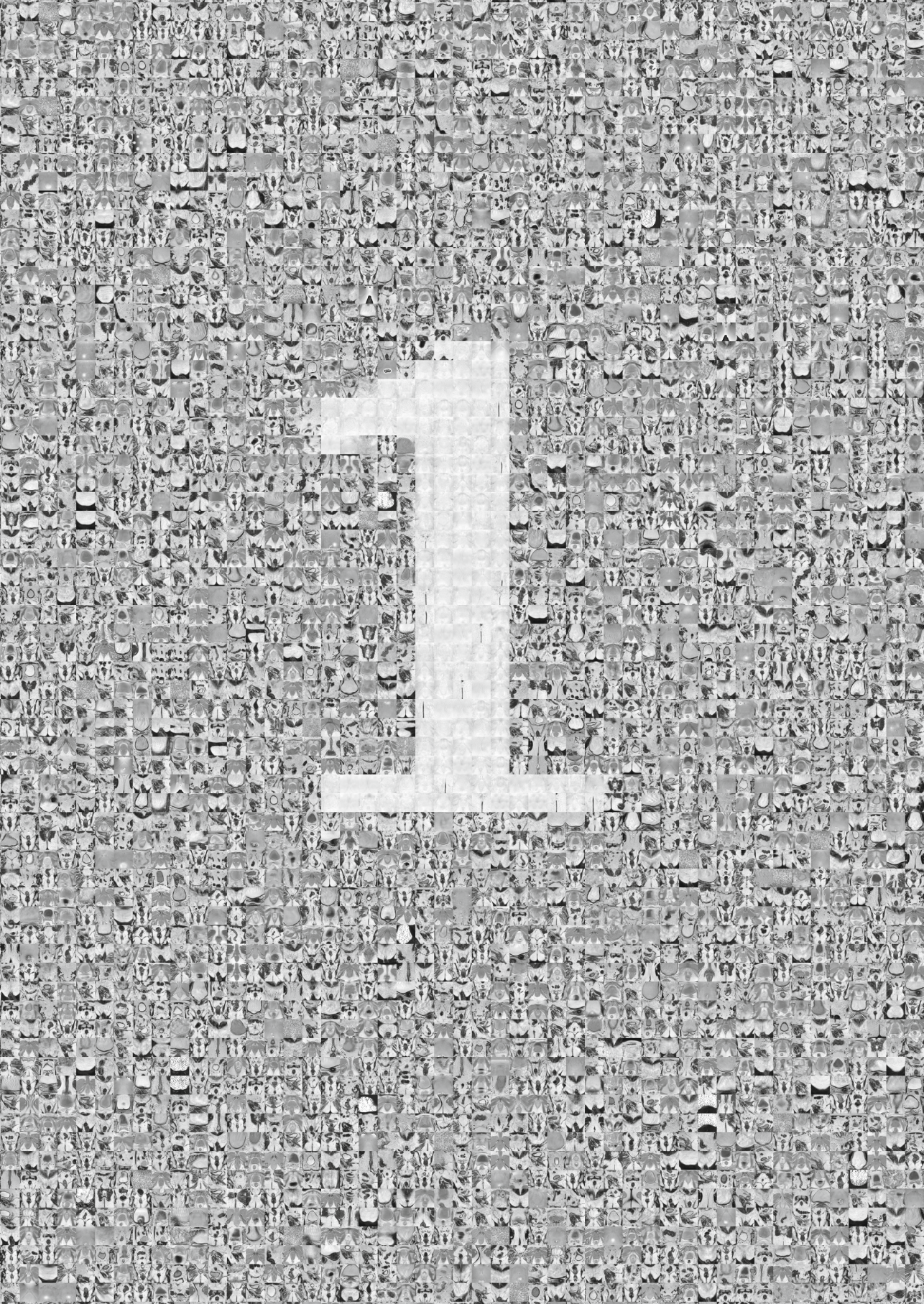
Daily quantitative MRI for radiotherapy response monitoring

Kooreman, E.S.



INTRODUCTION

MR-GUIDED RADIOTHERAPY

External beam radiotherapy (EBRT) is a type of cancer treatment that uses radiation from an outside source to treat the patient. Based on the disease type, stage, and location, a certain dose to the tumor is prescribed which is then given in small daily fractions over an extended period of time. In order to deliver a dose to the tumor while sparing the surrounding organs at risk (OARs), an accurate depiction of the tumor is necessary.

Magnetic resonance imaging (MRI) is a valuable imaging modality in oncology as images with soft tissue contrast can be generated. It is therefore often used in treatment planning for EBRT. However, treatment plans are usually made once before the start of treatment and do not take anatomical changes that happen between planning and before the start of the treatment into account. Moreover, for several tumor sites, it is known that the tumor can move significantly during treatment due to, for instance, respiratory or bowel motion. To account for these types of motion, margins around the tumor are taken into account that also are given a high radiation dose.

In 2009, a prototype machine that integrated an MRI scanner with a linear accelerator (MR-linac) was presented [1]. This machine would enable the acquisition of MR images of the patient during each treatment fraction and therefore allow adaptation of the treatment plan to the anatomy of the day. Using fast MRI sequences, it would be possible to image the motion of the tumor and reduce treatment margins.

In 2014, the first commercial MR-linac system became available [2]. This system, developed by ViewRay and called the MRIdian, combined a 0.35 T MRI system with three Cobalt-60 sources that were used for radiation. In 2017, the Cobalt-60 sources were replaced with a linac. In 2018, the second system, equipped with a 1.5 T MRI received a CE mark. This system, called the Unity MR-linac, was developed by Elekta and Philips. Our clinic acquired this system in 2016 as an early adopter, and the first patient was treated on it in 2018.

In our clinic, when a patient is eligible for treatment on the Unity MR-linac, a treatment plan is generated using images from a diagnostic MRI system with the patient positioned the same way as during treatment. On every treatment fraction, an MRI scan is performed on the MR-linac to assess the patient anatomy of the day, and when necessary the treatment plan is adapted. This can be done by either shifting the treatment plan to conform with the anatomy of the day, or by re-delineating the tumor and OARs and recalculating the treatment plan.

QUANTITATIVE MRI

MRI systems are designed to produce anatomical images where signal intensity differences between organs result in a clear depiction of the patient. Conventionally, these images are interpreted qualitatively where anomalous organ shapes or sizes, or hypo- or hyper-intense spots on an image with a certain weighting indicate the presence of some disease. In his book "Quantitative MRI of the Brain", Tofts compares this use of an MRI system to a photo camera, which makes anatomical images [3]. However, MRI systems are also capable of quantitative imaging (qMRI), transitioning from "just a camera" towards a scientific measuring device that allows quantitative measurements of the patient biology. Quantitative measures are theoretically easier to compare between institutes, systems, and patients. The challenge is to acquire quantitative parameters in a manner so that they are accurate, repeatable, and reproducible.

In general, to perform qMRI, multiple images are acquired in the same scanning session while changing the acquisition settings of the MRI scanner, generating images with different signal intensities. By fitting a model to these different signal intensities in each voxel of the image, a voxel-wise map of the underlying quantitative measure can be created.

Diffusion weighted imaging (DWI)

An example of the acquisition of such a quantitative map is given using DWI in Figure 1. For DWI, strong gradients are applied directly after the 90° excitation pulse and again after the 180° refocusