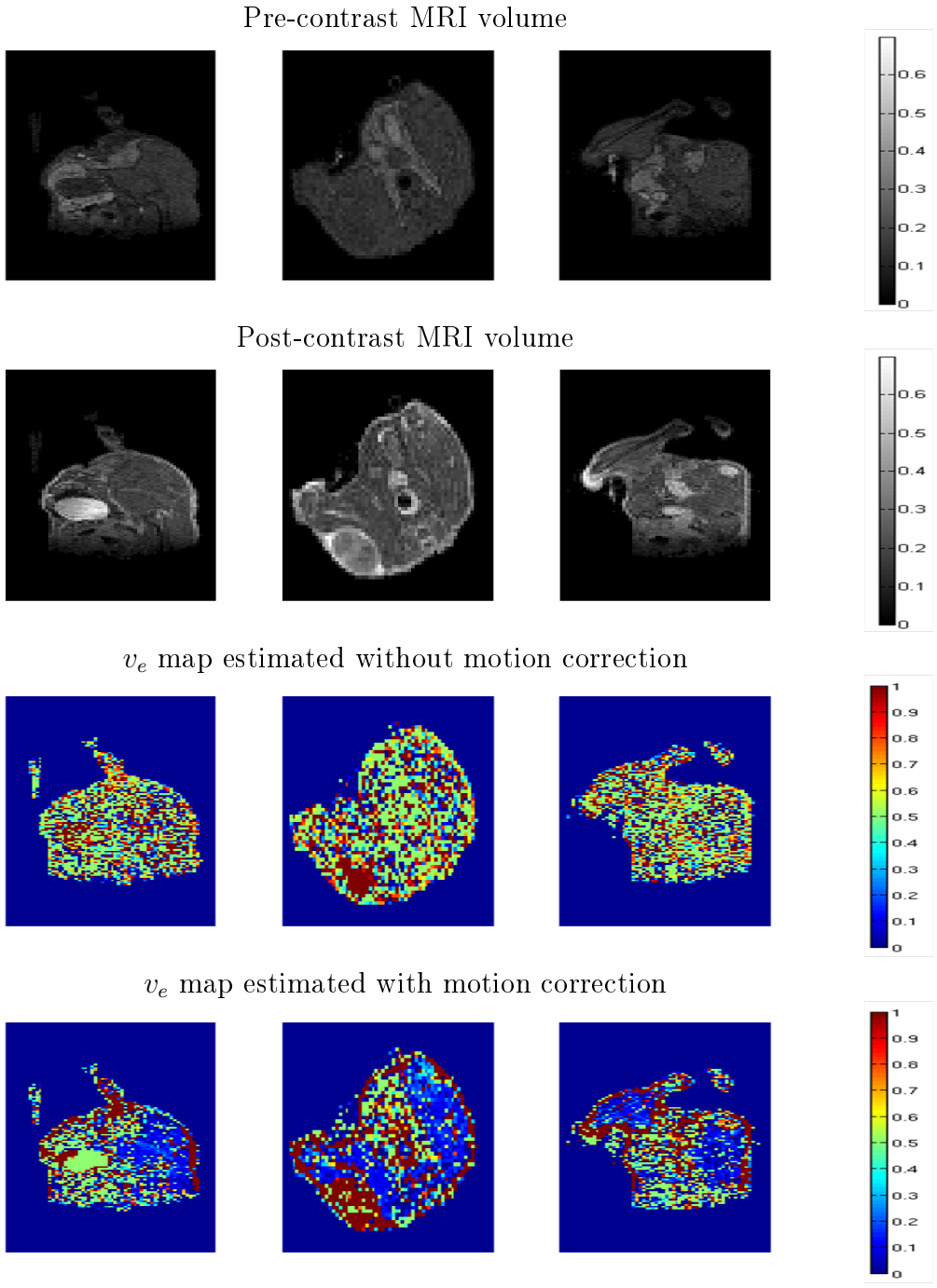

Figure 5.4: Axial (left column), coronal (middle column) and sagittal (right column) slices through the pre-contrast volume (top row), the post-contrast volume (2nd row), the v_e map estimated without motion correction (3rd row) and with motion correction (bottom row).

than the variances for k_{ep}. One possible explanation for this observation, is that the variance in the estimate of k_{ep} reflects a combined effect of uncertainty in the shape and amplitude of the curve, while the variance in v_e reflects uncertainty only in the amplitude of the curve. Thus, the former is likely to be higher.

In [Garpebring et al., 2012], one of the most significant contributors to uncertainty was found to be the bolus arrival time (BAT), or what we have referred to as the 'onset time'. In some papers, the onset time has been shown to be strongly correlated with areas of necrosis within the tumour [McPhee, 2012]. Thus, the maps for onset time also contain information that is potentially meaningful.

Figure 5.6 shows the maps for t_0 obtained from the VB algorithm with and without motion correction. It is important to note that the area with low t_0 (towards the bottom left of the middle frame) coincides with the area of high v_e, low k_{ep}, and low X_o, which are together indicative of a flat, low intensity curve, probably representing an area of necrosis within the tumour. The same area also shows a high variance for k_{ep} and t_0, thus showing that the estimates for this area have a higher degree of uncertainty than other areas.

Figure 5.7 shows the variances for onset time (i.e. the uncertainties in the estimates of onset time) as estimated by the VB algorithm. This Figure again indicates that the variances in the t_0 estimates are reduced substantially when motion correction is used, and the greatest reduction is observed at the tissue boundaries. However, it is also noticeable that the region with low t_0 maintains a high value for the variance with and without motion correction. This indicates a high level of uncertainty in the t_0 estimation for the necrotic areas of the tumour.

One other quantity that is treated as a model parameter in the VB algorithm framework is the intensity of the first time-point image (i.e., the baseline intensity – X_0). In MoRPE, the first time-point image is fixed as the reference frame, and the deforma-

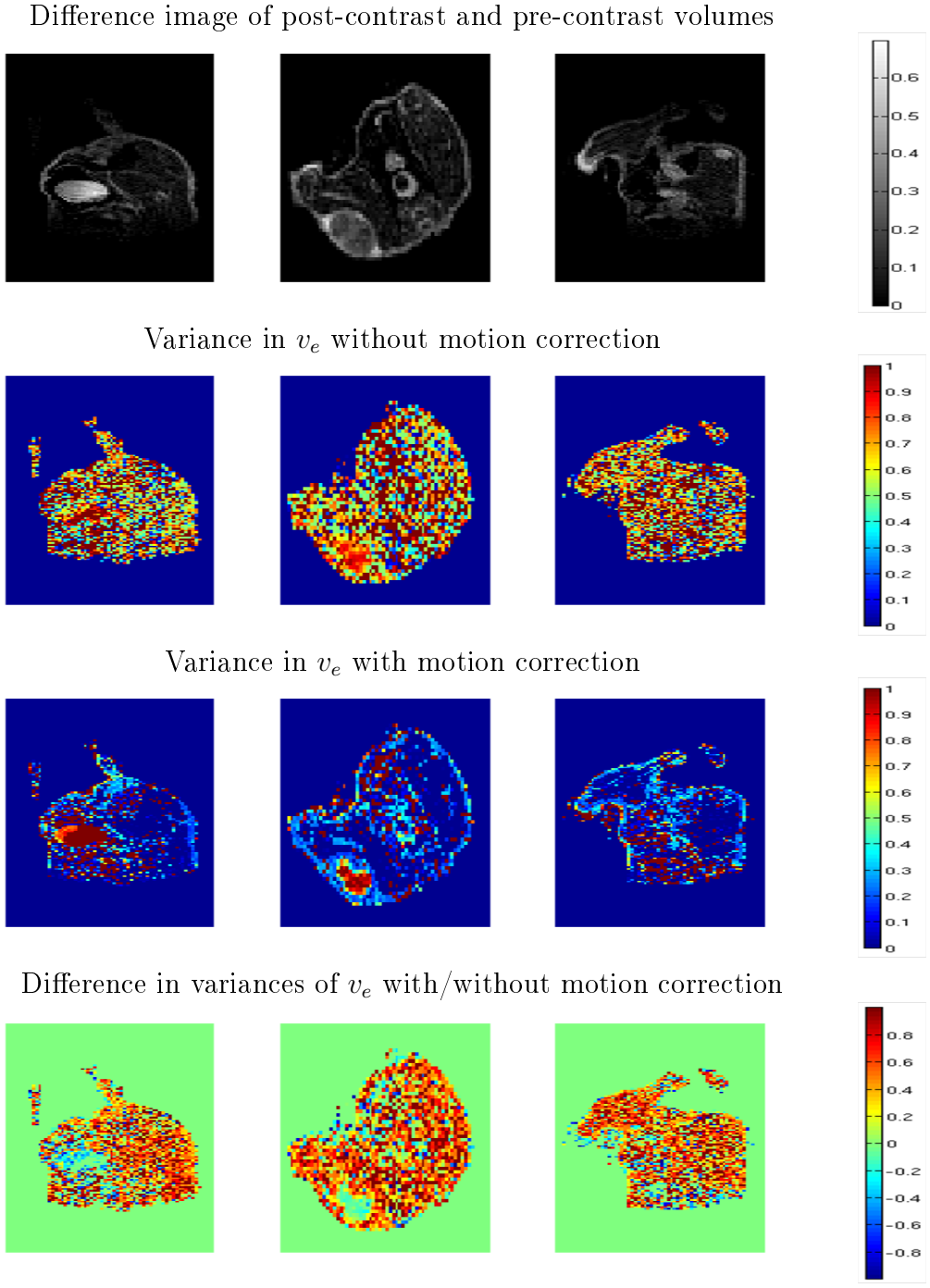

Figure 5.5: Axial (left column), coronal (middle column) and sagittal (right column) slices through the difference image of the post-contrast and pre-contrast volumes (top row), the maps showing the variance in v_e for the pre-clinical data without motion correction (2nd row) and with motion correction (3rd row) and their difference (bottom row).

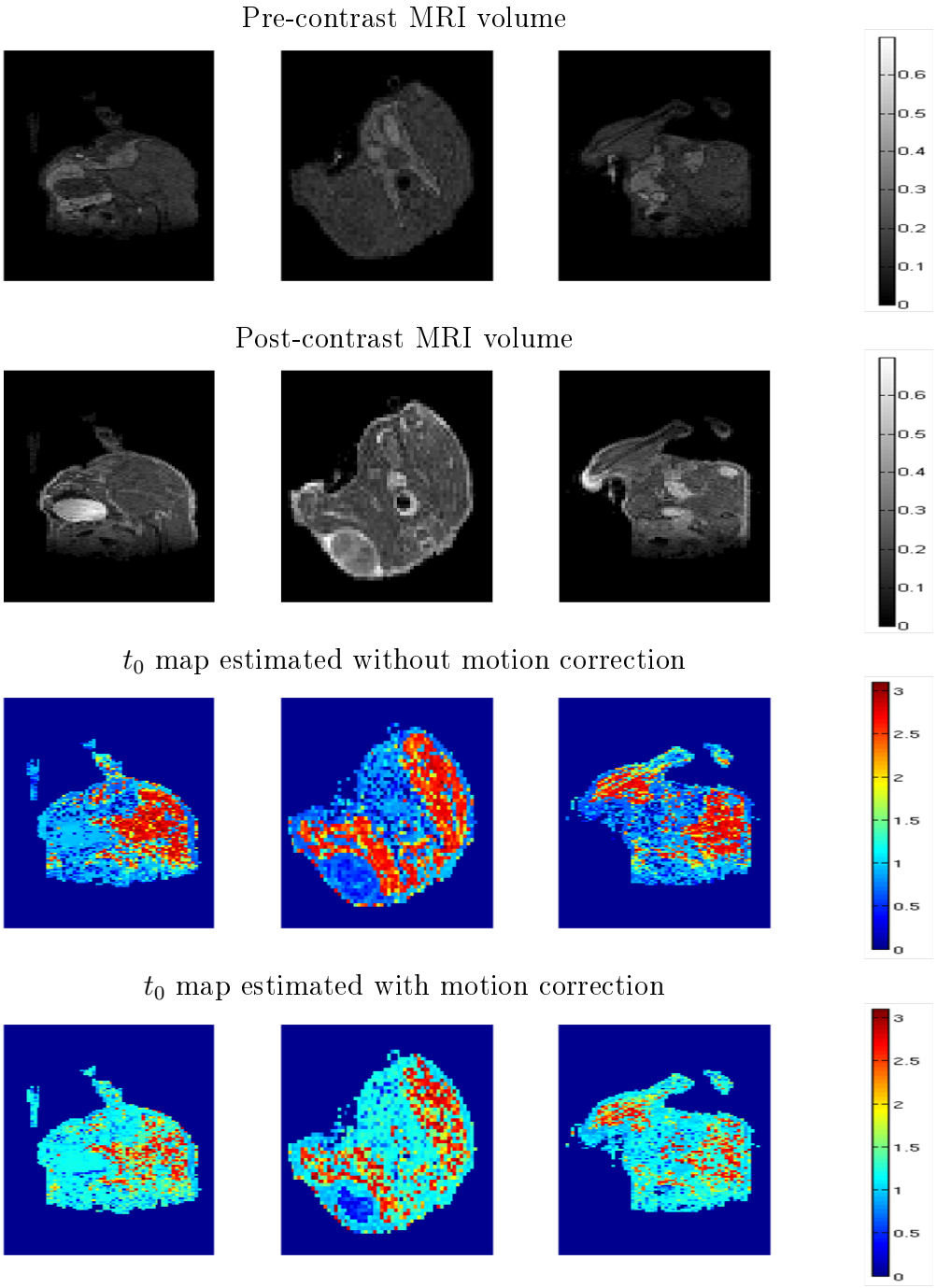

Figure 5.6: Axial (left column), coronal (middle column) and sagittal (right column) slices through the pre-contrast volume (top row), the post-contrast volume (2nd row), the volume showing the values for t_0 (in mins) estimated without motion correction (3rd row) and with motion correction (bottom row).

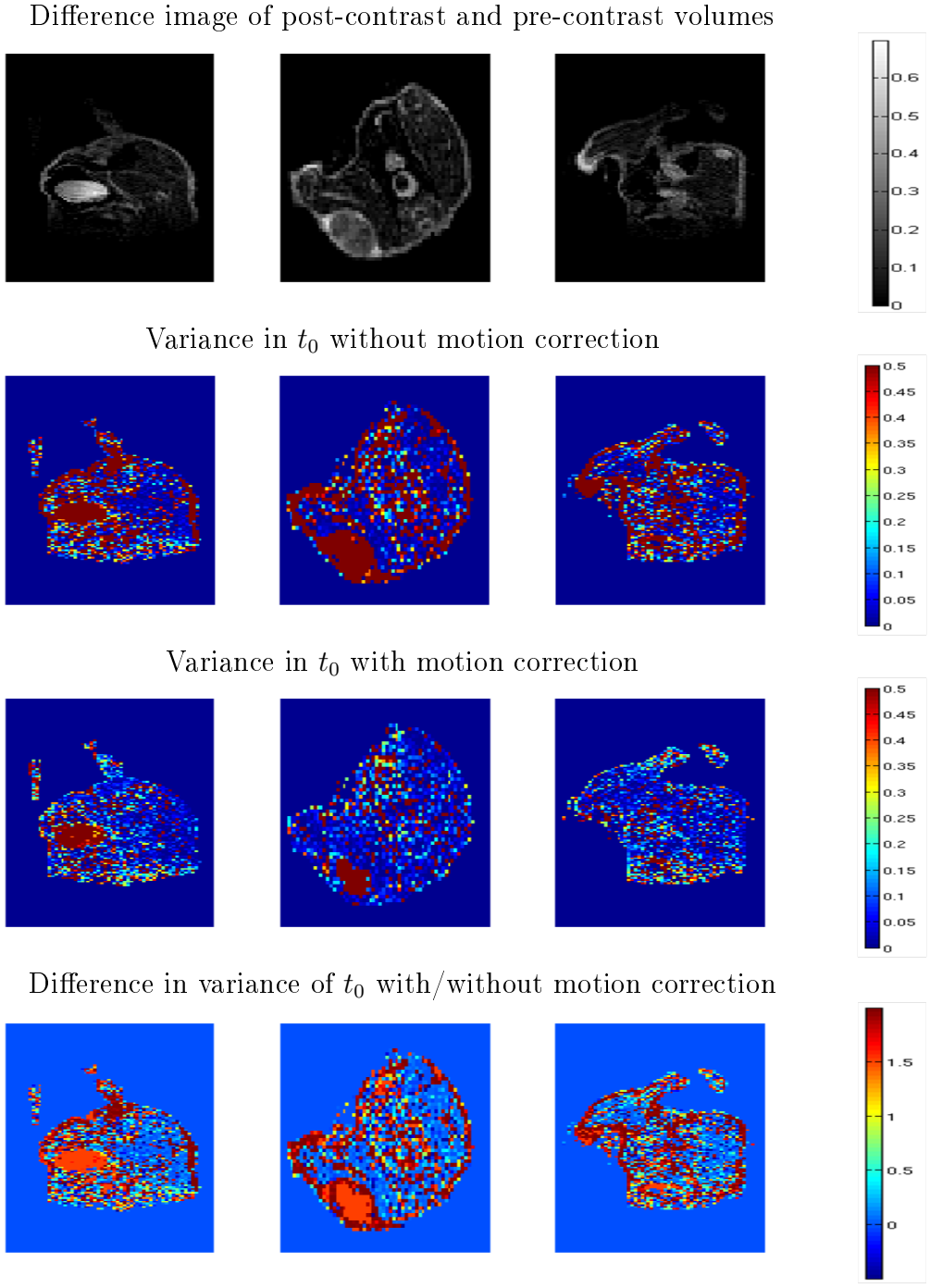

Figure 5.7: Axial (left column), coronal (middle column) and sagittal (right column) slices through the difference image of the post-contrast and pre-contrast volumes (top row), the maps showing the variance in t_0 for the pre-clinical data without motion correction (2nd row), with motion correction (3rd row), and their difference (bottom row).

tions that need to be applied are estimated with respect to this frame of reference. In other words, the transformations (T) estimated in MoRPE are applied to the other time-points in order to bring them into alignment with the baseline image.

The model prediction that is calculated at each iteration of MoRPE is also based on the intensities of each voxel in the first time-point image, and of course the PK parameter estimates. In the MATLAB inplementation of MoRPE, the first time-point intensities remain fixed, since no transformation is applied to the baseline image, and since X_0 is not treated as a model parameter. However, in the VB algorithm, the intensities of the first time-point image are also treated as one parameter of the model (in the same way that k_{ep}, v_e, T_{10} and onset time are other parameters of the model), and by setting a value for the prior variance as explained in Section 5.4 ($\sigma = 0.02$ for an image of size $128 \times 64 \times 64$ with intensities scaled between 0 and 1), the first time-point volumes is updated during the running of the VB algorithm in order to achieve a better fit to the PK model.

In principle, this is a better approach to follow, since it is entirely possible that some voxels in the baseline image are actually showing a corrupted value for the MRI intensity due to noise or other effects. For such corrupted voxels, keeping X_0 as a model parameter allows the VB algorithm to recover the original MRI intensity by using information from the other time-points. This also allows the baseline image to be deformed, if such a deformation serves to reduce the error between the deformed imageset and the model prediction. Applying motion correction further improves image fidelity, and also restores the image intensities so that they are closer to the values predicted by the model for each voxel. This process naturally reduces the level of uncertainty in the estimates for X_0 at each voxel. As Table 5.2 shows, the variance in the estimates of both T_{10} and X_0 were reduced by 15-20% of their original values as a result of including motion correction within the VB algorithm.

	k_{ep}	v_e	T_{10}	X_0	t_o
CV without motion correction	45.4%	31.1%	6.9%	8.7%	27.6%
CV with motion correction	32.7%	18.5%	5.8%	7.2%	16.3%
Mean reduction in CV	11.7%	12.6%	1.1%	1.5%	11.3%

Table 5.2: The coefficients of variation (CV) – i.e. the standard deviation divided by mean – for each of the five model-parameters, calculated for the pre-clinical data. The measures are calculated with and without motion correction, and the reduction in the CV values is reported in the last row.

5.5.2.2 Conclusions from Pre-clinical Experiments

The tests on the pre-clinical data clearly show that the uncertainty estimates (variances) associated with most of the PK parameters are reduced by virtue of performing motion correction during the iterations of the VB algorithm. In order to demonstrate this quantitively, we calculated the coefficients of variation (CV) – i.e. ratio of the standard deviation to the mean – for each model parameter, with and without applying motion correction. Table 5.2 shows the CVs calculated from the pre-clinical data with and without motion correction.

The variances for k_{ep}, v_e and t_0 are reduced by 30-40% of the values that were estimated without motion correction. This substantial reduction indicates that motion correction is important for calculating precise values of these three PK parameters. This decrease in the variances associated with the PK parameters is most likely caused by a reduction in the model-fit residual $k^T k$ (in Equation 5.18), which in turn influences the updates for the precision matrix, as shown in Equation 5.16.

5.5.3 Clinical Data

After testing the method on the pre-clinical data, we tested MoRPE on clinical data obtained in the rectal imaging trial (RIT) study, which has been reported in the previous chapter. The MRI sequence used to acquire each DCE-MRI scan was a T_1-weighted,

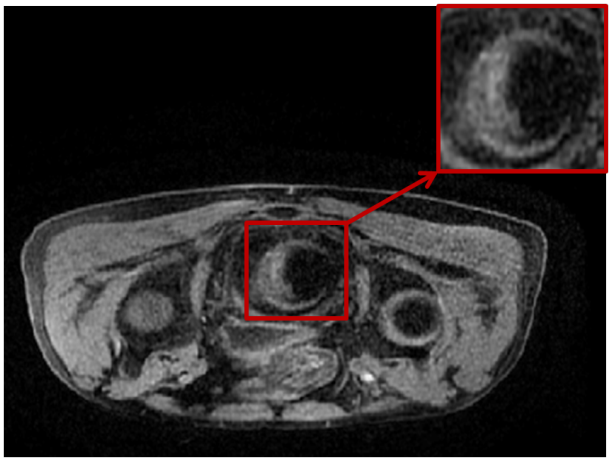

Figure 5.8: An axial slice of a pre-contrast volume, with the tumour ROI shown in the inset.

gradient-echo, fat-suppressed sequence (known as LAVA) with TR=4.5ms, TE=2.2ms and flip angle $\alpha = 12$ degrees. The contrast agent (gadobenate dimeglumine, also known as MultiHance) was injected via a peripheral vein after 4 initial acquisitions, and the MRI images were acquired every 12 seconds for the next 5 minutes. For most patients, 27 volumes were acquired during the DCE-MRI scan, with each volume being of dimensions $512 \times 512 \times 52$, and and each voxel having the dimension $1 \times 1 \times 2\,\mathrm{mm}^3$.

In order to determine the T_{10} map for each patient, MRI volumes were acquired at four different flip-angles (3, 9, 12 and 15 degrees). The variable flip-angle images were first registered to the first volume of the dynamic scan in order to ensure that the T_{10} map was in alignment with the dynamic sequence. The T_{10} map was then estimated from the variable flip-angle images using the procedure described in [Fram et al., 1987], and this map was used in the subsequent analysis.

5.5.3.1 Results on Clinical Data

The results on one example patient are discussed here to demonstrate the analysis process and the typical findings. Figure 5.8 shows an axial slice of the pre-contrast

volume of the clinical sample dataset. The tumour ROI is shown in the inset, and all parameter maps shown below represent this area. The main focus of this study, as in the case of the tests on pre-clinical data, was to assess the impact of motion correction on the PK parameter maps and the uncertainties (variance maps) associated with them, as estimated by the VB algorithm. Figure 5.9 shows that the variances estimates for k_{ep} are substantially reduced in many voxels in the outer areas of the tumour.

Figure 5.10 shows the variance associated with v_e estimated using the VB algorithm with and without motion correction. The reduction in the variances associated with v_e is more significant in the peripheral regions of the tumour, and is also relatively greater than the reduction in the variances associated with k_{ep}. This implies that motion correction is more effective in reducing the uncertainty in the amplitude of the curve, rather than the shape of the curve (that is governed by k_{ep} alone).

If we recall the equation for the concentration of contrast agent as expressed in the Tofts model is given by:

$$C(t) = AIF(t) \otimes k_{ep} v_e e^{-k_{ep} t} \tag{5.20}$$

$$= AIF(t) \otimes K^{trans} e^{-\frac{K^{trans}}{v_e} t} \tag{5.21}$$

It is more useful to consider the concentration of the contrast agent in terms of Equation 5.20, since it is clear from this equation that the parameter v_e only impacts the amplitude of the concentration time-curve, whereas the shape of the curve is affected only by k_{ep}. Thus, the uncertainty in these two parameters can loosely be interpreted as the uncertainty in the amplitude and shape of the curve (although it is important to remember that the uncertainty in k_{ep} reflects a combination of both).

After testing the method on one example patient from the RIT study, we performed

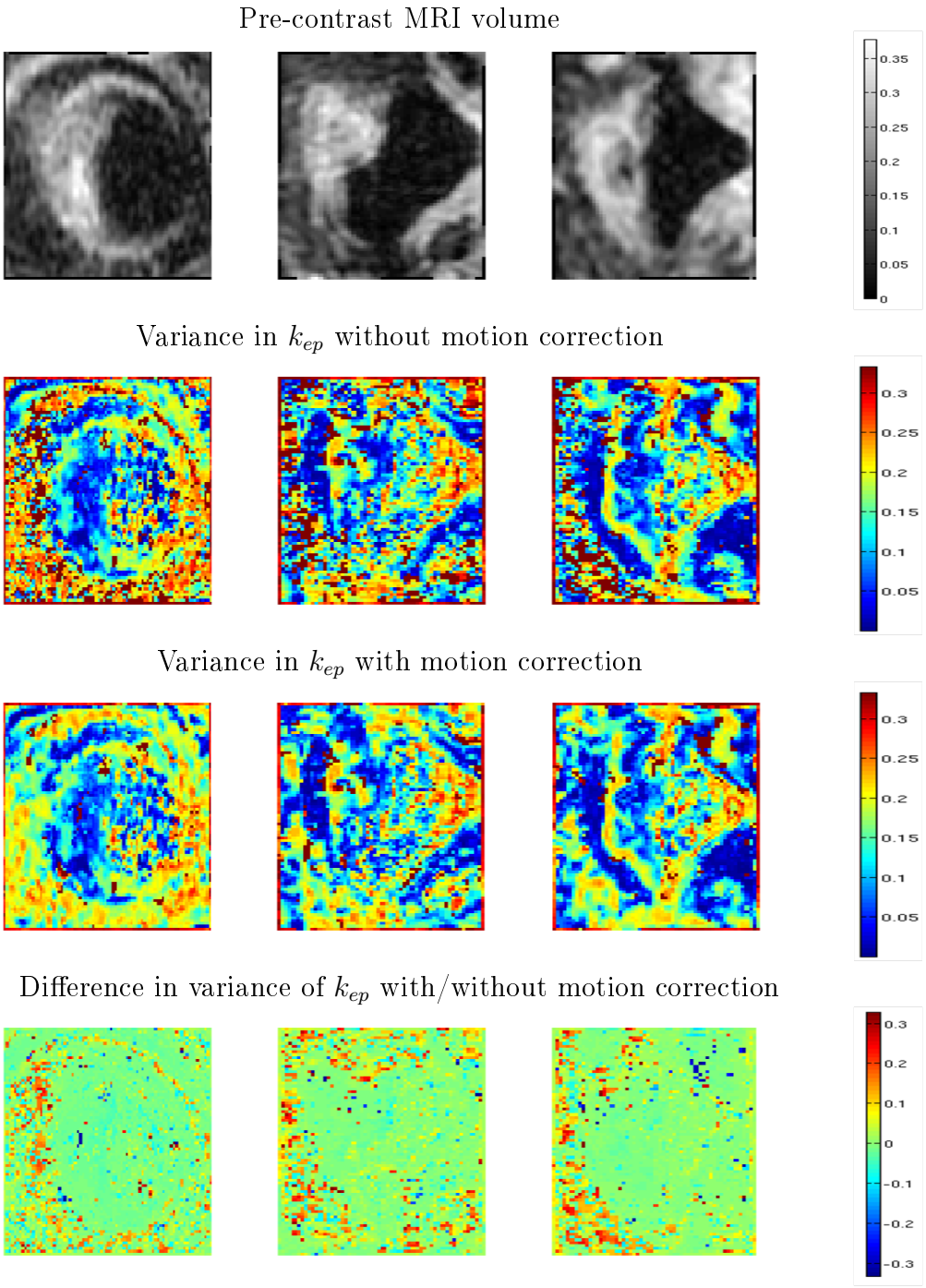

Figure 5.9: Axial (left column), coronal (middle column) and sagittal (right column) slices through the pre-contrast MRI volume (top row), and the maps showing the variance in k_{ep} for clinical data, estimated without motion correction (second row), with motion correction (third row), and their difference (bottom row).

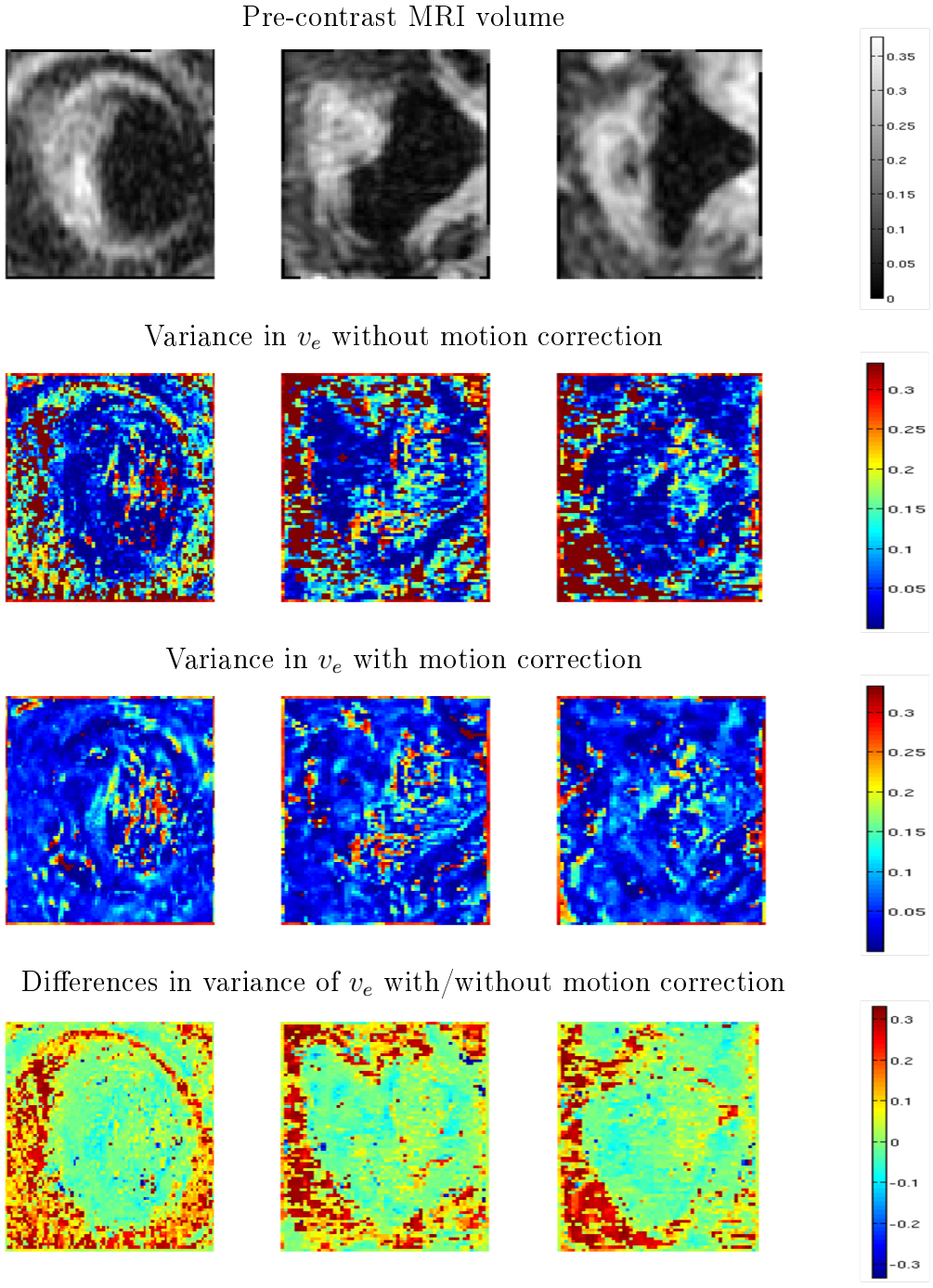

Figure 5.10: Axial (left column), coronal (middle column) and sagittal (right column) slices through the pre-contrast MRI volume (top row), and the maps showing the variance in v_e for clinical data, estimated without motion correction (second row), with motion correction (third row), and their difference (bottom row).

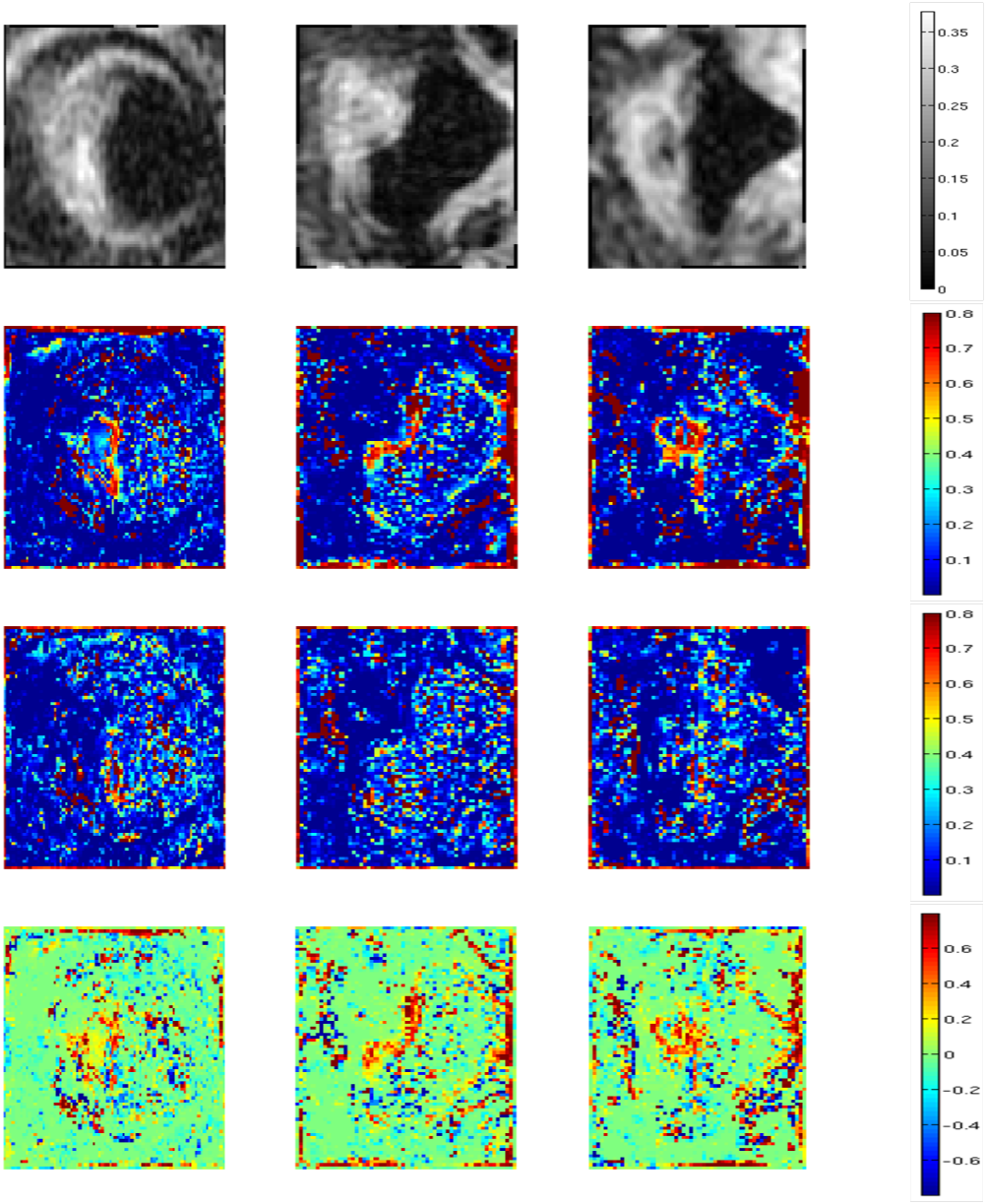

Figure 5.11: Axial (left column), coronal (middle column) and sagittal (right column) slices through the pre-contrast MRI volume (top row), and the maps showing the variance in t_0 for clinical data, estimated without motion correction (second row), with motion correction (third row), and their difference (bottom row).

	k_{ep}	v_e	T_{10}	X_0	t_o
CV before	25.1%	23.4%	8.6%	12.5%	21.5%
CV after	12.9%	17%	7.8%	10.2%	15.8%
Mean reduction in CV	12.3%	6.4%	0.8%	2.3%	5.7%
p-value	10^{-5}	0.004	0.01	0.007	0.002

Table 5.3: The coefficients of variation (standard deviation divided by mean) for each of the five model-parameters, averaged across all 20 patients. The measures are calculated before and after motion correction, and the p-values are estimated by performing the two-sample student t-test between the CV values for 1000 corresponding voxels picked out from each scan with and without motion correction.

a similar analysis in the pre-therapy scans for the other 19 patients from the RIT study. Apart from examining the distributions and variance maps for k_{ep}, v_e and t_0, we also examined the impact of motion-correction on the variances estimated for the other PK parameters – T_{10}, X_0.

5.5.3.2 Conclusions from Tests on Clinical Data

In our tests on clinical data for all twenty patients, we found that there was a consistent reduction in the variance values estimated for most of the PK parameters in the study. Instead of calculating the variances associated with each parameter, we calculated the ratio of the standard deviation to the mean (known as the coefficient of variation – CV) for each voxel, and then took the average of the CVs over all the voxels across the twenty scans, to get an aggregate value for each PK parameter. In order to determine whether the reductions in CV were statistically significant, we also tested the voxelwise values, without averaging, by performing a paired t-test between the CV values calculated with and without motion correction, for 20,000 voxels (1000 corresponding voxels being picked randomly from each scan). The results obtained for this entire analysis are presented in Table 5.3. Since the tumour ROIs had a size of about $100 \times 100 \times 52$, 1000 voxels were chosen randomly from each scan to avoid comparisons between too many neighbouring voxels in the pairwise t-test.

From Table 5.3, it is clear that there is a statistically significant ($p < 0.05$) reduction in the coefficients of variation corresponding to all five PK parameters due to motion correction. The greatest reduction takes place in the CV values of k_{ep} and t_0. These results indicate that motion correction is important for obtaining more precise estimates of PK parameters from clinical data. The reductions in the CV values for T_{10} and X_0 were comparatively lower than the reductions in the other three parameters, but still significant, when assessed on a voxel-by-voxel basis.

5.6 Different Causes of Uncertainty

One of the key questions that has been examined when looking at uncertainty in parameters derived from DCE-MRI, is the contribution that different factors make to the lack of precision in the PK parameters of interest. The issue of uncertainty in DCE-MRI has been analyzed in detail in several papers [Garpebring et al., 2012], [Buckley, 2002], [Schabel and Parker, 2008]. The most comprehensive analysis of uncertainties in different PK parameters has been carried out in [Garpebring et al., 2012], and there are many similarities in the VB framework and the mathematical framework used in that study. However, [Garpebring et al., 2012] divides the parameters into two classes – (1) the physiological parameters $p = [K^{trans}, v_e, v_p]$, and (2) the imaging parameters $q = [t_0, T_{10}, X_o, AIF]$. Here, the physiological parameters are assumed to be the real parameters of interest, and the uncertainty in their estimates is examined as a function of the noise in the imaging parameters. Although this analysis is very illuminating, it does not take motion into account, and does not allow us to determine the extent to which uncertainties in K^{trans}, v_e or t_0 contribute to the overall error between the signal and the PK model. The framework in [Garpebring et al., 2012] also does not enable a comparison of the uncertainty in K^{trans} with, say, the uncertainty in t_0 or T_{10}.

The entire rationale behind PK modeling of DCE-MRI data, is that the contrast uptake curves can be expressed as functions of some known and unknown parameters. However, in practice, even after the optimal values for the PK parameters have been determined, there is always some degree of 'model-fit error' or 'residual' between the model prediction and the actual data. This residual error could be caused by many factors – patient motion during the scan, thermal noise in the observed MRI signal, or simply a model failure. In the previous section, we have discussed the uncertainty/variance associated with the PK parameters. These values for the variance give us an estimate for how reliable the estimates for the PK parameters are.

Apart from this, the VB framework also provides voxel-wise estimates for how well the model represents the data. This information is contained in the estimates for the noise parameter s, which was discussed in Section 5.3. The noise parameter is updated according to Equation 5.18, and the final value for this parameter provides a measure for how noisy or reliable the data itself is, for each individual voxel, and to what extent it is explained by the model. We are interested in determining the principal causes for the estimate of noise in different voxels. In particular, we are interested in ascertaining the extent to which the estimates of noise are driven by motion during the scan, and by uncertainties in particular PK parameters.

5.6.1 Method of Allocating Uncertainties

In the VB framework described above, we treat all parameters equally, and there is no distinction between physiological and imaging parameters. We did not make any distinction because there are studies [McPhee, 2012] to show that the parameters such as the onset time (t_0) also contain important physiological information. There are also studies that show the impact of bias in the T_1-relaxation time affecting the measurements of the concentration of CA [Schabel and Parker, 2008]. Thus, we examined the

combined effect of all these factors on the estimates of noise associated with each voxel, rather than looking at the the effect of the uncertainty in one parameter on the uncertainty in another. From Equation 5.18, it is clear that the estimate for the total amount of noise in each voxel of the image is a combined function of the residuals, along with the Jacobian matrix of the data with respect to the PK parameters. In the VB framework, the noise is modeled as a Gamma distribution with parameters s and c. Since c is held constant, the amount of noise in the dataset is estimated by the shape parameter s.

$$\frac{1}{s} = \frac{1}{s_0} + \frac{1}{2}k^T k + \frac{1}{2}\mathrm{Tr}\left(\Sigma \cdot J^T J\right) \tag{5.22}$$

It is worth recalling here that s does not directly indicate the amount of noise in the dataset, but is only the hyperparameter that governs the distribution of the precision value ϕ, which in turn indicates the amount of noise in the dataset. In other words, the noise in the dataset is given by $\epsilon \sim N\left(0, \phi^{-1}\right)$, and the distribution of the precision parameter ϕ is given by $Ga\left(\phi;\, s, c\right)$. Since the VB framework allows us to determine the hyperparameters of the noise distribution independently for each voxel, we can then estimate the most probable value of ϕ at each voxel as the mode of the Gamma distribution, i.e $\hat{\phi} = (s-1)c$.

In other words, a larger estimate of s at any voxel implies that the noise precision is higher, and the model prediction is likely to be close to the actual data value for that voxel, whereas a smaller estimate of s implies that the error between the model prediction and the actual data is likely to be higher. Thus, the value of s in a particular voxel is a measure of how well the model represents the observed signal in that voxel. As can be seen from Equation 5.22, the estimate of s has two components. The first part is a direct measure of the error between the model prediction and the actual data. The second part denotes the trace of the product of the latest estimate of the covariance

matrix with the latest estimate of the Jacobians with respect to all the parameters.

Intuitively, the contribution of each parameter to the estimate of noise should be a combination of two factors: (1) The actual error between the model prediction and the data; (2) How much the model prediction changes if we change that particular parameter, i.e. the partial derivative of the model prediction with respect to the parameter in question. This is precisely the division between the two parts on the RHS of Equation 5.22. We can actually divide the term $M = \mathrm{Tr}\left(\Sigma \cdot J^T J\right)$ into different parts, each of which denotes the 'contribution' of one parameter to the overall estimate of noise.

Let us assume that our model has n parameters that need to be estimated. In other words, the set of model parameters can be denoted by $\{\theta_i\}_{i=1,..,n}$, and each element of the matrix $J^T J$ can be written as

$$\left[J^T J\right]_{i,j} = \frac{\partial f(\theta)}{\partial \theta_i} \cdot \frac{\partial f(\theta)}{\partial \theta_j} \tag{5.23}$$

Now, M can be expressed as:

$$M \;=\; \mathrm{Tr}\left(\left[\Sigma \cdot J^T J\right]\right) \;= \sum_{i=1,..,n} \left[\Sigma \cdot J^T J\right]_{i,i} \tag{5.24}$$

Each individual element of the trace of M can be represented as the contribution of the i^{th} PK parameter, and we can view the error component corresponding to each parameter θ_i, $Err(\theta_i)$ as:

$$Err(\theta_i) = \sum_j \left(cov\left(\theta_i, \theta_j\right) \cdot \left(\frac{\partial f(\theta)}{\partial \theta_j} \cdot \frac{\partial f(\theta)}{\partial \theta_i} \right) \right) \tag{5.25}$$

This can be interpreted as the contribution of θ_i to the noise estimate $1/s$. However, it is clear from Equation 5.22 that M is not the only component of the noise estimate. The other components are the initialization and the observed residual error between the model prediction and the actual data.

The residual error between the model prediction and the actual data ($k^T k$) can be a product of many factors. Firstly, there is patient motion during the scan, which may not be eliminated completely even after motion correction. Secondly, there is thermal noise in the dynamic sequence itself, which can not be eliminated by motion correction or by optimizing the PK parameters. The other reason could be a model-failure, i.e. the Tofts model may not accurately describe the passage of CA in certain parts of the tissue. Thus, it is not possible to attribute this error to any one of the individual PK parameters. In other words, since the PK parameters are supposed to have been optimized with respect to the error between the model-fit and the actual data, the remaining residual ($k^T k$) can be assumed to be caused by one of the extraneous factors listed above.

The optimization process in the VB framework initially assumes a very uninformative prior on the noise distribution, and hence the term $1/s_0$ can be neglected. Thus, while looking at the inverse of the final noise estimate, i.e. $1/s$, we can divide this into six parts – the contributions of the five PK parameters, as described in Equation 5.25, and the sixth part ($\frac{1}{2}k^T k$) reflecting the contribution of all other factors – motion, thermal noise and model-failure.

5.6.2 Findings on Sources of Uncertainty

There were primarily two questions that we were interested in answering while examining the contributions of different factors to the overall estimate of noise in the dataset:

1. How do the proportional contributions of different factors change due to motion correction?

2. What are the primary sources of noise in the overall analysis? Can we use the analysis of noise contributions to focus our attention on the calculation of mea-

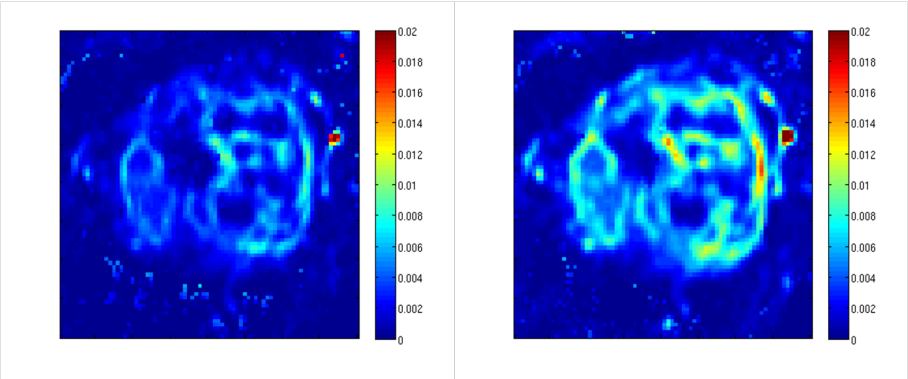

Figure 5.12: The noise parameter s, determined without motion correction (left), and with motion correction (right).

suring some parameters more accurately at the expense of others?

In order to address the first question, we investigated the percentage contributions of different factors to the overall noise estimate, across all voxels in twenty clinical datasets from the RIT study. For each dataset, we examined the error contribution of each of the six factors with and without motion correction. We were primarily interested in analyzing the effect of motion correction on the contribution of the residual error, since the residual error is representative of the error due to motion, and other factors such as thermal noise, etc.

Upon analyzing the change in the contribution of the residual to the total error estimate, we found that motion correction helped to substantially reduce the contribution of this factor, from an average of 30.7% without motion correction (measured across all voxels in the twenty datasets) to an average of 19.1% with motion correction. The factor that seems to compensate for this shift in balance, is the contribution of T_{10} to the overall error estimate. However, it is important to remember that here the shift only the percentage contributions to the total error, and in absolute values, the error estimate itself changes due to the process of motion correction, as shown in Figure 5.12. The fact that the contribution of T_{10} increases, while the contribution of the other fac-

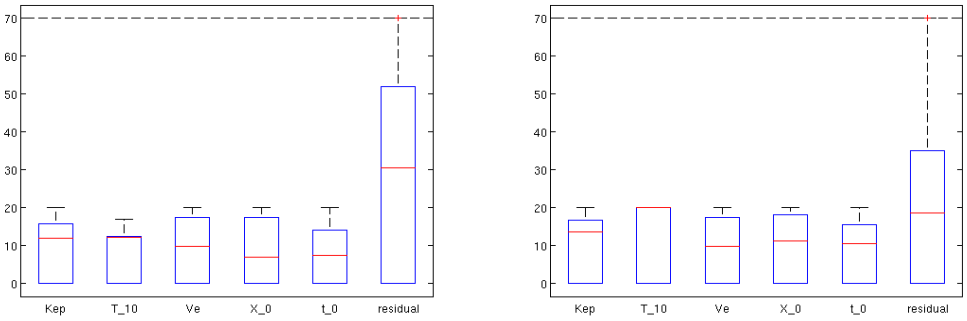

Figure 5.13: The percentage contributions of each PK parameter and the residual to the estimate of noise, calculated without motion correction (left) and with motion correction (right) across all voxels from twenty pre-therapy DCE-MRI scans.

tors remained the same or decreased, seems to indicate that in absolute terms, motion correction helped to reduce the error contribution of all other factors (especially the residual), but did not impact the contribution of T_{10}.

The percentage contribution of the pre-contrast MRI intensity (X_0) also increased marginally, while that of onset time (t_0) remained relatively unaffected by the process of motion correction, although in Section 5.5.3.2, the uncertainty in both their individual estimates has been shown to reduce. The analysis of noise contributions with and without motion correction clearly indicated two things:

1. The absolute value for the parameter controlling the Gamma distribution over the noise precision (s), increased as a result of motion correction, indicating that the VB estimate of the noise precision becomes higher when motion correction is incorporated into the analysis framework.

2. The percentage contribution of the 'model-fit' component (or the residual) also decreases considerably as a result of motion correction, and is compensated by an increased contribution from T_{10} and the pre-contrast MRI intensity (X_0).

The second observation outlined above can be seen clearly from the box-plots shown in Figure 5.13 for the percentage contributions of each parameter with and without motion correction, calculated over all twenty pre-therapy DCE-MRI scans. While the percentage contribution of the residual ($k^T k$) decreased from an average of 30.7% to 19.1%, the percentage contribution of T_{10} increased from 12.3% to 20.1%, and that of X_0 increased from 7.6% to 11%.

5.7 Correlations between Parameters

In the analysis of PK parameters derived from DCE-MRI data, one of the key questions has been the correlations (or lack thereof) between the patterns of different physiological parameters, shown through different parameter maps. This question is important because it impinges on the larger question pertaining to the clinical and prognostic value of different PK parameter. The general consensus is that K^{trans} or k_{ep} are the main PK parameters of interest, while v_e and other parameters are relatively less important. However, there are differing views about the importance of T_{10}, t_0, and on the extent of overlap between K^{trans}, k_{ep}, and other parameters. In this scenario, it is important to determine the extent to that different PK parameters provide overlapping or independent information.

There are various ways of examining the overlap in information provided by different PK parameter estimates. One way of examining this question would be to look at statistical measures of informational overlap such as conditional entropy or mutual information, based on the final PK parameter maps obtained. However, the VB framework provides us with a very useful tool for examining the correlations between different PK parameters by way of the posterior covariance matrix that emerges from the mathematical analysis. In many ways, this method of analysis is more fruitful, because it is

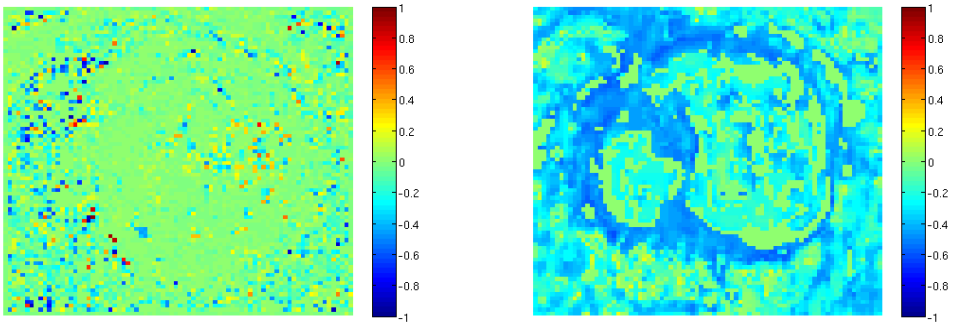

Figure 5.14: Correlation coefficients between k_{ep} and T_{10} (left) and between k_{ep} and v_e (right) for one example patient from the pre-therapy scan.

is the best estimate (with the constraints of the VB model) of the joint probability distribution of all of the parameters. In order to perform a quantitative comparison of the correlations between different pairs of PK parameters – say p_i and p_j – we calculated the correlation coefficient (CC) in the following manner:

$$CC\left(p_i,\, p_j\right) = \frac{Cov\left(p_i,\, p_j\right)}{\sigma_{p_i} \cdot \sigma_{p_j}} \tag{5.26}$$

5.7.1 Correlations in Clinical Data

In order to get an illustrative idea of how the different PK parameters are correlated with each other in clinical studies, we applied the VB framework to the twenty pre-therapy DCE-MRI scans acquired during the RIT study, and studied the mean and standard deviations of the correlation coefficients observed across different parameter pairs. The results obtained for one example patient are shown in Figs 5.14 – 5.18.

The correlation coefficients between k_{ep} and T_{10} and between k_{ep} and v_e are shown in Figure 5.14. It is clear from the Figure that there is little or no correlation between k_{ep} and T_{10}, while there is a clearly discernible negative correlation between k_{ep} and v_e. In the tests across all twenty patients, we found an average correlation coefficient of

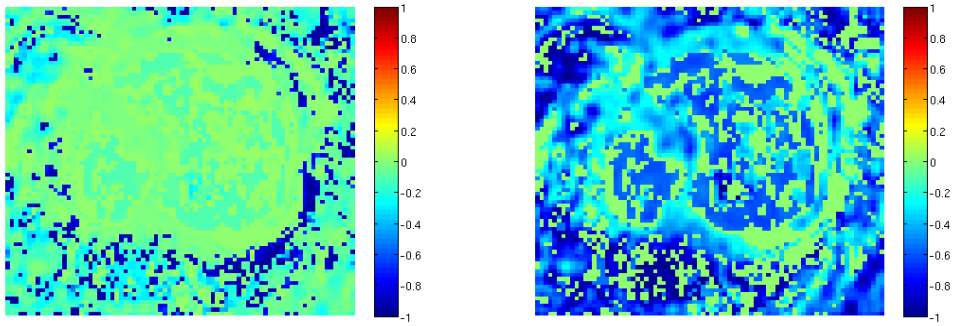

Figure 5.15: Correlation coefficients between k_{ep} and X_0 (left) and between k_{ep} and t_0 (right) for one example patient from the pre-therapy scan.

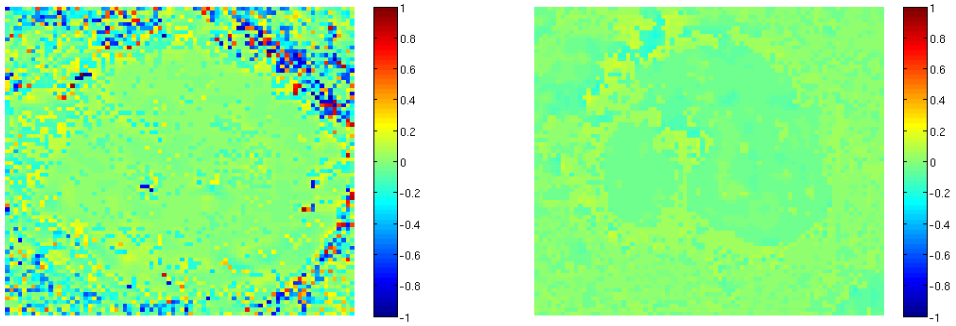

Figure 5.16: Correlation oefficients between v_e and T_{10} (left), and between v_e and X_0 (right) for one example patient from the pre-therapy scan.

-0.11 between k_{ep} and v_e.

Figure 5.15 shows that there is almost no correlation between k_{ep} and X_0, but a strongly negative correlation between k_{ep} and t_0. The average value for the correlation coefficient measured across all twenty scans, between k_{ep} and t_0, was found to be -0.23. Similarly Figure 5.16 shows that the correlation coefficients between v_e and T_{10} as well as v_e and X_0 are both negligible. However, while the correlation map between v_e and X_0 has a discernible structure, the correlation map between v_e and T_{10} seems more noisy.

In the same manner, we can examine the correlation coefficients between the onset

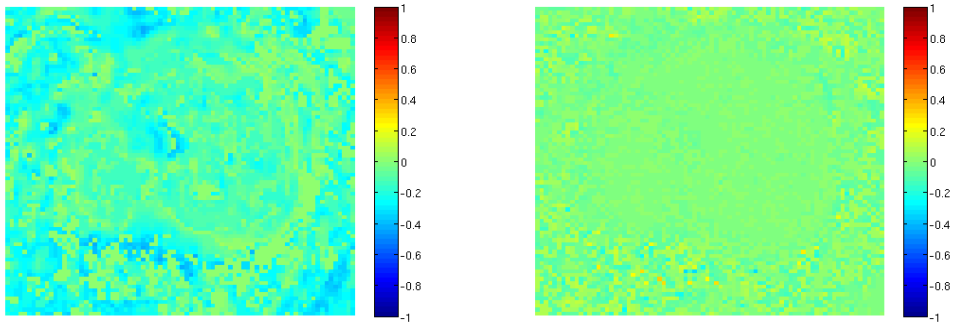

Figure 5.17: Correlation oefficients between v_e and t_0 (left), and between T_{10} and t_0 (right) for one example patient from the pre-therapy scan.

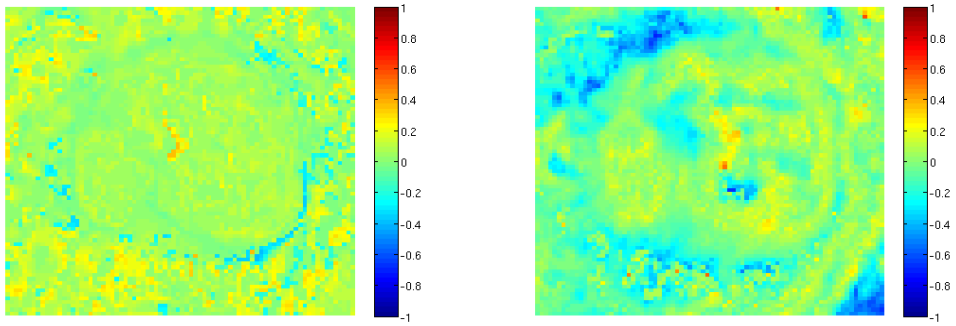

Figure 5.18: Correlation oefficients between X_0 and t_0 (left), and between X_0 and T_{10} (right) for one example patient from the pre-therapy scan.

time (t_0) and the other PK parameters as well. Figure 5.17 shows that the correlation coefficients between t_0 and v_e has low values, with an average of -0.2 for this patient, but only -0.07 on average across all twenty scans. The correlation between t_0 and T_{10} is seen to be negligible.

On examining the correlation coefficients between the pre-contrast MRI intensity (X_0) and t_0 on the one hand, and X_0 and T_{10} on the other, we find that both have structurally similar maps, but the former pair has higher values overall than the latter.

The analysis of the correlation coefficients between different PK parameters revealed a few interesting observations. The three pairs of correlations, which were noticeable,

were the following (in increasing order of absolute magnitude):

1. Correlations between k_{ep} and v_e: A negative correlation ($mean = -0.11$) is seen between k_{ep} and v_e, shown in Figure 5.14.

2. Correlations between T_{10} and t_0: Figure 5.18 shows that there is a reasonably strong ($mean = 0.16$) positive correlation between T_{10}, and the onset time (t_0).

3. Correlations between k_{ep} and t_0: Figure 5.15 shows that there is an even stronger ($mean = -0.23$) negative correlation between the transfer coefficient k_{ep} and the onset time t_0.

The three observations outlined above led us to the next obvious question that arises – i.e. what is the correlation between K^{trans} and t_0? In order to examine the correlations between a derived parameter such as $K^{trans}(= k_{ep} \cdot v_e)$ and t_0, we need to examine how the VB mathematical framework can be used to calculate variances and covariances between derived parameters.

5.7.2 Variances and Correlations for Derived Parameters

The variances and covariances for derived parameters can be estimated using the covariance matrix that emerges from the VB framework, as well as a Taylor series approximation that allows us to approximate the variation in a derived parameter using gradients of the derived parameter with respect to the other primary variables, and the covariance matrix between the variables themselves.

In order to estimate the uncertainties for derived parameters (P), such as $K^{trans} = k_{ep} \cdot v_e$, we use the Taylor series expansion of $P = g(\theta) = g(k_{ep}, v_e)$, which is: $g(\theta + \partial\theta) = g(\theta) + \nabla g.(\partial\theta)^T$, to approximate the variance of P as:

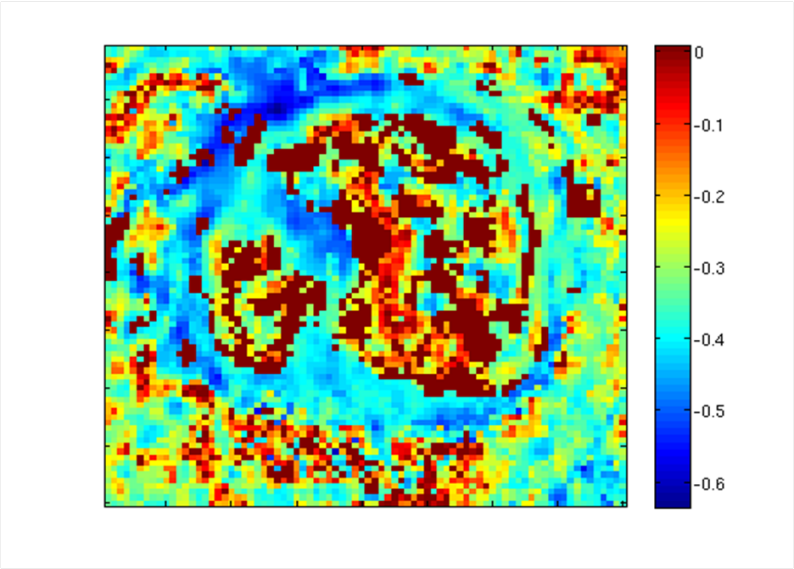

Figure 5.19: Correlation coefficients between K^{trans} and t_0 for one example patient from the pre-therapy scan.

$$\begin{aligned} var(P) &= E\left\{(P - P_{mean})^2\right\} &\approx& \quad E\left\{(\nabla g \cdot (\partial\theta)^T)^2\right\} \\ &= (\nabla g) \cdot E\left\{(\partial\theta)^T \cdot (\partial\theta)\right\} \cdot (\nabla g)^T &=& \quad (\nabla g) \cdot \Sigma \cdot (\nabla g)^T \end{aligned} \tag{5.27}$$

Here Σ denotes the covariance matrix of the PK parameters obtained as part of the VB estimation. In this manner, the variance of K^{trans} can be calculated as:

$$var(K^{trans}) = [v_e,\, k_{ep}] \cdot cov(k_{ep},\, v_e) \cdot [v_e,\, k_{ep}]^T \tag{5.28}$$

Similarly, the covariance between K^{trans} and t_0 can be estimated as:

$$\begin{aligned} cov(K^{trans}, t_0) &= E\left\{(K^{trans} - K^{trans}_{mean})(t_0 - t_{0,mean})\right\} &\approx& \quad E\left\{(\nabla g \cdot (\partial\theta)^T \partial t_0\right\} \\ &= (\nabla g) \cdot E\left\{(\partial\theta)^T \cdot \partial t_0\right\} \end{aligned} \tag{5.29}$$

Here, we can replace $(\partial\theta)^T \cdot \partial t_0$ by the vector of covariances $[cov(k_{ep}, t_0),\, cov(v_e, t_0)]$. Upon examining the covariance between K^{trans} and t_0, we found that there was a strong

negative correlation between the two, as shown in Figure 5.19. The average value of the correlation coefficient measured across all twenty scans, between K^{trans} and t_0 was found to be -0.32, which was the strongest (negative) correlation measured between any two parameters.

5.8 Uncertainties and Prediction of Response to CRT

In the previous sections, we have investigated the effect of motion correction on the uncertainties and covariances that are calculated by the VB algorithm. However, since the ultimate purpose of developing these tools for the analysis of DCE-MRI data is to improve the prediction of response to therapy, and hence facilitate patient stratification, it is important for us to determine whether the incorporation of uncertainty into the DCE-MRI analysis improves prediction of therapy.

In order to study the effect of uncertainty on the prediction of response to therapy, two key issues need to be dealt with after the variances have been calculated using the VB algorithm. The first is how to convert the variances into some measure of relative uncertainty or a weighting. The second issue is how this weighting should be integrated into the analysis of the parameter maps. Since we have so far looked at the cumulative distribution functions (CDFs) of the parameter maps acquired before and after therapy, we are now interested in examining how these CDFs are altered as a result of weighting by uncertainties.

The values for the variance that emerge from the VB algorithm are quite meaningful in themselves, since they indicate the range within that most parameter values lie. However, from Equations 5.17 and 5.16 it is clear that the final values for the variance that are obtained are a function of the initialization, as well as the Jacobians calculated at each iteration. The value of the Jacobians, in turn, can be influenced by a variety

of factors, such as the total amount of noise and motion in that particular scan, the spatial variations in the true parameter maps, etc. Thus, if the values of the parameters were to be weighted by the values of precisions themselves, then it would be highly likely that the parameter maps from a scan with more motion/noise would show an artificially decreased set of values as opposed to a scan with less motion/noise. This would artificially obscure the information that we are interested in, and make it more difficult to make accurate clinical predictions based on the parameter maps.

In order to avoid this problem, it was necessary to normalize the variances estimated for each voxel in some manner. We did so by dividing the value of variance by the sum of variances estimated across all voxels. Thus, in order to compute each parameter map *with* uncertainty, the parameter map was divided point-wise by its corresponding variance (normalized by the average variance calculated across all voxels). Thus, for each parameter p, and each location $(x,\ y,\ z)$, we calculated a new parameter value np, in the following manner:

$$np_{(x,y,z)} = \frac{p_{(x,y,z)}}{\sigma_{(x,y,z)}} \cdot \overline{\sigma} \tag{5.30}$$

Here, $\sigma_{(x,y,z)}$ denotes to the standard deviation estimated for the PK parameter p at the location $(x,\ y,\ z)$, and $\overline{\sigma}$ denotes the average value for the standard deviation, calculated across all voxels in the tumour being analysed. Once the new parameter map had been computed, the KS-distance between the pre- and post-therapy K^{trans} distribution for each patient was calculated in exactly the same manner, as explained in Chapter 4.

Table 5.4 shows the same statistical measures that we had computed earlier in Section 4.4, with and without incorporating uncertainty into the calculations. From the Table, it is clear that incorporation of uncertainties does not improve the results significantly, and in fact, in some cases, it makes the results slightly worse. This is most probably due to the fact that we are still computing measures (KS-distance) that depend on the histograms of the PK parameter distributions, rather than the spatial

Measure	Weinmann w/o	Weinmann with	Orton w/o	Orton with
p-value w/o	4×10^{-3}	2×10^{-4}	3.5×10^{-4}	5×10^{-7}
p-value with	3.5×10^{-3}	1.4×10^{-4}	4×10^{-4}	10^{-6}
LOOV w/o	0.26	0.14	0.17	0
LOOV with	0.23	0.12	0.19	0

Table 5.4: Statistical measures for the difference between the partial responder ($n = 11$) and non-responder ($n = 9$) groups for the KS-distance measured between the pre- and post-therapy scans using each of the four methods, with and without incorporating uncertainty. LOOV represents the proportion of samples mis-classified during leave-one-out validation.

distribution itself. Thus, the advantages of incorporating uncertainty seem to be lost in the process of data compression that takes place in the process of calculating statistical measures that do not take spatial distributions into account.

In the next chapter, we will investigate whether the inclusion of uncertainty measures causes any improvement in the prediction of therapeutic outcome, when the spatial distribution of PK parameters is being taken into account, and when measures of heterogeneity are being calculated.

Chapter 6

Heterogeneity within Tumours

One of the other factors affecting disease outcome in cancer, which has been alluded to in earlier chapters but not accorded the importance that it deserves, is tumour heterogeneity. Our understanding of what constitutes 'heterogeneity', and how it affects treatment outcome is still at a nascent stage, but there is a growing consensus that it has enormous significance for predicting the outcome of various treatment protocols [Yang and Knopp, 2011].

Apart from the size and location of the tumour, it is now well-established in the clinical community that the extent of vascularization in and around the tumour, and its pattern of growth are also of major prognostic value. In Chapter 4, we have shown that the measurement of perfusion through PK modeling can be used to effectively distinguish between responders and non-responders to chemoradiotherapy in a cohort of colorectal cancer patients. In a previously presented study [Bhushan et al., 2013], the importance of spatial heterogeneity and uncertainty in predicting response to therapy has been demonstrated. The present chapter describes in detail the motivation behind the methods used in [Bhushan et al., 2013], and also provides a more comprehensive analysis of the impact of heterogeneity and uncertainty in different PK parameter maps

on the prediction of response to therapy.

6.1 Previous Work

6.1.1 Statistical Measures of Heterogeneity

The simplest measure of heterogeneity that has been reported in literature is the standard deviation (σ) of quantitative or semi-quantitative parameters calculated from DCE-MRI data. It has been shown that for primary breast tumours, the classification of malignant and benign lesions can be improved by taking the standard-deviation of K^{trans} into account, in addition to its mean [Issa et al., 1999]. The importance of statistical measures of heterogeneity in the classification of benign and malignant lesions has also been investigated in [Ertas et al., 2007]. While looking at the normalized maximum intensity-time ratio, it was found that the entropy and standard deviation were both found to have a high diagnostic accuracy in classifying suspicious regions correctly.

Although the task of lesion classification is different from predicting response to therapy, even for the task of predicting response to chemotherapy in breast cancer patients, it has been found that the range of K^{trans} values displays greater sensitivity than the mean [Padhani et al., 2006]. In this study, it was also demonstrated that the range of K^{trans} values decreased in those patients who showed a response to chemotherapy (as assessed from pathology), while no other kinetic parameter was successful in predicting response. However, this study did not involve any motion correction, and the comparison between histological data and DCE-MRI statistics was performed by calculating the DCE-MRI parameters from a single slice manually chosen by clinicians, since it was not possible to do volumetric analysis with the software available at that time. The histograms of relative signal intensity (RSI) values calculated across all pixels were

analyzed in [Mayr et al., 2000], and it was found that the $10th$ percentile RSI value combined with the number of pixels in the tumour region had a 100% success rate in being able to predict subsequent tumour control vs recurrence in radiotherapy treatment of 16 cervical cancer patients. However, since the RSI is a measure that depends on the specifics of the scanner and imaging protocol that is used, it is difficult to assess the reproducibility of these results. In another study focusing on high-grade gliomas, the median of the K^{trans} distribution did not correlate with either grade or treatment outcome for tumours, whereas the 95th percentile value was found to be closely correlated to overall survival [Mills et al., 2006]. There are several other studies that use measures derived from the histograms – different percentile values, skewness, kurtosis, etc – that can be used to quantify the degree of heterogeneity in any parameter map.

However, the utility of all such measures is limited, because they neglect the spatial distribution of parameter values. When values calculated independently across voxels are compressed into a summary value or even a histogram (as was done to calculate the KS-distance between pre- and post-therapy K^{trans} distributions), there is always a distinct possibility that the arrangement of pixel values may change completely, with the histogram remaining the same. Thus, it is important to incorporate information about the positions of different PK parameter values with respect to each other in a manner that can be interpreted easily by clinicians.

6.1.2 Taking Spatial Arrangements Into Account

Current knowledge about the biological growth and spread of tumours also indicates why the spatial distribution of PK-parameter values may be very important. PK parameters such as K^{trans} are thought to represent perfusion and permeability. Thus, it is quite likely that areas with high K^{trans} are more likely to receive a higher dosage of cytotoxic agents than other parts. Also, since oxygenation is also linked to blood flow,

areas of high K^{trans} may also be correlated to areas of hypoxia. Hypoxia can have severe implications for the radiosensitivity of different areas, and therefore on the effectiveness of radiotherapy. Similarly, the onset time has been shown to be correlated with areas of necrosis, which could be indicative of how effective chemotherapy or radiotherapy has been for the patient.

Thus, calculating measures of heterogeneity based on the spatial distribution of K^{trans} values, or values of onset times, can be very beneficial. As we can see from the previous section, the histogram analysis of tumour regions has certainly achieved some success in the estimation of heterogeneity in tumours. Although the usual method for calculating histograms does not take the spatial distribution of values into account, different methods of estimating histograms have also been proposed – most notably the Non-parameteric estimation method or the "NP-Windows" method employed in [Dowson et al., 2008] and [Joshi and Brady, 2009]. One possible reason for the success of histogram based values (standard deviation, kurtosis, different percentiles) in predicting response may be that in many tumours, the high and low K^{trans} (or other parameter) portions of the histogram correspond to distinct regions of the image space. Thus, looking at the 95th percentile or above of the K^{trans} histogram, is somewhat similar to "hot-spot" analysis of tumours, where the clinician identifies the brightest area of the tumour ROI from the K^{trans} map, and then calculates values based only on that area [Mussurakis et al.,]. Thus, if the hot-spot is placed accurately, then it may be approximately equivalent to examining the maximum intensity part of the histogram. However, examining the histogram itself is likely to be more inaccurate since the hot-spot may contain low-intensity voxels within it, and since its shape and size may vary with respect to surrounding areas (thus affecting its representation in the histogram).

Hot-spot analysis has been used widely in breast cancer to calculate quantities such as the peak enhancement ratio (PER), and the maximum intensity-time ratio (MITR).

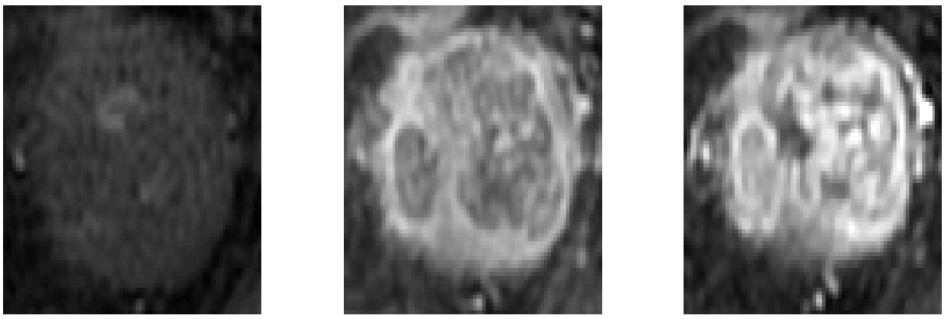

Figure 6.1: Axial slices of the tumour from a dynamic scan, taken at $t = 10$ sec, $t = 34$ sec, and $t = 2$ mins after injection of contrast agent.

Although Mussuraki and colleagues [Mussurakis et al.,] found that there was no statistically significant advantage of using PER derived from hot-spots in order to distinguish between benign and malignant tumours, another study [Liney et al., 1999] found that while looking at MITR instead of PER, the hot-spot analysis performed much better than the entire tumour. In other solid tumours, including colorectal tumours, what is often apparent in a DCE-MRI scan is the early enhancement of the rim, followed by a slow diffusion on CA into the inner parts of the tumour, as seen in Figure 6.1. In some tumours, the CA never reaches the innermost part of the tumour, indicating the presence of a necrotic core. In order to capture this effect, some groups have evolved methods of segmenting the tumour into a poorly enhancing core, and a strongly enhancing periphery [Jiang et al., 2004], [Checkley et al., 2003].

This concept has been extended and used to calculate a measure of "radial heterogeneity" in [Benjaminsen et al., 2004]. In this paper, the authors have divided the image of the tumour into five concentric sectors, bounded by lines drawn at distances of $n.R/5$ $(n = 1, .., 5)$ from the tumour center. Here the authors were essentially looking at the distribution of K^{trans} in the five concentric layers going from the centre of the tumour to its periphery, and it was shown that there was a high degree of correlation between the normalized average K^{trans}, and the normalized blood supply in each of

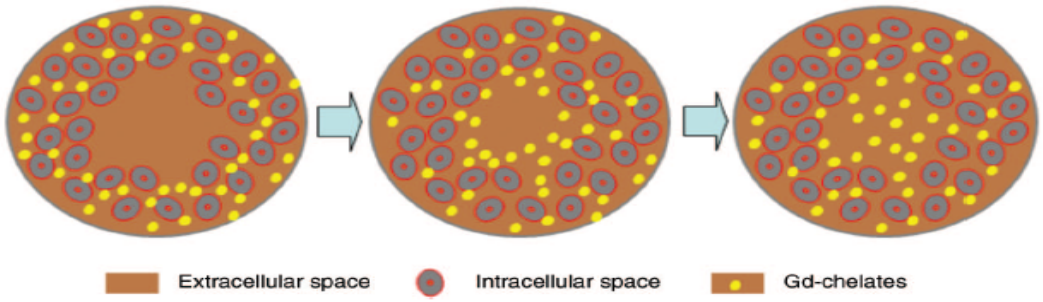

Figure 6.2: The process of contrast-agent diffusion from the outer rim to the centre of the tumour, as shown in [Benjaminsen et al., 2004].

the five sectors. Further, the four different categories of tissue as determined from histology, had increasing levels of average K^{trans} corresponding to increasing levels of vascularization.

The idea of a well-perfused rim and a poorly enhancing core of a tumour comes from observations of tumour lesions and solid tumours, which have been confirmed by many groups. The observations from DCE-MRI scans actually correspond to the behavior of tumour perfusion – where the contrast agent reaches the outer rim of the tumour first, and then slowly diffuses inwards, as shown in Figure 6.2. In order to mathematically capture and quantify this diffusion process, another parameter known as the contrast diffusion coefficient (CDC) has been derived in [Jia et al., 2008]. In order to calculate this quantity, the authors first estimate the signal intensity gradient at each position in the 2D image space by adding the absolute difference between each pixel and its four neighbours. Thus, the signal intensity gradient at each position (i, j) is given by:

$$\nabla S_{i,j} = \frac{1}{4}\left(\mid I_{i,j} - I_{i-1,j} \mid + \mid I_{i,j} - I_{i+1,j} \mid + \mid I_{i,j} - I_{i,j-1} \mid + \mid I_{i,j} - I_{i,j+1} \mid\right) \qquad (6.1)$$

The total amount of diffusion observed in the tumour at each time-point t of the

scan is then derived as follows:

$$I_{\nabla S}(t) = \frac{1}{\overline{S(t)}} \sum_{i,j} \nabla S_{i,j}(t) \tag{6.2}$$

Here, $\overline{S(t)}$ refers to the average intensity across the tumour ROI at time t. The signal intensity gradients across all locations in the tumour ROI (total N pixels) are then summed to arrive at the total diffusion at time t. As one would expect, the magnitude of total diffusion decreases gradually from the time of arrival of contrast agent at the tumour periphery, until it is uniformly diffused across the tumour ROI. Thus, the authors determine a decay factor b by fitting an exponential decay function given by:

$$I_{\nabla S}(t) = k \cdot \exp(-b \cdot t) + m \tag{6.3}$$

In order to derive the CDC using Equation 6.3, the authors solve the standard diffusion equation for one idealised case by imposing three boundary conditions: (1) spherical shape of the tumour, (2) impermeability of the CA at the tumour boundary, and (3) uniform enhancement of the tumour rim. Subject to these simplified assumptions, the authors derive the following linear relationship between the decay factor b and the CDC:

$$CDC = b\frac{A_{ROI}}{\pi(4.4934)^2} \tag{6.4}$$

Here A_{ROI} represents the size of the tumour. The physical interpretation of Equation 6.4 is that the CDC represents the combined effect of how big the tumour is, and how quickly the contrast-agent diffuses into it. However, the truth is that in many tumours, we do not simply observe a uniform diffusion from the periphery into the center. We observe that some vascular structures exhibit enhancement first, and then the CA diffuses from the vascular structures into the interstitial spaces (Figure 6.4).

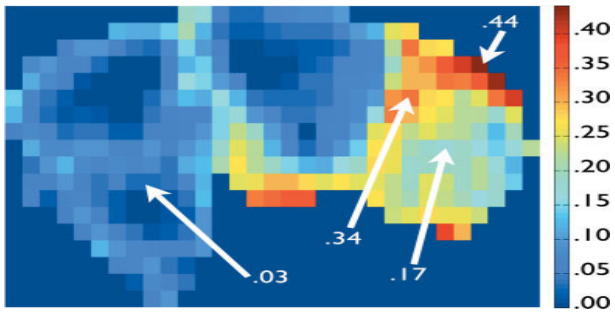

Figure 6.3: The K^{trans} map showing the heterogeneity of a tumour. Figure taken from [Rose et al., 2009b].

The diffusion model assumed in [Jia et al., 2008] holds true for tumours that are very poorly vascularized. However, it does not hold true for the tumours that have clearly defined vascular structures within them. For such tumours, it is necessary to develop an alternative method for characterizing their heterogeneity.

There are a series of other approaches for characterizing heterogeneity based on the fractal nature of the PK-parameter maps [Rose et al., 2009b], [Dzik-Jurasz et al., 2004]. In order to calculate the fractal measures on PK parameter maps for 3D tumours, the parameter maps first need to be projected onto a 4th dimension. Thus, the K^{trans} value at each voxel represents the 'height' at that position, and the fractal dimension or surface area is calculated in a similar way to how it would be done for a 2D image. In order to quantify the fractal dimension, two main approaches are used. One is related to calculating the Renyi fractal dimension d_q, which is estimated as:

$$d_q = \lim_{s \to 0} \frac{\log \sum_i p_i^q}{(1-q) \log \frac{1}{s}} \quad and \; d_1 = \lim_{q \to 1} d_q \tag{6.5}$$

Here s is the scale and p_i is the normalized value of the i-th voxel [Rose et al., 2009b]. The Renyi fractal dimension for $q = 0$ is known as the box dimension, which has been used in various applications. However, the box dimension is usually applied in the case of binary images, where the actual values at different positions is not important. In

the case of PK-parameter maps, since the value of the parameters at each position is of great importance, the higher order Renyi dimensions and other measures, such as the surface area, or the ratio of the surface area to the volume have been preferred. In order to calculate the surface area of the parameter map, the basic quantity that needs to be computed is a sum of the differences between the values at each pixel and its neighbours. Thus, the surface area can easily be approximated by:

$$S_T = \frac{1}{2} \sum_{x \in T} \left[\sum_{y \in \Omega_x} (|f(x) - f(y)|) \right] + N_T \tag{6.6}$$

Here, S_T refers to the surface area for the tumour T, N_T refers to the total number of voxels within the tumour, and $f(x)$ denotes the value of the chosen PK parameter at position x. The neighbourhood of x is denoted by Ω_x (six neighbours in 3D, and four in 2D) and a factor of $\frac{1}{2}$ is needed in order to prevent double-counting of the same surface area. In order to normalize this quantity, the measure that is usually examined is the ratio of the surface-area with the volume, i.e. S_T/N_T.

The ratio of the surface area to the volume of a PK parameter map, or the Renyi fractal dimension, essentially quantifies how fractured, or discontinuous the parameter map is. If one thinks of a K^{trans} map, it is not hard to imagine that the highest surface-area to volume ratio will be achieved when the map consists of alternating high and low K^{trans} values throughout the map. Thus, such a measure does not highlight the importance of features such as enhancing rims or noticeable vascular structures within the tumour, which are known to be physiologically meaningful.

In the clinical and pre-clinical data available to us, we observed that in the entire DCE-MRI scan (Figure 6.1) and the K^{trans} maps obtained for these datasets (Figure 6.4), two things were clear:

1. The enhancing rim not only appears brighter than the surrounding areas, but also

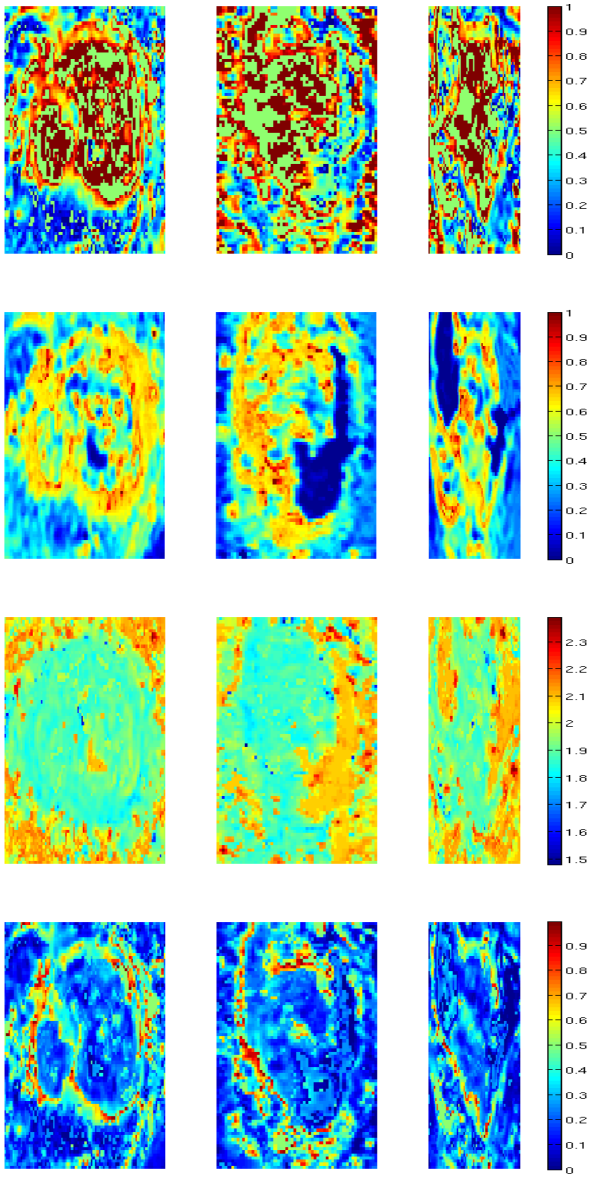

Figure 6.4: Axial (left column), sagittal (middle column) and coronal (right column) slices of the k_{ep} map in min^{-1} (top row), the K^{trans} map in min^{-1}(second row), the onset time in min (third row), and a ratio of the K^{trans} map with the map of onset times in min^{-2} (bottom row) for a rectal adenocarcinoma.

enhances earlier than the surrounding tissue, as seen in Figure 6.1.

2. The enhancing rim has higher K^{trans} values than surrounding areas, whereas the necrotic core has a lower value, as seen in the dark blue areas in the second row of Figure 6.4.

3. The maps for K^{trans} and k_{ep} appear to be inversely correlated with the map for the onset time, t_0, as can be seen from the second and third row of Figure 6.4.

6.2 The Acceleration Constant

It is important to remember here that the negative correlation between K^{trans} and t_0 seen here was also found when correlations emerging from the VB framework were examined in Section 5.7. A close examination of the actual spatial maps K^{trans} and t_0 further confirms the findings from the previous chapter. This negative correlation is neither meaningless, nor a coincidence. Physiologically, the ratio $k_a = K^{trans}/t_0$ (having units min^{-2}) can be likened to a measure of the acceleration experienced by the bolus when it comes into contact with the tissue at that location, and is henceforth referred to as the 'acceleration constant'. It signifies the ability of the vasculature to respond to the force exerted by the bolus, per unit mass.

The early enhancement of the rim and some of the intra-tumoural vessels can be seen even more clearly in the map of the onset time. However, in order to capture the effect of both the early enhancement and the higher K^{trans} values, we have chosen to examine the ratio of K^{trans} to the onset time. This gives the clearest visualization of the enhancing rim and of the vasculature within the tumour. It is important to note that even though the acceleration constant may appear to be equal to the slope of the contrast uptake curve during its initial period, this is actually not the case. The slope is a quantity that is quite sensitive to where it is calculated, and normally it is

calculated as an average between the time that that CA reaches the concerned voxel (t_0), and the time-to-peak. In order to illustrate the difference between the slope of the contrast-uptake curve and the suggested quantity k_a, let us examine the equation for the total concentration of contrast-agent again:

$$C_T(t) = DK^{trans}\left\{3.99\left(\frac{e^{-\frac{K^{trans}}{v_e}t} - e^{-0.144t}}{0.144 - \frac{K^{trans}}{v_e}}\right) + 4.78\left(\frac{e^{-\frac{K^{trans}}{v_e}t} - e^{-0.011t}}{0.0111 - \frac{K^{trans}}{v_e}}\right)\right\} \quad (6.7)$$

This form of the equation assumes the Weinmann population-averaged AIF, but all other cases would be mathematically similar. Now in order to calculate the slope of this contrast uptake curve with respect to time, we would obtain:

$$\begin{aligned}\frac{dC_T(t)}{dt} &= DK^{trans}3.99\left(\frac{\frac{-K^{trans}}{v_e}\cdot e^{-\frac{K^{trans}}{v_e}t} + 0.144\cdot e^{-0.144t}}{0.144 - \frac{K^{trans}}{v_e}}\right) \\ &\quad +4.78\left(\frac{\frac{-K^{trans}}{v_e}\cdot e^{-\frac{K^{trans}}{v_e}t} + 0.0111\cdot e^{-0.011t}}{0.0111 - \frac{K^{trans}}{v_e}}\right) \quad (6.8)\end{aligned}$$

In order to calculate the slope at any particular time t_p, we can simply replace t with t_p in the equation above. In this manner, the slope at the time of arrival of bolus (onset time) can also be calculated, or averaged over the period between t_0 and the time-to-peak, t_{peak}. However, these quantities would convey a different meaning from the ratio $k_a = K^{trans}/t_0$. One of the reasons that we have chosen to look at k_a is also because we wanted to incorporate the information contained in the onset time, which is not reflected in the other Tofts model parameters, and that has been shown to be correlated with areas of necrosis within the tumour [McPhee, 2012].

The onset time has been looked at in the context of semi-quantitative values such as the maximum-intensity time ratio (MITR) in [Liney et al., 1999]. However, like all

other semi-quantitative values, the MITR also suffered from inconsistency across different scanners and imaging protocols. With k_a, we hope to overcome these shortcomings, and look at the combined prognostic value of early onset time as well as high K^{trans}.

6.3 Heterogeneity measures

As discussed in the previous section, there are many existing ways of characterizing the the heterogeneity in PK parameter maps. In this study, we have compared various simple measures of heterogeneity. An examination of the raw DCE-MRI data available to us made it clear that most of the tumours had a clearly defined enhancing rim, as well as a intra-tumoural vascular structures. It was also evident that the enhancing rim and intra-tumoural vascular structures not only appeared brighter than surrounding areas, but also enhanced sooner than the neighbouring areas. Thus, apart from looking at only the K^{trans} and k_{ep} maps, we also examined the map of the acceleration constant $k_a = K^{trans}/t_0$, after scaling the values of t_0 so that the minimum was set to 1 (min). This scaling was applied to ensure that discrepancies between different patients, in the time of injection and the time taken for the CA to reach the tumour area did not affect the value of k_a across patients.

In order to calculate the heterogeneity, we used some simple measures to quantify different effects observed in the data. The measures that we have developed can be applied to any parameter, and we will examine their application to various PK parameter maps. First, we wanted to develop a measure that associates a higher value with continuous patches or lines (as observed in the enhancing rim and 'hot-spots') showing high parameter values. In order to do this, we calculated the quantity:

$$C_T = \sum_{x \in T} \left[f(x) \left(\sum_{y \in \Omega_x} \exp\left(\frac{-(f(x) - f(y))^2}{2\sigma} \right) \right) \right] \tag{6.9}$$

Here, C_T refers to the 'continuity' of parameter values (represented by the function f) in the tumour. The position of central voxels are denoted by x, and the neighbourhood of each central voxel is denoted by Ω_x. This measure essentially gives a higher weight to continuous patches of high parameter values in the tumour region, and accords lower values to pixels that may have a high parameter value themselves, but are surrounded by lower valued pixels. In this sense, the continuity measure serves the opposite purpose to the fractal measures described in equations (6.5) and (6.6). However, neither of these two approaches alone is particularly good at highlighting the enhancing rim and vascular structures noted earlier.

It is clear from Equation 6.6, which the ratio of the surface area to the volume is actually nothing but a normalized sum of the absolute values of the gradients in the images. Calculating gradients is one straightforward way of capturing the edges within an image, and hence all linear and line-based structures such as rims and vessels. However, both the gradient and the Renyi fractal measures are quite sensitive to noise, and accord equal importance to all changes in intensity, thus there is a need for looking at more robust measures that can capture the structures that we are interested in.

One set of measures that we can examine, in order to achieve our objective, are measures derived from the structure tensor:

$$ST = [I_x,\ I_y,\ I_z]^T \cdot [I_x,\ I_y,\ I_z] \tag{6.10}$$

$$= \begin{bmatrix} I_x^2 & I_xI_y & I_xI_z \\ I_xI_y & I_y^2 & I_yI_z \\ I_xI_z & I_yI_z & I_z^2 \end{bmatrix} \tag{6.11}$$

Here ST represents the structure tensor calculated at the location (x, y, z), and I_x, I_y, I_z denote the image gradients along the x, y and z axes respectively. The structure tensor and the Hessian have been examined in many papers, both in medical imaging

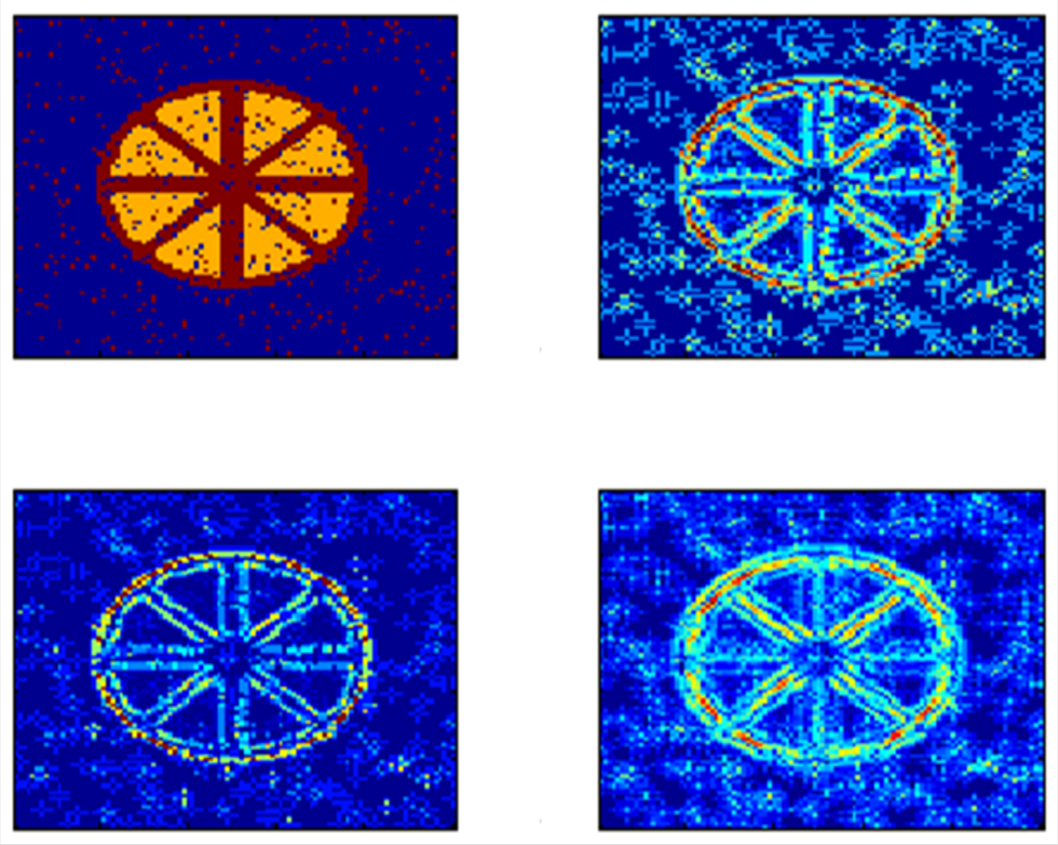

Figure 6.5: An artificial parameter map (top left), the absolute value of the gradients (top right), the largest eigenvalues of the structure-tensor eigen-decomposition (bottom left) and the largest eigenvalues from the Hessian eigen-decomposition (bottom right). Each image is scaled between the minimum and the maximum values for that image.

and in computer vision, as indicators of different features in the image. It is worth noting that for noisy data, the image first requires some Gaussian smoothing (by a convolution with a Gaussian kernel) as a pre-processing step. The most meaningful way of extracting information from the structure tensor, or the Hessian, is to look at the eigenvalue decomposition of the structure tensor, and then examine the local structure of the image at each point by examining the eigenvalues and eigenvectors at each position. We can get an idea of how the absolute values of the gradient compare with the largest eigenvalues of the Hessian and the structure-tensor decomposition by looking at a simple synthetic example.

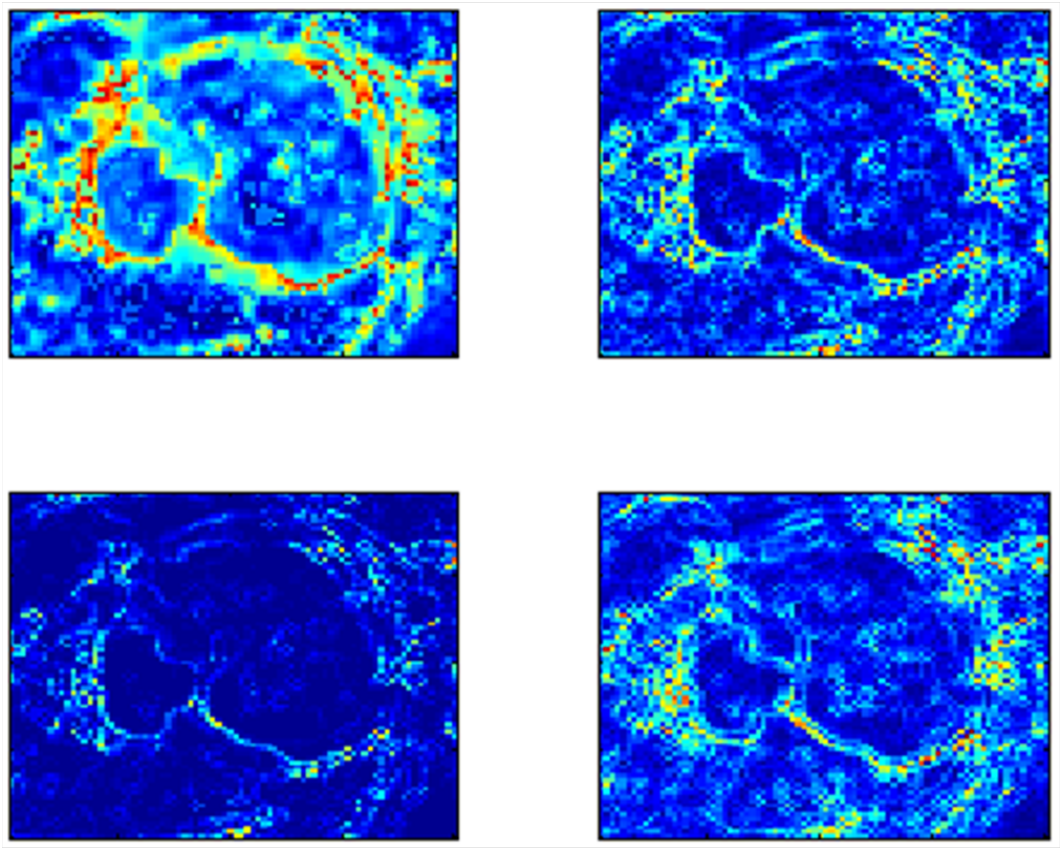

Figure 6.6: The K^{trans} map for a patient (top left), the absolute value of the gradients (top right), the largest eigenvalues of the structure-tensor eigen-decomposition (bottom left) and the largest eigenvalues from the Hessian eigen-decomposition (bottom right). Each image is scaled between the minimum and the maximum values for that image.

It can be seen from Figure (6.5), that absolute values of the gradients, as well as the largest eigenvalues of the Hessian are both more sensitive to noise than the largest eigenvalue of the structure tensor. However, both the sensitivity to noise and the appearance of the eigenvalues are subject to the scale of analysis and the size of the kernel used. Thus, it is necessary to look at a real example to get an idea of the features that each of these measures is able to highlight.

The eigenvalues of the Hessian have also been shown to have the ability to identify various important features such as vessels and airways in medical images [Wiemker et al., 2012]. However, due to the low resolution of the DCE-MRI data ($\sim 1mm^3$), the vascular struc-

tures and rim of the tumour are only seen as faint lines rather than as clearly defined vessels.

After some preliminary tests, we decided that calculating radial heterogeneity or rim-enhancement alone was not enought to characterize the irregular shapes of enhancing vascular structures within the tumours (as seen in Figure 6.4). We also found that the eigenvalues of the Hessian were too sensitive to noise, and therefore unreliable. On the other hand, the continuity measure C_T was too insensitive to changes in the parameter values, and required a careful calibration of the parameter σ, hence this was also abandoned and not used in any of the subsequent analyses. Finally, for each parameter map P, we chose to examine the following set of measures:

1. A measure of the fractal nature of the parameter map was estimated by calculating a ratio of surface-area to the volume of the parameter hypersurface. A simple method for approximating the surface-area of a 3D parameter map is to sum up the difference in parameter values between all neighbouring voxels and then divide by 2 to avoid double-counting of the same area. If we assume that the voxels are isotropic, and have length 1 unit in all dimensions, then:

$$S_T = \frac{1}{2} \sum_{x \in T} \left[\sum_{y \in \Omega_x} |P(x) - P(y)| \right] / N_T \tag{6.12}$$

   Here T represents the tumour region, N_T represents the number of voxels in the tumour region, and Ω_x represents the six neighbours of the voxel x in 3D. The volume of the 3D parameter map under these assumptions can be taken to be $V_T = \sum_{x \in T} P(x)$.

2. Since S_T measures only the difference between neighbouring voxels, without taking the value of the central voxel into account, we define another measure 'sharpness', which is obtained by a point-wise multiplication between the gradient map

by the parameter map P itself:

$$Sharpness_T = \frac{1}{2}\sum_{x\in T} P(x)\left[\sum_{y\in\Omega_x}(|P(x)-P(y)|)\right]/N_T \tag{6.13}$$

This simple measure is useful for characterizing the rim and other intra-tumoural vascular structures, which are visible as sharp, high-value edges in the PK parameter maps.

3. Another method for characterizing rims and other linear structures in the parameter map is to examine the structure tensor of the parameter map:

$$ST = \left[\frac{\partial P}{\partial x},\ \frac{\partial P}{\partial y},\ \frac{\partial P}{\partial z}\right]^T \cdot \left[\frac{\partial P}{\partial x},\ \frac{\partial P}{\partial y},\ \frac{\partial P}{\partial z}\right] \tag{6.14}$$

The eigenvalues of the structure tensor ($[e_1, e_2, e_3]$, where $e_1 > e_2 > e_3$) can then be used to characterize the linearity or 'planarity' of the parameter map at every position. Since we expect strongly linear parts of the parameter map to be characterized by one dominant eigenvalue, we measure:

$$Planarity_T = \sum_{x\in T} P(x)\frac{(e_1(x)-(e_2(x)+e_3(x)))}{e_1(x)+e_2(x)+e_3(x)}/N_T \tag{6.15}$$

Figure 6.7 shows the tumour delineation, along with the K^{trans} map and the k_a map for a patient who did not respond to CRT, while Fig. 6.8 displays the same information for a patient that responded positively to CRT. These Figures illustrate the differences between the typical K^{trans} and k_a distributions between responders and non-responders. While there are many irregularly shaped high-K^{trans} regions in the tumour region for the non-responder, the responder shows more linear high-K^{trans} regions closer to the tumour boundaries. This observation is even more pronounced in the case of the k_a

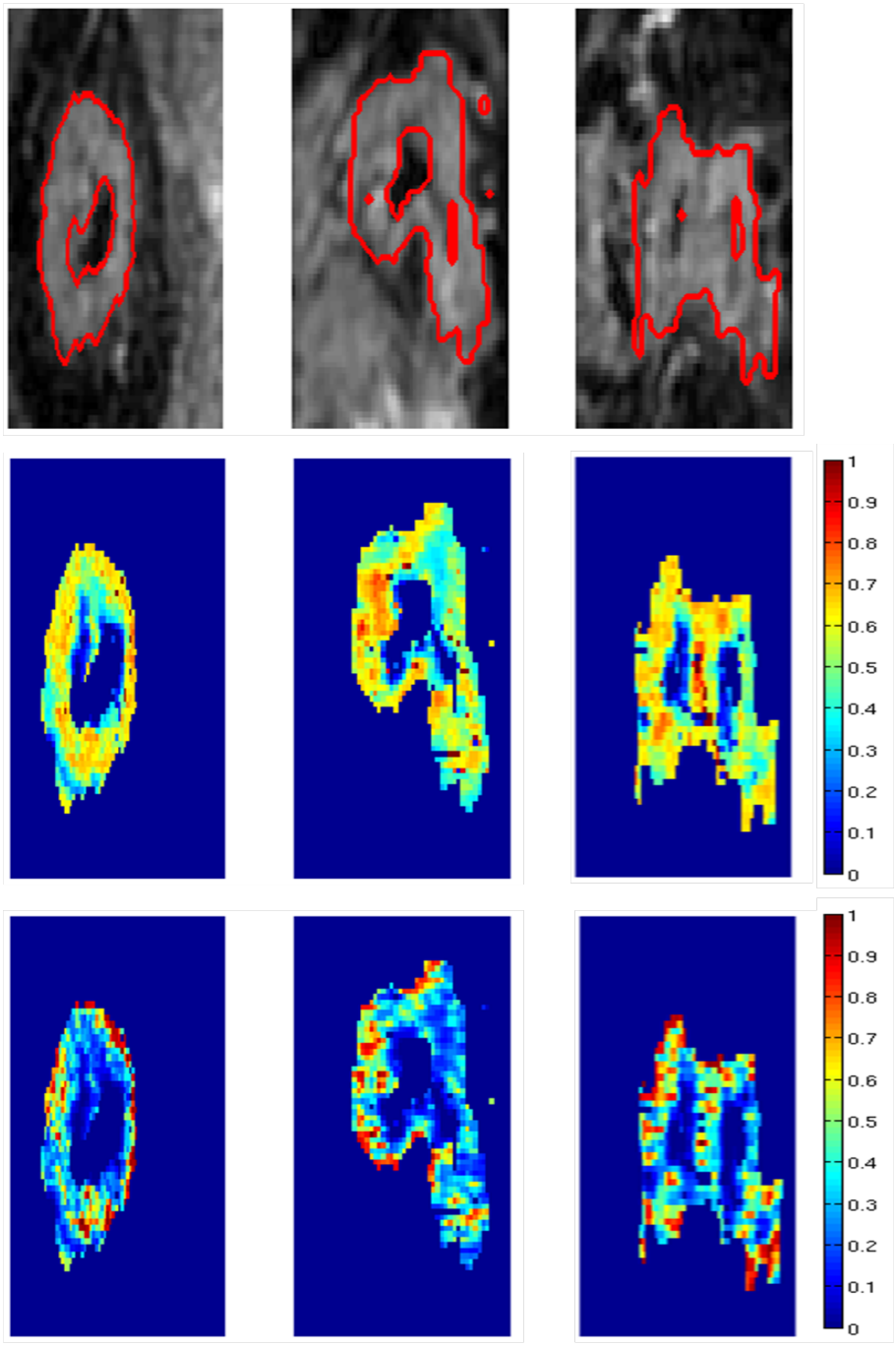

Figure 6.7: The tumour outline on the pre-contrast DCE-MRI image (top), the K^{trans} map in min^{-1} (middle) and a ratio of the K^{trans} map with the map of onset times in min (bottom) for a rectal adenocarcinoma of a non-responding patient.

maps for both patients.

6.4 Integrating Uncertainty and Heterogeneity

In the previous section we have examined various methods for calculating heterogeneity of PK parameter maps. However, as we have seen in Chapter 5, the parameter maps themselves are subject to uncertainties arising from noise, motion and other factors involved in the MRI acquisition process. In Section 5.8, we saw that integrating uncertainty into the prediction of response to therapy did not make any appreciable difference. The primary reason for this was the fact that the measure we computed (KS-distance) was relatively unaffected by changes in the uncertainty of parameter values, since it is calculated from the histogram, which in turn was unaffected by changes in the spatial arrangement of values.

We would have expected to see a significant impact of weighting by uncertainty on the KS-distance calculations, if the uncertainties in parameters were significantly correlated with the values of the parameters themselves, rather than their locations. However, in the datasets that we have analyzed, what is more apparent is that the uncertainty is higher near the boundaries of the tumour, and in areas that undergo more motion during the scan. There seems to be little correlation between higher values of the PK parameters and higher amounts of uncertainty. In other words, the uncertainty estimates have a greater correlation with the spatial distribution of parameters, than with the distribution of the values.

Thus, if we were to calculate measures of heterogeneity directly from the PK parameter maps without taking their precisions into account, we would allow the uncertainties in these maps to propagate, and perhaps be magnified in the calculation of derived measures. As discussed inthe previous chapter, we are not interested in weighting by the

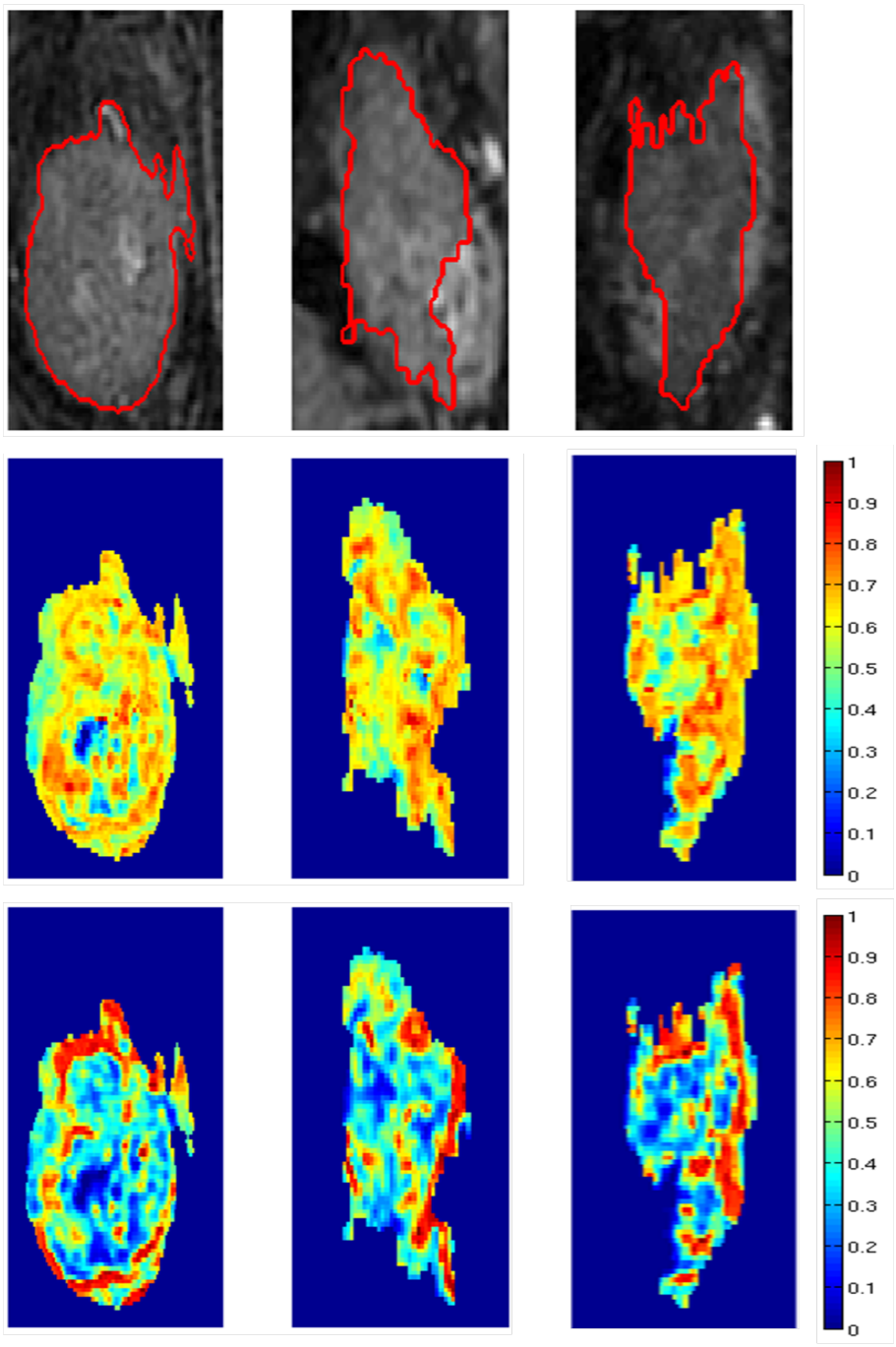

Figure 6.8: The tumour outline on the pre-contrast DCE-MRI image (top), the K^{trans} map (middle) and a ratio of the K^{trans} map with the map of onset times (bottom) for a rectal adenocarcinoma of a responding patient.

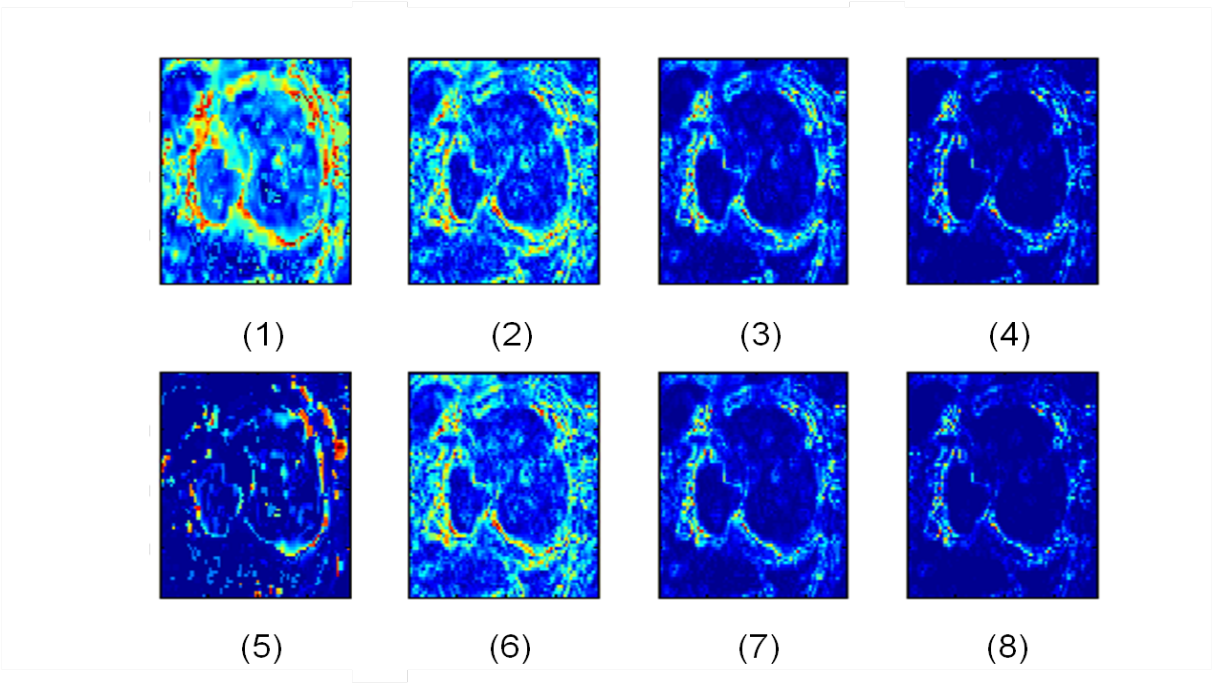

Figure 6.9: The k_a map for a tumour (1), and the maps showing S_T/V_T (2), sharpness (3) and planarity (4). The uncertainty in k_a (5) is then used to scale each of the other three maps: S_T/V_T (6), sharpness (7) and planarity (8). Each image is scaled between the minimum and the maximum values for that parameter map.

actual variances or precisions that are calculated in the VB framework, since this could artificially alter results for datasets that have a higher overall level of uncertainty. Instead, we are more interested in a relative weighting that accords a higher importance to voxels that have a lower estimate of uncertainty. In order to achieve this, we divide the estimate of each voxel by the normalized variance – i.e. the variance divided by the mean variance across the whole tumour. In the following Section, we calculate various measures of heterogeneity from the PK parameter maps, and weighted by their relative uncertainties. We will also see how these weightings affect the overall results that are calculated.

6.5 Experiments

To compare the predictive values of all these measures of heterogeneity, we computed each measure independently on pre-therapy DCE-MRI scans obtained for twenty patients with locally advanced rectal adenocarcinomas. All the patients were imaged before and after five weeks of long-course CRT, and the modified rectal cancer regression Grade (m-RCRG) was calculated when the tumour was resected at eight weeks. According to the m-RCRG, eleven of the twenty patients were found to be partial responders (m-RCRG= 2), while the other nine were found to be non-responders (m-RCRG= 3).

For acquiring and processing all the DCE-MRI data in this study, every effort was made to follow the international consensus developed by the Quantitative Imaging Biomarkers Alliance [Ashton, 2010]. All the data were acquired on a 1.5T GE scanner using a T_1-weighted, gradient-echo, fat-suppressed sequence (LAVA) with TR=4.5ms, TE=2.2ms and flip angle $\alpha = 12$ degrees. Four variable flip angle images (at 3, 9, 12 and 15 degrees) were acquired before the dynamic sequence, and the T_{10} was estimated using the procedure described in [Fram et al., 1987]. This map was used in the subsequent analysis. The contrast agent (MultiHance) was injected via a peripheral vein after 4 initial acquisitions and the MRI images were then acquired every 12 seconds for the next 5 minutes. For most patients, 27 volumes were acquired during the DCE-MRI scan, with each volume having a resolution of $1 \times 1 \times 2\,\mathrm{mm}^3$.

An expert radiologist segmented the tumour on high-resolution T_2-weighted volumes of the tumour, acquired before the dynamic sequence. This T_2-weighted volume was rigidly registered to the first volume of the dynamic scan, and the VB motion correction and parameter estimation algorithm was applied to an ROI containing the entire tumour microenvironment. The heterogeneity measures were then computed within the segmented tumour from the pre-therapy scan for each patient. We calculated each of these three measures for each of the three chosen parameters (k_a, K^{trans} and k_{ep}), as

well as versions that are weighted by uncertainties associated with them (Figure 6.9), as calculated from the VB algorithm. In order to compute each parameter map *with* uncertainty, the parameter map was divided point-wise by its corresponding variance (normalized by the sum of all variances across all voxels).

In order to compare the predictive values of all the measures of heterogeneity (computed within the segmented tumour), they were compared with the modified rectal cancer regression Grade (m-RCRG) for each patient, calculated after resection at eight weeks. According to the m-RCRG, eleven of the twenty patients were found to be partial responders (m-RCRG= 2), while the other nine were found to be non-responders (m-RCRG= 3).

The ability of each heterogeneity measure to predict response was statistically evaluated in two ways: (1) by performing a two-sample student t-test and computing the p-value between the quantities measured for the responder and non-responder groups; and (2) leave-one-out validation (LOOV) – each sample was classified (by fitting a multivariate normal density to each group and then assigning the sample value to the highest probability group) using all other samples as the training data, and the average proportion of samples that were mis-classified was recorded(Table 6.1). The heterogeneity measures were also compared with the mean and standard deviations of each parameter, in order to evaluate the importance of using spatial distributions as opposed to statistical summary values.

6.6 Results

In our tests, we compared three parameters: k_{ep}, K^{trans} and k_a. For each parameter, we calculated the p-value and the LOOV value for five different measures: mean, standard deviation, S_T, *sharpness* and *planarity*. To assess the impact of incorporating uncer-

		Mean		St-Dev		S_T		Sharpness		Planarity	
Uncertainty $\rightarrow$		w/o	with	w/o	with	w/o	with	w/o	with	w/o	with
k_{ep}	p-value	0.10	0.09	0.21	0.16	0.13	0.10	0.09	***0.08***	0.12	0.10
	LOOV	0.52	***0.43***	0.61	0.53	0.51	0.48	0.46	***0.43***	0.50	0.46
K^{trans}	p-value	0.09	0.07	0.13	0.11	0.04	0.03	0.04	***0.02***	0.07	0.05
	LOOV	0.42	0.38	0.48	0.41	0.24	0.21	0.25	***0.19***	0.38	0.33
k_a	p-value	0.04	0.03	0.08	0.07	0.03	0.02	0.01	0.01	0.004	***0.002***
	LOOV	0.28	0.25	0.41	0.38	0.25	0.20	0.14	0.10	0.10	***0.05***

Table 6.1: Statistical measures for the difference between the partial responder ($n = 11$) and non-responder ($n = 9$) groups for five measures using three PK parameters. Each measure is calculated on each parameter map with and without taking uncertainty into account. LOOV represents the proportion of samples mis-classified during leave-one-out validation. The minimum value in each row is indicated in bold.

tainty, we calculated each measure for each parameter with and without the uncertainty weighting.

From all the parameters that were tested, k_a was found to be the best predictor of response to therapy, showing a significant ($p < 0.05$) difference between responders and non-responders for the mean, S_T, sharpness and planarity measures (Figure 6.10). In general, we found that the measures of spatial heterogeneity showed consistently better results than the mean and standard deviation. It is also clear from Table 6.1, which the inclusion of uncertainty in the calculations of different measures consistently improved the ability of each measure to distinguish between responders and non-responders.

In the tests, it was found that the measure *planarity* calculated on the k_a map was most effective in distinguishing between the responder and non-responder groups – both on the basis of the t-tests, as well as the LOOV analysis. It is worthwhile remembering that due to the scaling of the variances calculated at each voxel, the overall heterogeneity measure with uncertainty will always be lower than the one calculated without uncertainty. Figure 6.10 illustrates how the *planarity* values corresponding to k_a vary across all 20 patients, with and without incorporating uncertainty. As the Figure shows, the gap between the responder and non-responder groups is wider when

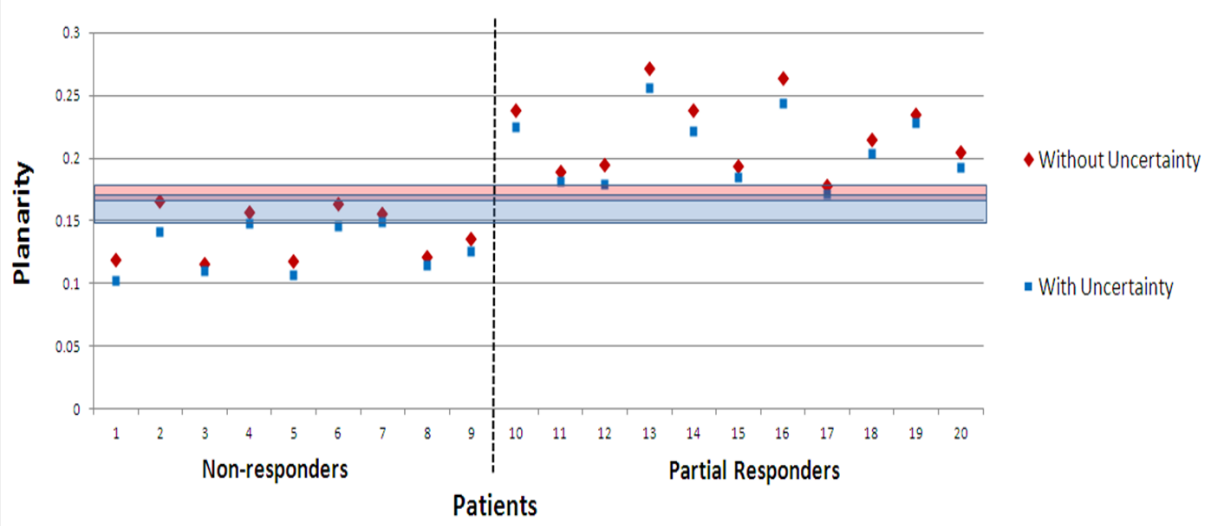

Figure 6.10: Plots for the measured 'planarity' in the k_a map across all 16 patients. The red dots indicate the values obtained *without* incorporating uncertainty, and the red gap (0.01) is the minimum difference between responders and non-responders. The blue dots indicate the values obtained *with* uncertainty, and the blue gap (0.03) is the minimum difference between responders and non-responders.

uncertainty is taken into account.

6.7 Discussion

There are various conclusions that can be drawn on the basis of Table 6.1. The results represent a combination of three factors:

1. The prognostic value of different PK parameters: k_{ep}, K^{trans} and k_a.

2. The advantages/disadvantages of spatial heterogeneity measures over summary measures such as mean and variance.

3. The advantages/disadvantages of incorporating uncertainty into the calculation of the spatial heterogeneity measures and the summary measures.

In order to examine each of these questions, we can revisit Table 6.1, and plot the p-values obtained for each of the three PK parameters individually. Figure 6.11 shows

that the distinction between the responder and non-responder groups becomes clearer (i.e. p-values become smaller) if we use K^{trans} instead of k_{ep}, and becomes clearer still, if we use k_a instead of K^{trans}. Another observation that can be made, is that the p-values are consistently lower when uncertainty is incorporated, for each of the summary/heterogeneity measures.

As far as the heterogeneity measures themselves are concerned, k_{ep} and K^{trans} show the strongest p-values for the the heterogeneity metric *Sharpness*, whereas k_a shows the strongest p-value for *Planarity*. It is interesting to note that for all three measures, the mean itself is a more reliable measure than the variance across all patients. This is probably because many non-responder tumours have a wide range of parameter values, leading to a high value for the variance. However, in such tumours, the high and low values are restricted to distinct areas of the tumour, leading to a low value for other heterogeneity metrics and the mean itself. On the other hand, responder tumours – with high and low-parameter values neighboring each other – also show high values for variance, thus making it hard to distinguish between the two on the basis of the variance alone.

Since the heterogeneity measures *do* take the spatial arrangements into account, they perform better than the variance in all cases, and better than the mean in most cases. To be specific, the heterogeneity measure *sharpness* performs better than the mean for all parameters, according to the p-value as well as the LOOV measure. The other spatial heterogeneity measures – S_T and *planarity* – perform better than the mean for the parameters K^{trans} and k_a, though they perform marginally worse for the parameter k_{ep}.

It is evident from Table 6.1 that k_{ep} is relatively worse than K^{trans} in distinguishing between the two patient groups. This is possibly because k_{ep} captures the heterogeneity in the shape of the contrast-uptake curve, while K^{trans} captures changes in both

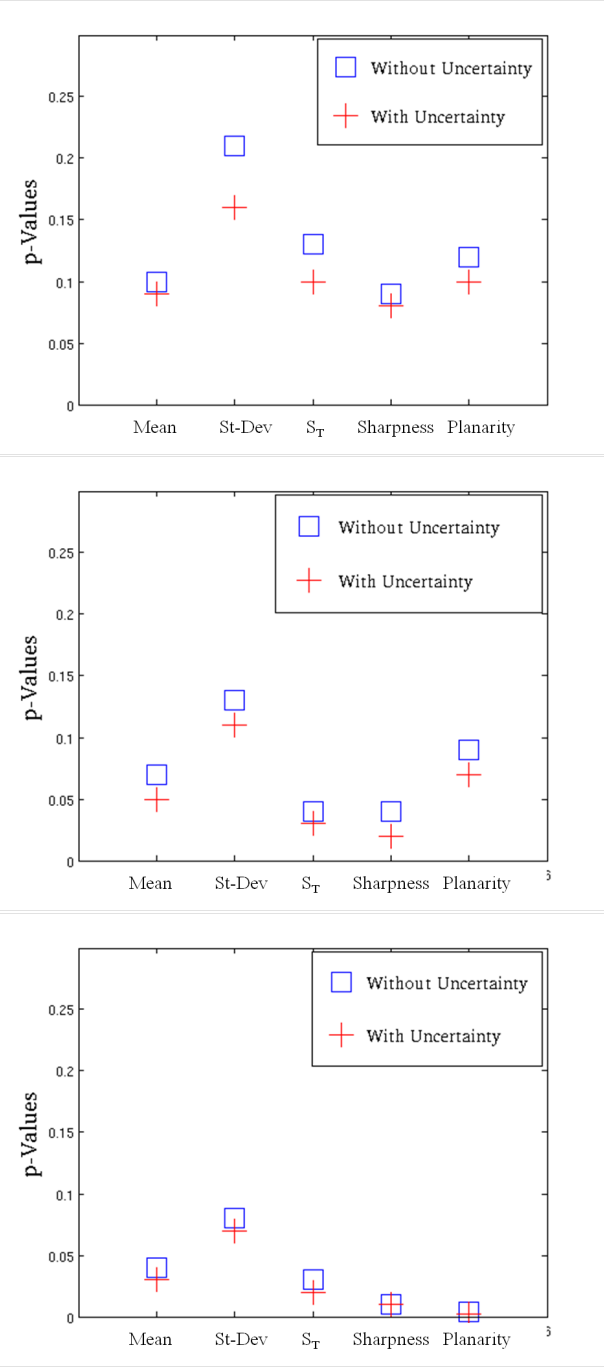

Figure 6.11: The p-values corresponding to each summary/heterogeneity measure for k_{ep} (top), K^{trans} (middle) and k_a (bottom) calculated across the group of 20 patients.

the shape and the amplitude of the curve, both of which are likely to be indicative of response to CRT. On the other hand, the acceleration constant, k_a, captures the heterogeneity in the shape, amplitude and starting point of the curve.

The reason that k_a performs so much better than the other parameters, is the fact that it represents a physiological characteristic of the tissue – namely its ability to respond to the force exerted by the bolus, per unit mass. The strong negative correlation between the onset time, t_0, and K^{trans}, as well as the sharp maps for k_a indicate that the ratio between K^{trans} and t_0 does indeed have a biological meaning. This ability of the vasculature to respond to a force exerted by the bolus, is likely to be indicative of how well the vasculature would respond to chemotherapeutic drugs delivered to that area.

6.8 Conclusion

The results indicate that spatial heterogeneity of PK parameters in general, and the acceleration constant, $k_a = K^{trans}/t_o$, in particular, are very important for predicting response to therapy on the basis of only the pre-therapy DCE-MRI scan. The results also show that the inclusion of uncertainty in the calculation of heterogeneity measures causes a marked improvement in their ability to distinguish between responders and non-responders to CRT.

In future, it would be interesting to examine whether k_a has the same prognostic value in tests on other kinds of data (not just colorectal cancer data), and whether the negative correlation between K^{trans} and t_0, is evident in other data as well. We would also need to examine whether the k_a maps display the same kind of linear/planar structures in tumours that have more irregular and asymmetrical shapes.

The measures of spatial heterogeneity used in this study were motivated by the data

itself, and on the physical characteristics of the tumours being examined. Although these characteristics (a fast-enhancing rim, intra-tumoural vascular structures, etc) are generic to most tumours, it is quite possible that other kinds of oncological data may necessitate the development of other kinds of spatial heterogeneity measures. Thus more tests are required to test the robustness of the heterogeneity measures used in this study.

In the tests conducted in this study, we found that the inclusion of uncertainty in PK parameter estimates improves the ability of spatial heterogeneity measures in distinguishing between responder and non-responder groups. In future, it is important to examine whether this finding is replicated in other kinds of data, especially in cases where motion can be constrained. It is also necessary to include individually determined AIFs within the PK model, and then determine whether uncertainty in the AIF estimate has a significant impact on prediction of response to therapy.

Chapter 7

Conclusions and Outlook

As discussed in Chapter 2, DCE-MRI data can provide very valuable insights into tumour physiology. However, one of the greatest difficulties in the analysis of DCE-MRI data, and in making clinical predictions based on it, has been the corruption of the data due to patient motion occurring during a scan.

7.1 Summary

In this thesis, we first presented the MoRPE framework for motion correction and PK parameter estimation in Chapter 3. The method was tested with different deformation models, and it was found that the diffeomorphic Demons method resulted in a lower TRE than the regular Demons algorithm in a test on simulated data. The regular Demons algorithm also resulted in negative Jacobians (indicating folding of the tissue) in parts of the deformation field. Thus, we opted to use diffeomorphic Demons for the subsequent tests on simulated and real data.

The primary objective of the tests on simulated data was to compare the efficacy of MoRPE with other conventional similarity measures, in being able to recover the original PK parameter maps and deformations that are used to generate the simulated

data. In the tests, it was found that MoRPE resulted in the lowest error between the recovered and original parameter maps. It also provided the lowest TRE, implying that it performed better than the other algorithms in being able to recover the deformations applied to the simulated data.

The tests on simulated data indicated that conventional similarity measures such as SSD, MI and NCC were unable to cope with the large changes in intensities between consecutive time frames, caused by the fast uptake of CA. The results showed that MoRPE performed much better than the other methods, since it was was based on a physiological model of contrast uptake.

The method that performed best after MoRPE in the tests on simulated data, was the algorithm using NCC as the cost function. Thus, in our subsequent tests on clinical data, we restricted ourselves to comparisons between MoRPE and NCC. The main purpose behind the tests on clinical data, was to determine whether the analysis of DCE-MRI data using MoRPE enables patient stratification or prediction of response in a more effective manner than using other algorithms. The MoRPE framework was applied to data from a colorectal cancer study in Chapter 4, and was found to be highly successful in distinguishing between responders and non-responders on the basis of the difference between K^{trans} maps measured before and after five weeks of CRT.

Within the framework of MoRPE, we were seeking to investigate the impact of (a) motion correction; (b) different PK parameters; and (c) the choice of the AIF; on our ability to distinguish between responders and non-responders to CRT. Although the tests on clinical data were conducted on a reasonably small cohort of patients, we demonstrated that using MoRPE enables a better discrimination between responders and non-responders than estimating PK-parameters through direct curve-fits or using NCC-based motion correction followed by curve-fits. The results showed that K^{trans} is the PK parameter that is most successful in discriminating between responders and

non-responders, followed by k_{ep} and then by v_e. We also showed that using the Orton AIF within the framework of MoRPE provides better results than using the Weinmann AIF.

The results in Chapter 4 clearly demonstrated that the motion correction method can have a big impact on the difference between PK parameter maps estimated at different time points (for the same patient), and therefore on the prediction of response based on these differences. The results showed that there is a clear clinical benefit of using MoRPE instead of NCC-based motion correction. The results also showed that the choice of the AIF substantially affects predictions about the response to CRT, and that there is a noticeable clinical benefit of using AIF models that simulate bolus arrival in a realistic manner.

A quantitative analysis of how well the PK model represents the data under different circumstances, was presented in Chapter 5 using a VB mathematical framework. In Section 5.5.1, the tests on simulated data showed that there is broad agreement between the results obtained with MoRPE, and with the VB algorithm, when tested on the same data. The results also showed that the error between the ground truth PK parameter maps and the ones estimated using MoRPE or the VB algorithm, was usually around $0.5 - 1$ standard deviations (as estimated by the VB algorithm). This showed that the variances being estimated by the VB algorithm were realistic, and in accordance with other results.

The tests on pre-clinical data demonstrated that the variances estimated for k_{ep}, v_e and t_0 were reduced by 20-40% of the values that were estimated before motion correction (Section 5.5.2.2), while the variances in X_0 and T_{10} saw a marginal reduction. The same findings were reinforced by the tests on clinical data in Section 5.5.3.2. This substantial reduction indicates that motion correction is especially important for calculating precise values of k_{ep}, v_e and t_0.

The analysis of noise contributions before and after motion correction in Section 5.6 further demonstrated the benefits of motion correction by showing that the VB estimate of the noise precision becomes higher when motion correction is incorporated into the analysis framework. The relatively high contribution of T_{10} without *and* with motion correction indicates that imaging protocols need to be designed in a manner that facilitates accurate computation of the T_{10} map. The high contribution of t_0, and the prognostic value of $k_a = K^{trans}/t_0$ also indicates the importance of determining the onset time more accurately, and perhaps using imaging protocols with a higher temporal resolution.

The results in Section 5.6.2 also showed that the percentage contribution of the 'model-fit' component (or the residual) decreases considerably as a result of motion correction, and is compensated by an increased contribution from T_{10} and the pre-contrast MRI intensity (X_0). The analysis of the correlation coefficients between different PK parameters in Section 5.7 revealed that the strongest absolute correlation is found between K^{trans} and t_0 (mean = -0.32), followed by the correlation between k_{ep} and t_0 (mean = -0.23).

In spite of all these findings, the incorporation of uncertainties in calculating the KS-distance between K^{trans} maps showed no significant impact on the prediction of response to therapy. The most likely reason for this was the fact that the uncertainties associated with PK parameters were related to their spatial distributions, and not to their magnitudes.

In order to investigate the importance of spatial distributions of PK parameters, in Chapter 6 we discuss the importance of estimating heterogeneity in PK parameter maps. Different measures of heterogeneity calculated from the maps of K^{trans}, k_{ep}, and a new derived biomarker $k_a = K^{trans}/t_o$, were used to predict outcome to therapy from the pre-therapy scan alone. The results in Section 6.6 demonstrated the importance

of spatial heterogeneity measures of PK parameters in general, and the acceleration constant, (k_a) in particular, in predicting response to therapy. The results also showed that the inclusion of uncertainty in the calculation of heterogeneity measures causes a marked improvement in their ability to distinguish between responders and non-responders to CRT, even though it had had no significant impact on prediction on the basis of KS-distance.

The results in Chapter 5 and 6 indicate that the uncertainties or variances associated with PK parameter estimates for individual voxels are more closely associated with their physical locations rather than their magnitudes. The uncertainties are higher for voxels located near tissue boundaries, where motion is likely to have a larger effect.

7.2 Future Work

In this thesis, we have used the Tofts PK model (with the Weinmann/Orton AIF) throughout, but as discussed in Section 4.5.3, it is possible that the use of other PK models and parameters may further improve the ability of the MoRPE framework to predict response to therapy. This needs to be investigated in the future. The impact of using inidvidually determined AIFs on prediction of response also needs to be investigated, and the MoRPE framewok needs to be extended so that it can incorporate the same.

The comparison between different PK models and between population-averaged AIFs and individually determined AIFs can be carried out in exactly the same manner as the comparison between the Orton and Weinmann AIF was carried out in Section 4.3 – i.e. by comparing their ability to distinguish between responders and non-responders in a clinical study of patients with solid tumours receiving neoadjuvant CRT.

We have examined the uncertainties in several PK parameters as well as T_{10} and

X_0 in Chapter 5. However, an aspect that needs to be examined in the future, is the uncertainty in the parameters governing the shape of the AIF itself. In order to facilitate a comprehensive analysis of the uncertainty in the AIF and its impact on prediction of response, it is necessary to calculate the AIF individually for each patient from high temporal resolution data, using the most appropriate method available [Yang et al., 2004].

In Chapter 4 we had used the difference between PK parameter maps estimated before and after CRT to predict the response to therapy, whereas in Chapter 6, we had used heterogeneity measures calculated from only the pre-therapy scan to do the same. In the near future, we will investigate whether the accuracy of predictions can be improved further by examining the difference in the heterogeneity measures between pre- and post-therapy scans. We will also investigate the prognostic value of other heterogeneity measures, and of other combinations of PK parameters.

The prognostic value of the acceleration constant (k_a) defined in Chapter 6 needs to be investigated on other data for other kinds of solid tumours, and the physiological meaning of 'acceleration' needs to be further examined with reference to the differential equations involved in defining PK models. The heterogeneity measures 'sharpness' and 'planarity' were found to be very successful in predicting response to therapy in the colorectal cancer patients studied in this thesis. However, we need to apply these measures to data for breast and liver cancer, to test how robust and generic these measures are.

One of the recent developments in cancer treatment has been the advent of anti-angiogenic drugs such as Nelfinavir, which are meant to normalize the vasculature and thereby ensure more effective delivery of drugs or oxygen to the tumour tissues. The role of DCE-MRI can be extremely crucial in the task of assessing the impacts of anti-angiogenic drugs on the tumour tissues, as has been shown in [George et al., 2001] and

[Sauerbier et al., 2009]. In this thesis, we have used the MoRPE framework to predict response to CRT in colorectal cancer patients. In future, this method can be tested for prediction of response to other kinds of treatment protocols (such as administration of anti-angiogenic drugs in conjunction with CRT) on other kinds of solid tumours.

In [Wynn et al., 2010], a survey conducted amongst British surgeons showed the extremely low level of confidence in any measure showing complete response to neoadjuvant chemo/radiotherapy in colorectal cancer patients, even though the percentage of complete responders is thought to be about 15%. This confidence is even lower when it comes to non-invasive methods of determining response such as DCE-MRI. However, with rapidly improving clinical protocols for CRT, the percentage of complete responders is bound to increase, and thus it is imperative to develop non-invasive methods of determining the extent of response to CRT. This thesis presents a probabilistic framework for motion correction and PK analysis of DCE-MRI data, that can be used to make clinical predictions about the response of colorectal cancer patients to CRT. The thesis also presents methods for estimating heterogeneity and uncertainty in PK parameter maps, which will be built upon in the near future to improve the accuracy of clinical predictions in different kinds of cancer data and clinical settings.

List of Publications

1. M Bhushan, JA Schnabel, M Chappell, F Gleeson, M Anderson, J Franklin, JM Brady, M Jenkinson. **The Impact of Heterogeneity and Uncertainty on Prediction of Response to Therapy Using Dynamic MRI Data.** Proc. Medical Image Computing and Computer Assisted Intervention (MICCAI) 2013 (to appear).

2. M Bhushan, M Enescu, C Galerneau, M Anderson, F Gleeson, M Brady, JA Schnabel, M Jenkinson. **Quantifying Response to CRT in Colorectal Cancer Patients from Dynamic Imaging.** Proc. Novel technology that shapes radiology: EIBIR presents IMAGINE, European Congress of Radiology (ECR), 2013.

3. M Bhushan, JA Schnabel, LN Tanner, F Gleeson, JM Brady, M Jenkinson. **Simultaneous Bayesian Estimation of Motion and Pharmacokinetic Parameters in dceMRI Sequences for Discrimination of Responders in Colorectal Cancer.** Proc. International Symposium for Magnetic Resonance in Medicine (ISMRM) 2012.

4. M Bhushan, JA Schnabel, MP Heinrich L Risser, M Brady, M Jenkinson. **Motion Correction and Parameter Estimation in dceMRI sequences: Application to Colorectal Cancer.** Proc. Medical Image Computing and Computer

Assisted Intervention (MICCAI) 2011.

5. B Irving, L Tanner, M Enescu, M Bhushan, EJ Hill, J Franklin, JA Schnabel and M Brady. **Personalized Estimation of the Arterial Input Function for Improved Pharmacokinetic Modelling of Colorectal Cancer using dceMRI.** MICCAI Workshop on Abdominal Imaging, 2013 (to appear).

6. M Enescu, M Bhushan, EJ Hill, J Franklin, EM Anderson, RA Sharma and JA Schnabel. **pCT Derived Arterial Input Function for Improved Pharmacokinetic Analysis of Longitudinal dceMRI for Colorectal Cancer.** Proc. SPIE Medical Imaging, 2013.

7. MP Heinrich, M Jenkinson, M Bhushan, T Matin, F Gleeson, M Brady, JA Schnabel. **NonMIND: Modality Independent Neighbourhood Descriptor for Multi-modal Deformable Registration.** Proc. Medical Image Analysis, Vol. 16(7) Oct. 2012, Special Issue on MICCAI 2011.

8. MP Heinrich, M Jenkinson, M Bhushan, T Matin, F Gleeson, M Brady, JA Schnabel. **Non-local Shape Descriptor: A new Similarity Metric for Deformable Multimodal Registration.** Proc. Medical Image Computing and Computer Assisted Intervention (MICCAI) 2011.

Bibliography

[Armitage et al., 2005] Armitage, P., Behrenbruch, C., Brady, M., and Moore, N. (2005). Extracting and visualizing physiological parameters using dynamic contrast-enhanced Magnetic Resonance Imaging of the breast. *Medical Image Analysis*, 9(4):315–329.

[Arsigny et al., 2006] Arsigny, V., Commowick, O., Pennec, X., and Ayache, N. (2006). A log-Euclidean framework for statistics on diffeomorphisms. *MICCAI: International Conference on Medical Image Computing and Computer-Assisted Intervention*, 9(Pt 1):924–931.

[Arsigny et al., 2005] Arsigny, V., Pennec, X., and Ayache, N. (2005). Polyrigid and polyaffine transformations: a novel geometrical tool to deal with non-rigid deformations - application to the registration of histological slices. *Medical Image Analysis*, 9(6):507–523.

[Ashburner, 2007] Ashburner, J. (2007). Fast diffeomorphic image registration algorithm. *NeuroImage*, 38(1):95–113.

[Ashton, 2010] Ashton, E. (2010). Quantitative MR in multi-center clinical trials. *Journal of Magnetic Resonance Imaging : JMRI*, 31(2):279–288.

[Atkin et al., 2006] Atkin, G., Taylor, N. J., Daley, F. M., Stirling, J. J., Richman, P., Glynne-Jones, R., D'Arcy, J. A., Collins, D. J., and Padhani, A. R. (2006). Dynamic

contrast-enhanced Magnetic Resonance Imaging is a poor measure of rectal cancer angiogenesis. *British Journal of Surgery*, 93(8):992–1000.

[Avants et al., 2008a] Avants, B. B., Epstein, C. L., Grossman, M., and Gee, J. C. (2008a). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1):26.

[Avants et al., 2008b] Avants, B. B., Epstein, C. L., Grossman, M., and Gee, J. C. (2008b). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1):26.

[Barboriak et al., 2008] Barboriak, D. P., MacFall, J. R., Viglianti, B. L., and Dewhirst Dvm, M. W. (2008). Comparison of three physiologically-based pharmacokinetic models for the prediction of contrast agent distribution measured by dynamic MR imaging. *Journal of Magnetic Resonance Imaging : JMRI*, 27(6):1388–1398.

[Bateman et al., 2009] Bateman, A. C., Jaynes, E., and Bateman, A. R. (2009). Rectal cancer staging post neoadjuvant therapy–how should the changes be assessed? *Histopathology*, 54(6):713–721.

[Beg et al., 2005a] Beg, M. F., Miller, M. I., Trouvé, A., and Younes, L. (2005a). Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157.

[Beg et al., 2005b] Beg, M. F., Miller, M. I., Trouvé, A., and Younes, L. (2005b). Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157.

[Benjaminsen et al., 2004] Benjaminsen, I. C., Graff, B. r. A., Brurberg, K. G., and Rofstad, E. K. (2004). Assessment of tumor blood perfusion by high-resolution dynamic contrast-enhanced MRI: a preclinical study of human melanoma xenografts. *Magnetic Resonance in Medicine*, 52(2):269–76.

[Berger, 1985] Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer.

[Berns et al., 2003] Berns, E. M. J. J., Klijn, J. G. M., Look, M. P., Grebenchtchikov, N., Vossen, R., Peters, H., Geurts-Moespot, A., Portengen, H., van Staveren, I. L., Meijer-van Gelder, M. E., Bakker, B., Sweep, F. C. G. J., and Foekens, J. A. (2003). Combined Vascular Endothelial Growth Factor and TP53 Status Predicts Poor Response to Tamoxifen Therapy in Estrogen Receptor-positive Advanced Breast Cancer. *Clin. Cancer Res.*, 9(4):1253–1258.

[Bertout et al., 2008] Bertout, J. A., Patel, S. A., and Simon, M. C. (2008). The impact of O2 availability on human cancer. *Nat Rev Cancer*, 8(12):967–975.

[Bhushan et al., 2013] Bhushan, M., Schnabel, J. A., Chappell, M., Gleeson, F., Anderson, M., Franklin, J., Brady, S. M., and Jenkinson, M. (2013). The Impact of Heterogeneity and Uncertainty on Prediction of Response to Therapy Using Dynamic MRI Data. *MICCAI: International Conference on Medical Image Computing and Computer-Assisted Intervention*, 8149:316–323.

[Bhushan et al., 2011] Bhushan, M., Schnabel, J. A., Risser, L., Heinrich, M. P., Brady, J. M., and Jenkinson, M. (2011). Motion correction and parameter estimation in dceMRI sequences: application to colorectal cancer. *MICCAI: International Conference on Medical Image Computing and Computer-Assisted Intervention*, 14(Pt 1):476–483.

[Bhushan et al., 2012] Bhushan, M., Schnabel, J. A., Tanner, L., Gleeson, F., Brady, J. M., and Jenkinson, M. (2012). Simultaneous Bayesian Estimation of Motion and Pharmacokinetic Parameters in Dynamic Contrast-Enhanced MRI for the Discrimination of Responders in Colorectal Cancer. *ISMRM: International Symposium on Magnetic Resonance in Medicine*, 1(Pt 1):1974.

[Bradley et al., 2008] Bradley, D. P., Tessier, J. L., Checkley, D., Kuribayashi, H., Waterton, J. C., Kendrew, J., and Wedge, S. R. (2008). Effects of AZD2171 and vandetanib (ZD6474, Zactima) on haemodynamic variables in an SW620 human colon tumour model: an investigation using dynamic contrast-enhanced MRI and the rapid clearance blood pool contrast agent, P792 (gadomelitol). *NMR in biomedicine*, 21(1):42–52.

[Bro-Nielsen and Gramkow, 1996] Bro-Nielsen, M. and Gramkow, C. (1996). Fast fluid registration of medical images. *Visualization in biomedical computing*, pages 265–276.

[Brown et al., 1997] Brown, L. F., Detmar, M., Claffey, K., Nagy, J. A., Feng, D., Dvorak, A. M., and Dvorak, H. F. (1997). Vascular permeability factor/vascular endothelial growth factor: a multifunctional angiogenic cytokine. *EXS*, 79:233–69.

[Buckley, 2002] Buckley, D. L. (2002). Uncertainty in the analysis of tracer kinetics using dynamic contrast-enhanced T1-weighted MRI. *Magnetic Resonance in Medicine*, 47(3):601–606.

[Buonaccorsi et al., 2005] Buonaccorsi, G. A., Roberts, C., Cheung, S., Watson, Y., Davies, K., Jackson, A., Jayson, G. C., and Parker, G. J. M. (2005). Tracer kinetic model-driven registration for dynamic contrast enhanced MRI time series. *MICCAI: International Conference on Medical Image Computing and Computer Assisted Intervention*, 8(Pt 1):91–98.

[Chan et al., 2003] Chan, H. M., Chung, A. C. S., Yu, S. C. H., Norbash, A., and Wells III, W. M. (2003). Multi-modal image registration by minimizing kullback-leibler distance between expected and observed joint class histograms. *MICCAI: International Conference on Medical Image Computing and Computer-Assisted Intervention*.

[Chappell et al., 2009] Chappell, M., Groves, A., Whitcher, B., and Woolrich, M. (2009). Variational Bayesian Inference for a Nonlinear Forward Model. *IEEE Transactions on Signal Processing*, 57(1):223–236.

[Checkley et al., 2003] Checkley, D., Tessier, J. J., Kendrew, J., Waterton, J. C., and Wedge, S. R. (2003). Use of dynamic contrast-enhanced MRI to evaluate acute treatment with ZD6474, a VEGF signalling inhibitor, in PC-3 prostate tumours. *British journal of cancer*, 89(10):1889–95.

[Christensen, 1999] Christensen, G. (1999). Consistent linear-elastic transformations for image matching. *Information processing in medical imaging*, pages 224–237.

[Chung et al., 2002] Chung, A., Wells, W., Norbash, A., and Grimson, W. (2002). Multi-modal image registration by minimising kullback-leibler distance. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2002*, pages 525–532.

[Collignon et al., 1995] Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., and Marchal, G. (1995). Automated multi-modality image registration based on information theory. *Information processing in medical imaging*, 14(6):263.

[Cooper et al., 2000] Cooper, R. A., Carrington, B. M., Loncaster, J. A., Todd, S. M., Davidson, S. E., Logue, J. P., Luthra, A. D., Jones, A. P., Stratford, I., Hunter, R. D., and West, C. M. (2000). Tumour oxygenation levels correlate with dynamic contrast-enhanced Magnetic Resonance Imaging parameters in carcinoma of the cervix. *Ra-*

diotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology, 57(1):53–9.

[Coupe et al., 2008] Coupe, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., and Barillot, C. (2008). An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Transactions on Medical Imaging*, 27(4):425–441.

[D'Agostino et al., 2002] D'Agostino, E., Maes, F., Vandermeulen, D., and Suetens, P. (2002). Viscous fluid model for multimodal non-rigid image registration using mutual information. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2002*, pages 541–548.

[De Naeyer et al., 2011] De Naeyer, D., Verhulst, J., Ceelen, W., Segers, P., De Deene, Y., and Verdonck, P. (2011). Flip angle optimization for dynamic contrast-enhanced MRI-studies with spoiled gradient echo pulse sequences. *Physics in medicine and biology*, 56(16):5373–95.

[Deoni et al., 2003] Deoni, S. C., Rutt, B. K., and Peters, T. M. (2003). Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state. *Magnetic Resonance in Medicine*, 49(3):515–526.

[Dinter et al., 2009] Dinter, D. J., Horisberger, K., Zechmann, C., Wenz, F., Brade, J., Willeke, F., Neff, K. W., Schoenberg, S. O., and Hofheinz, R. D. (2009). Can dynamic MR imaging predict response in patients with rectal cancer undergoing cetuximab-based neoadjuvant chemoradiation? *Onkologie*, 32(3):86–93.

[Donaldson et al., 2010a] Donaldson, S. B., Buckley, D. L., O'Connor, J. P., Davidson, S. E., Carrington, B. M., Jones, A. P., and West, C. M. L. (2010a). Enhancing fraction measured using dynamic contrast-enhanced MRI predicts disease-free survival in patients with carcinoma of the cervix. *British journal of cancer*, 102(1):23–6.

[Donaldson et al., 2010b] Donaldson, S. B., West, C. M. L., Davidson, S. E., Carrington, B. M., Hutchison, G., Jones, A. P., Sourbron, S. P., and Buckley, D. L. (2010b). Comparison of tracer kinetic models for T1-weighted dynamic contrast-enhanced MRI: application in carcinoma of the cervix. *Magnetic Resonance in Medicine*, 63(3):691–700.

[Dowson et al., 2008] Dowson, N., Kadir, T., and Bowden, R. (2008). Estimating the joint statistics of images using nonparametric windows with application to registration using mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(10):1841–1857.

[Du et al., 2010] Du, J., Karimi, A., Wu, Y., Korosec, F. R., Grist, T. M., and Mistretta, C. A. (2010). Automated vessel segmentation using cross-correlation and pooled covariance matrix analysis. *Magnetic Resonance Imaging*.

[Dvorak et al., 1995] Dvorak, H. F., Brown, L. F., Detmar, M., and Dvorak, A. M. (1995). Vascular permeability factor/vascular endothelial growth factor, microvascular hyperpermeability, and angiogenesis. *The American journal of pathology*, 146(5):1029–39.

[Dzik-Jurasz et al., 2004] Dzik-Jurasz, A. S., Walker-Samuel, S., Leach, M. O., Brown, G., Padhani, A., George, M., and Collins, D. J. (2004). Fractal parameters derived from analysis of DCE-MRI data correlates with response to therapy in rectal carcinoma. In *12th International Society for Magnetic Resonance in Medicine*, page 2503, Kyoto, Japan.

[Ertas et al., 2007] Ertas, G., Gulcur, H. O., and Tunaci, M. (2007). Improved lesion detection in MR mammography: three-dimensional segmentation, moving voxel sam-

pling, and normalized maximum intensity-time ratio entropy. *Academic Radiology*, 14(2):151–61.

[Ferrant et al., 1999] Ferrant, M., Warfield, S., Guttmann, C., Mulkern, R., Jolesz, F., and Kikinis, R. (1999). 3D image matching using a finite element based elastic deformation model. *Medical Image Computing and Computer-Assisted Intervention–MICCAI'99*, pages 202–209.

[Fluckiger et al., 2009] Fluckiger, J. U., Schabel, M. C., and Dibella, E. V. R. (2009). Model-based blind estimation of kinetic parameters in dynamic contrast enhanced (DCE)-MRI. *Magnetic Resonance in Medicine*, 62(6):1477–86.

[Foryś and Mokwa-Borkowska, 2005] Foryś, U. and Mokwa-Borkowska, A. (2005). Solid tumour growth analysis of necrotic core formation. *Mathematical and Computer Modelling*, 42(5):593–600.

[Fram et al., 1987] Fram, E. K., Herfkens, R. J., Johnson, G. A., Glover, G. H., Karis, J. P., Shimakawa, A., Perkins, T. G., and Pelc, N. J. (1987). Rapid calculation of T1 using variable flip angle gradient refocused imaging. *Magnetic Resonance Imaging*, 5(3):201–208.

[Fritz-Hansen et al., 1996] Fritz-Hansen, T., Rostrup, E., Larsson, H. B. W., Søndergaard, L., Ring, P., and Henriksen, O. (1996). Measurement of the arterial concentration of gd-dtpa using MRI: a step toward quantitative perfusion imaging. *Magnetic Resonance in Medicine*, 36(2):225–231.

[Garpebring et al., 2012] Garpebring, A., Brynolfsson, P., Yu, J., Wirestam, R., Johansson, A., Asklund, T., and Karlsson, M. (2012). Uncertainty estimation in dynamic contrast-enhanced MRI. *Magnetic Resonance in Medicine*, 69:992–1002.

[George et al., 2001] George, M. L., Dzik-Jurasz, A. S. K., Padhani, A. R., Brown, G., Tait, D. M., Eccles, S. A., and Swift, R. I. (2001). Non-invasive methods of assessing angiogenesis and their value in predicting response to treatment in colorectal cancer. *British Journal of Surgery*, 88(12):1628–1636.

[Gollub et al., 2012] Gollub, M. J., Gultekin, D. H., Akin, O., Do, R. K., Fuqua, J. L., Gonen, M., Kuk, D., Weiser, M., Saltz, L., Schrag, D., Goodman, K., Paty, P., Guillem, J., Nash, G. M., Temple, L., Shia, J., and Schwartz, L. H. (2012). Dynamic contrast enhanced-MRI for the detection of pathological complete response to neoadjuvant chemotherapy for locally advanced rectal cancer. *Eur Radiol*, 22(4):821–831.

[Hashemi et al., 2010] Hashemi, R. H., Bradley, W. G., and Lisanti, C. J. (2010). *MRI: The basics*. Lippincott Williams & Wilkins.

[Hawighorst et al., 1998] Hawighorst, H., Weikel, W., Knapstein, P. G., Knopp, M. V., Zuna, I., Schönberg, S. O., Vaupel, P., and van Kaick, G. (1998). Angiogenic activity of cervical carcinoma: assessment by functional Magnetic Resonance Imaging-based parameters and a histomorphological approach in correlation with disease outcome. *Clinical Cancer Research*, 4(10):2305–12.

[Hayton et al., 1997] Hayton, P., Brady, M., Tarassenko, L., and Moore, N. (1997). Analysis of dynamic {MR} breast images using a model of contrast enhancement. *Medical Image Analysis*, 1(3):207–224.

[Hill et al., 1994] Hill, D. L., Studholme, C., and Hawkes, D. J. (1994). Voxel similarity measures for automated image registration. *Article of SPIE*, 2359:205.

[Hockel et al., 1996] Hockel, M., Schlenger, K., Aral, B., Mitze, M., Schaffer, U., and Vaupel, P. (1996). Association between Tumor Hypoxia and Malignant Progression in Advanced Cancer of the Uterine Cervix. *Cancer Res.*, 56(19):4509–4515.

[Hollander and Wolfe, 1973] Hollander, M. and Wolfe, D. A. (1973). Nonparametric statistical methods.

[Issa et al., 1999] Issa, B., Buckley, D. L., and Turnbull, L. W. (1999). Heterogeneity analysis of Gd-DTPA uptake: improvement in breast lesion differentiation. *Journal of Computer Assisted Tomography*, 23(4):615–621.

[Jackson et al., 2007] Jackson, A., O'Connor, J. P. B., Parker, G. J. M., and Jayson, G. C. (2007). Imaging tumor vascular heterogeneity and angiogenesis using dynamic contrast-enhanced Magnetic Resonance Imaging. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 13(12):3449–3459.

[James et al., 1999] James, K., Eisenhauer, E., Christian, M., Terenziani, M., Vena, D., Muldal, A., and Therasse, P. (1999). Measuring Response in Solid Tumors: Unidimensional Versus Bidimensional Measurement. *JNCI Journal of the National Cancer Institute*, 91(6):523–528.

[Jemal et al., 2011] Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., and Forman, D. (2011). Global cancer statistics 2011. *CA: A Cancer Journal for Clinicians*, 61(2):69–90.

[Jemal et al., 2009] Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., and Thun, M. J. (2009). Cancer statistics, 2009. *CA: A Cancer Journal for Clinicians*, 59(4):225.

[Jenkinson et al., 2012] Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., and Smith, S. M. (2012). FSL. *NeuroImage*, 62(2):782–790.

[Jia et al., 2008] Jia, G., O'Dell, C., Heverhagen, J. T., Yang, X., Liang, J., Jacko, R. V., Sammet, S., Pellas, T., Cole, P., and Knopp, M. V. (2008). Colorectal liver metastases: contrast agent diffusion coefficient for quantification of contrast enhancement heterogeneity at MR imaging. *Radiology*, 248(3):901–9.

[Jiang et al., 2004] Jiang, L., Zhao, D., Constantinescu, A., and Mason, R. P. (2004). Comparison of BOLD contrast and Gd-DTPA dynamic contrast-enhanced imaging in rat prostate tumor. *Magnetic Resonance in Medicine*, 51(5):953–60.

[Joshi and Brady, 2009] Joshi, N. and Brady, M. (2009). Non-Parametric Mixture Model Based Evolution of Level Sets and Application to Medical Images. *International Journal of Computer Vision*, 88(1):52–68.

[Karsa et al., 2010] Karsa, L., Lignini, T. A., Patnick, J., Lambert, R., and Sauvaget, C. (2010). Dimensions of the CRC problem. *Best Practice & Research Clinical Gastroenterology*, 24(4):381–396.

[Koenderink, 1984] Koenderink, J. J. (1984). The structure of images. *Biological cybernetics*, 50(5):363–70.

[Lee et al., 2000] Lee, J. C., Chow, N. H., Wang, S. T., and Huang, S. M. (2000). Prognostic value of vascular endothelial growth factor expression in colorectal cancer patients. *European journal of cancer (Oxford, England : 1990)*, 36(6):748–53.

[Li et al., 2011] Li, X., Welch, E. B., Arlinghaus, L. R., Chakravarthy, A. B., Xu, L., Farley, J., Loveless, M. E., Mayer, I. A., Kelley, M. C., Meszoely, I. M., Means-Powell, J. A., Abramson, V. G., Grau, A. M., Gore, J. C., and Yankeelov, T. E. (2011). A novel AIF tracking method and comparison of DCE-MRI parameters using individual and population-based AIFs in human breast cancer. *Physics in Medicine and Biology*, 56(17):5753–5769.

[Liang and Lauterbur, 2000] Liang, Z.-P. and Lauterbur, P. C. (2000). *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*. SPIE Optical Engineering Press.

[Liney et al., 1999] Liney, G. P., Gibbs, P., Hayes, C., Leach, M. O., and Turnbull, L. W. (1999). Dynamic contrast-enhanced MRI in the differentiation of breast tumors: user-defined versus semi-automated region-of-interest analysis. *Journal of Magnetic Resonance Imaging : JMRI*, 10(6):945–9.

[Loncaster et al., 2002] Loncaster, J. A., Carrington, B. M., Sykes, J. R., Jones, A. P., Todd, S. M., Cooper, R., Buckley, D. L., Davidson, S. E., Logue, J. P., Hunter, R. D., and West, C. M. L. (2002). Prediction of radiotherapy outcome using dynamic contrast enhanced MRI of carcinoma of the cervix. *International journal of radiation oncology, biology, physics*, 54(3):759–67.

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

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

[Marquardt, 1963] Marquardt, D. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431 – 441.

[Mayr et al., 2000] Mayr, N. A., Yuh, W. T., Arnholt, J. C., Ehrhardt, J. C., Sorosky, J. I., Magnotta, V. A., Berbaum, K. S., Zhen, W., Paulino, A. C., Oberley, L. W., Sood, A. K., and Buatti, J. M. (2000). Pixel analysis of MR perfusion imaging in predicting radiation therapy outcome in cervical cancer. *J Magnetic Resonance Imaging*, 12(6):1027–1033.

[McGrath et al., 2009] McGrath, D. M., Bradley, D. P., Tessier, J. L., Lacey, T., Taylor, C. J., and Parker, G. J. M. (2009). Comparison of model-based arterial input

functions for dynamic contrast-enhanced MRI in tumor bearing rats. *Magnetic Resonance in Medicine*, 61(5):1173–84.

[McPhee, 2012] McPhee, K. C. (2012). Delayed Bolus Arrival Time with High Molecular Weight Contrast Agent, an Indicator of Necrosis. *Annual Proceedings of International Society of Magnetic Resonance in Medicine*, 20.

[Mills et al., 2006] Mills, S. J., Patankar, T. A., Haroon, H. A., Baleriaux, D., Swindell, R., and Jackson, A. (2006). Do cerebral blood volume and contrast transfer coefficient predict prognosis in human glioma? *AJNR American Journal of Neuroradiology*, 27(4):853–858.

[Mussurakis et al.,] Mussurakis, S., Buckley, D. L., and Horsman, A. Dynamic MRI of invasive breast cancer: assessment of three region-of-interest analysis methods. *Journal of computer assisted tomography*, 21(3):431–8.

[Naish et al., 2009] Naish, J. H., Kershaw, L. E., Buckley, D. L., Jackson, A., Waterton, J. C., and Parker, G. J. M. (2009). Modeling of contrast agent kinetics in the lung using T1-weighted dynamic contrast-enhanced MRI. *Magnetic Resonance in Medicine*, 61(6):1507–14.

[Orlowski et al., 2012] Orlowski, K., Rohrer Bley, C., Zimmermann, M., Vuong, V., Hug, D., Soltermann, A., Broggini-Tenzer, A., and Pruschy, M. (2012). Dynamics of tumor hypoxia in response to patupilone and ionizing radiation. *PloS one*, 7(12):e51476.

[Orton et al., 2007] Orton, M. R., Collins, D. J., Walker-Samuel, S., D'Arcy, J. a., Hawkes, D. J., Atkinson, D., and Leach, M. O. (2007). Bayesian estimation of pharmacokinetic parameters for DCE-MRI with a robust treatment of enhancement onset time. *Physics in medicine and biology*, 52(9):2393–408.

[Orton et al., 2008] Orton, M. R., D'Arcy, J. A., Walker-Samuel, S., Hawkes, D. J., Atkinson, D., Collins, D. J., and Leach, M. O. (2008). Computationally efficient vascular input function models for quantitative kinetic modelling using DCE-MRI. *Physics in Medicine and Biology*, 53(5):1225–1239.

[Padhani, 2002] Padhani, A. R. (2002). Dynamic contrast-enhanced MRI in clinical oncology: Current status and future directions. *Journal of Magnetic Resonance Imaging*, 16(4):407–422.

[Padhani et al., 2006] Padhani, A. R., Hayes, C., Assersohn, L., Powles, T., Makris, A., Suckling, J., Leach, M. O., and Husband, J. E. (2006). Prediction of clinico-pathologic response of breast cancer to primary chemotherapy at contrast-enhanced MR imaging: initial clinical results. *Radiology*, 239(2):361–374.

[Park, 2003] Park, J. O. (2003). Measuring Response in Solid Tumors: Comparison of RECIST and WHO Response Criteria. *Japanese Journal of Clinical Oncology*, 33(10):533–537.

[Parker et al., 2006] Parker, G. J. M., Roberts, C., Macdonald, A., Buonaccorsi, G. A., Cheung, S., Buckley, D. L., Jackson, A., Watson, Y., Davies, K., and Jayson, G. C. (2006). Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI. *Magnetic Resonance in Medicine*, 56(5):993–1000.

[Patankar et al., 2005] Patankar, T. F., Haroon, H. A., Mills, S. J., Baleriaux, D., Buckley, D. L., Parker, G. J., and Jackson, A. (2005). Is volume transfer coefficient (K(trans)) related to histologic grade in human gliomas? *American Journal of Neuroradiology*, 26(10):2455–2465.

[Pluim et al., 2000] Pluim, J. P., Maintz, J. B., and Viergever, M. A. (2000). Image registration by maximization of combined mutual information and gradient information. *IEEE Transactions on Medical Imaging*, 19(8):809–814.

[Pluim et al., 2003] Pluim, J. P. W., Maintz, J. B. A., and Viergever, M. A. (2003). Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004.

[Rajwade et al., 2006] Rajwade, A., Banerjee, A., and Rangarajan, A. (2006). A New Method of Probability Density Estimation with Application to Mutual Information Based Image Registration. *Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2(1640968):1769–1776.

[Reinmuth et al., 2003] Reinmuth, N., Parikh, A. A., Ahmad, S. A., Liu, W., Stoeltzing, O., Fan, F., Takeda, A., Akagi, M., and Ellis, L. M. (2003). Biology of angiogenesis in tumors of the gastrointestinal tract. *Microscopy research and technique*, 60(2):199–207.

[Risholm et al., 2013] Risholm, P., Janoos, F., Norton, I., Golby, A. J., and Wells, W. M. (2013). Bayesian characterization of uncertainty in intra-subject non-rigid registration. *Medical Image Analysis*, 17(5):538–55.

[Risholm et al., 2010] Risholm, P., Pieper, S., Samset, E., and Wells, W. M. (2010). Summarizing and visualizing uncertainty in non-rigid registration. *MICCAI: International Conference on Medical Image Computing and Computer-Assisted Intervention*, 13(Pt 2):554–61.

[Roche et al., 2000] Roche, A., Malandain, G., and Ayache, N. (2000). Unifying maximum likelihood approaches in medical image registration. *International Journal of Imaging Systems and Technology*, 11(1):71–80.

[Roche et al., 1999] Roche, A., Malandain, G., Ayache, N., and Prima, S. (1999). Towards a better comprehension of similarity measures used in medical image registration. *Medical Image Computing and Computer-Assisted Intervention–MICCAI'99*, pages 555–566.

[Rose et al., 2007] Rose, C. J., Mills, S., O'Connor, J. P. B., Buonaccorsi, G. A., Roberts, C., Watson, Y., Whitcher, B., Jayson, G., Jackson, A., and Parker, G. J. M. (2007). Quantifying heterogeneity in dynamic contrast-enhanced MRI parameter maps. *Med Image Comput Comput Assist Interv*, 10(Pt 2):376–384.

[Rose et al., 2009a] Rose, C. J., Mills, S. J., O'Connor, J. P. B., Buonaccorsi, G. A., Roberts, C., Watson, Y., Cheung, S., Zhao, S., Whitcher, B., Jackson, A., and Parker, G. J. M. (2009a). Quantifying spatial heterogeneity in dynamic contrast-enhanced MRI parameter maps. *Magnetic Resonance in Medicine*, 62(2):488–499.

[Rose et al., 2009b] Rose, C. J., Mills, S. J., O'Connor, J. P. B., Buonaccorsi, G. A., Roberts, C., Watson, Y., Cheung, S., Zhao, S., Whitcher, B., Jackson, A., and Parker, G. J. M. (2009b). Quantifying spatial heterogeneity in dynamic contrast-enhanced {MRI} parameter maps. *Magnetic Resonance in Medicine*, 62(2):488–499.

[Rubner et al., 2000] Rubner, Y., Tomasi, C., and Guibas, L. J. (2000). The Earth Mover's Distance as a Metric for Image Retrieval. *International Journal of Computer Vision*, 40(2):99–121.

[Rueckert et al., 1998] Rueckert, D., Hayes, C., Studholme, C., Summers, P., Leach, M., and Hawkes, D. (1998). Non-rigid registration of breast MR images using mutual information. *Medical Image Computing and Computer-Assisted Interventation—MICCAI'98*, pages 1144–1152.

[Rueckert et al., 2002] Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L. G., Leach, M. O., and Hawkes, D. J. (2002). Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721.

[Sauerbier et al., 2009] Sauerbier, S., Palmowski, M., Vogeler, M., Nagursky, H., Al-Ahmad, A., Fisch, D., Hennig, J., Schmelzeisen, R., Gutwald, R., and Fasol, U. (2009). Onset and maintenance of angiogenesis in biomaterials: in vivo assessment by dynamic contrast-enhanced MRI. *Tissue Eng Part C Methods*, 15(3):455–462.

[Schabel and Morrell, 2009] Schabel, M. C. and Morrell, G. R. (2009). Uncertainty in T(1) mapping using the variable flip angle method with two flip angles. *Physics in Medicine and Biology*, 54(1):N1–8.

[Schabel and Parker, 2008] Schabel, M. C. and Parker, D. L. (2008). Uncertainty and bias in contrast concentration measurements using spoiled gradient echo pulse sequences. *Physics in Medicine and Biology*, 53(9):2345–2373.

[Schmid et al., 2009a] Schmid, V. J., Whitcher, B., Padhani, A. R., Taylor, N. J., and Yang, G.-Z. (2009a). A Bayesian hierarchical model for the analysis of a longitudinal dynamic contrast-enhanced MRI oncology study. *Magnetic Resonance in Medicine*, 61(1):163–174.

[Schmid et al., 2009b] Schmid, V. J., Whitcher, B., Padhani, A. R., Taylor, N. J., and Yang, G.-Z. (2009b). A Bayesian hierarchical model for the analysis of a longitudinal dynamic contrast-enhanced MRI oncology study. *Magnetic Resonance in Medicine*, 61(1):163–174.

[Schnabel et al., 2003] Schnabel, J. A., Tanner, C., Castellano-Smith, A. D., Degenhard, A., Leach, M. O., Hose, D. R., Hill, D. L. G., and Hawkes, D. J. (2003).

Validation of nonrigid image registration using finite-element methods: application to breast MR images. *IEEE Transactions on Medical Imaging*, 22(2):238–247.

[Shan et al., 2011] Shan, Z. Y., Mateja, S. J., Reddick, W. E., Glass, J. O., and Shulkin, B. L. (2011). Retrospective evaluation of PET-MRI registration algorithms. *Journal of digital imaging*, 24(3):485–93.

[Shibuya, 2008] Shibuya, M. (2008). Vascular endothelial growth factor-dependent and -independent regulation of angiogenesis. *BMB reports*, 41(4):278–86.

[Siegel et al., 2013] Siegel, R., Naishadham, D., and Jemal, A. (2013). Cancer statistics, 2013. *CA: a cancer journal for clinicians*, 63(1):11–30.

[Simonsen et al., 2012] Simonsen, T. G., Gaustad, J.-V., Leinaas, M. N., and Rofstad, E. K. (2012). Vascular abnormalities associated with acute hypoxia in human melanoma xenografts. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, 105(1):72–8.

[Simpson et al., 2013] Simpson, I. J. A., Woolrich, M. W., Andersson, J. L. R., Groves, A. R., and Schnabel, J. A. (2013). Ensemble learning incorporating uncertain registration. *IEEE Transactions on Medical Imaging*, 32(4):748–56.

[Springer et al., 2010] Springer, F., Steidle, G., Martirosian, P., Claussen, C. D., and Schick, F. (2010). Effects of in-pulse transverse relaxation in 3D ultrashort echo time sequences: analytical derivation, comparison to numerical simulation and experimental application at 3T. *Journal of magnetic resonance (San Diego, Calif. : 1997)*, 206(1):88–96.

[Srikanchana et al., 2004] Srikanchana, R., Thomasson, D., Choyke, P., and Dwyer, a. (2004). A comparison of pharmacokinetic models of dynamic contrast enhanced MRI.

Proceedings. 17th IEEE Symposium on Computer-Based Medical Systems, pages 361–366.

[St Lawrence and Lee, 1998] St Lawrence, K. S. and Lee, T. Y. (1998). An adiabatic approximation to the tissue homogeneity model for water exchange in the brain: I. Theoretical derivation. *Journal of Cerebral Blood Flow and Metabolism*, 18(12):1365–77.

[Studholme et al., 1999] Studholme, C., Hill, D. L. G., Hawkes, D. J., et al. (1999). An overlap invariant entropy measure of 3D medical image alignment. *Pattern recognition*, 32(1):71–86.

[Sun and Yuan, 2006] Sun, W. and Yuan, Y.-X. (2006). *Optimization Theory and Methods: Nonlinear Programming (Google eBook)*. Springer.

[Tanner, 2010] Tanner, L. N. (2010). Functional Imaging Markers for Tumour Characterisation.

[Therasse et al., 2000] Therasse, P., Arbuck, S. G., Eisenhauer, E. A., Wanders, J., Kaplan, R. S., Rubinstein, L., Verweij, J., Van Glabbeke, M., van Oosterom, A. T., Christian, M. C., and Gwyther, S. G. (2000). New Guidelines to Evaluate the Response to Treatment in Solid Tumors. *JNCI Journal of the National Cancer Institute*, 92(3):205–216.

[Thirion, 1998] Thirion, J. P. (1998). Image matching as a diffusion process: an analogy with Maxwell's demons. *Medical Image Analysis*, 2(3):243–60.

[Tofts, 1997] Tofts, P. S. (1997). Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *Journal of Magnetic Resonance Imaging*, 7(1):91–101.

[Tofts et al., 1995a] Tofts, P. S., Berkowitz, B., and Schnall, M. D. (1995a). Quantitative analysis of dynamic Gd-DTPA enhancement in breast tumors using a permeability model. *Magnetic Resonance in Medicine*, 33(4):564–568.

[Tofts et al., 1995b] Tofts, P. S., Berkowitz, B., and Schnall, M. D. (1995b). Quantitative Analysis of Dynamic Gd-DTPA Enhancement in Breast Tumors Using a Permeability Model. *Magnetic Resonance in Medicine*, 33(4):564–568.

[Tofts et al., 1999] Tofts, P. S., Brix, G., Buckley, D. L., Evelhoch, J. L., Henderson, E., Knopp, M. V., Larsson, H. B., Lee, T. Y., Mayr, N. A., Parker, G. J., Port, R. E., Taylor, J., and Weisskoff, R. M. (1999). Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusable tracer: standardized quantities and symbols. *J Magnetic Resonance Imaging*, 10(3):223–232.

[Tofts and Kermode, 1991] Tofts, P. S. and Kermode, A. G. (1991). Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts. *Magnetic Resonance in Medicine*, 17(2):357–367.

[Trouvé, 1998] Trouvé, A. (1998). Diffeomorphisms Groups and Pattern Matching in Image Analysis. *International Journal of Computer Vision*, 28(3):213–221.

[Turnbull, 2009] Turnbull, L. W. (2009). Dynamic contrast-enhanced MRI in the diagnosis and management of breast cancer. *NMR in Biomedicine*, 22(1):28–39.

[Van den Elsen et al., 1995] Van den Elsen, P. A., Maintz, J. A., Pol, E. D., and Viergever, M. A. (1995). Automatic registration of CT and MR brain images using correlation of geometrical features. *IEEE transactions on medical imaging*, 14(2):384–396.

[Vercauteren et al., 2007] Vercauteren, T., Pennec, X., Perchant, A., and Ayache, N. (2007). Non-parametric diffeomorphic image registration with the demons algorithm.

MICCAI: International Conference on Medical Image Computing and Computer-Assisted Intervention, 10(Pt 2):319–26.

[Vercauteren et al., 2008] Vercauteren, T., Pennec, X., Perchant, A., and Ayache, N. (2008). Symmetric log-domain diffeomorphic registration: A demons-based approach. *MICCAI: Internationa Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 754–761.

[Vercauteren et al., 2009] Vercauteren, T., Pennec, X., Perchant, A., and Ayache, N. (2009). Diffeomorphic demons: efficient non-parametric image registration. *NeuroImage*, 45(1 Suppl):S61–72.

[Viola and Wells III, 1997] Viola, P. and Wells III, W. M. (1997). Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154.

[Wang et al., 2008] Wang, Y., Huang, W., Panicek, D. M., Schwartz, L. H., and Koutcher, J. A. (2008). Feasibility of using limited-population-based arterial input function for pharmacokinetic modeling of osteosarcoma dynamic contrast-enhanced MRI data. *Magnetic Resonance in Medicine*, 59(5):1183–1189.

[Weickert, 1997] Weickert, J. (1997). *A review of nonlinear diffusion filtering*, volume 1252 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg.

[Weidner, 1995] Weidner, N. (1995). Intratumor microvessel density as a prognostic factor in cancer. *The American journal of pathology*, 147(1):9–19.

[Weinberg, 2013] Weinberg, R. A. (2013). *The Biology of Cancer*. Taylor & Francis Group.

[Weinmann et al., 1984] Weinmann, H. J., Brasch, R. C., Press, W. R., and Wesbey, G. E. (1984). Characteristics of gadolinium-DTPA complex: a potential NMR contrast agent. *American journal of roentgenology*, 142(3):619.

[Wiemker et al., 2012] Wiemker, R., Klinder, T., Bergtholdt, M., Meetz, K., Carlsen, I. C., and B Ulow, T. (2012). A Radial Structure Tensor and Its Use for Shape-Encoding Medical Visualization of Tubular and Nodular Structures. *IEEE transactions on visualization and computer graphics*.

[Woods et al., 1993] Woods, R. P., Mazziotta, J. C., et al. (1993). MRI-PET registration with automated algorithm. *Journal of computer assisted tomography*, 17(4):536.

[Wynn et al., 2010] Wynn, G. R., Bhasin, N., Macklin, C. P., and George, M. L. (2010). Complete clinical response to neoadjuvant chemoradiotherapy in patients with rectal cancer: opinions of British and Irish specialists. *Colorectal Disease*, 12(4):327–333.

[Yang et al., 2004] Yang, C., Karczmar, G. S., Medved, M., and Stadler, W. M. (2004). Estimating the arterial input function using two reference tissues in dynamic contrast-enhanced MRI studies: fundamental concepts and simulations. *Magnetic Resonance in Medicine*, 52(5):1110–1117.

[Yang and Knopp, 2011] Yang, X. and Knopp, M. V. (2011). Quantifying tumor vascular heterogeneity with dynamic contrast-enhanced Magnetic Resonance Imaging: a review. *Journal of Biomedicine and Biotechnology*, 2011:732848.

[Zahra et al., 2007] Zahra, M. A., Hollingsworth, K. G., Sala, E., Lomas, D. J., and Tan, L. T. (2007). Dynamic contrast-enhanced MRI as a predictor of tumour response to radiotherapy. *The Lancet Oncology*, 8(1):63–74.

[Zhang et al., 2008] Zhang, X. M., Yu, D., Zhang, H. L., Dai, Y., Bi, D., Liu, Z., Prince, M. R., and Li, C. (2008). 3D dynamic contrast-enhanced MRI of rectal carcinoma

at 3T: correlation with microvascular density and vascular endothelial growth factor markers of tumor angiogenesis. *Journal of Magnetic Resonance Imaging : JMRI*, 27(6):1309–16.

[Zitova and Flusser, 2003] Zitova, B. and Flusser, J. (2003). Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000.

[Zollei et al., 2007] Zollei, L., Jenkinson, M., Timoner, S., and Wells, W. (2007). Marginalized MAP approach and EM optimization for pair-wise registration. *IPMI: International Conference on Information Processing in Medical Imaging*, pages 662–674.

[Zöllner et al., 2009] Zöllner, F. G., Sance, R., Rogelj, P., Ledesma-Carbayo, M. J., Rørvik, J., Santos, A., and Lundervold, A. (2009). Assessment of 3D dce-MRI of the kidneys using non-rigid image registration and segmentation of voxel time courses. *Computerized Medical Imaging and Graphics*, 33(3):171–181.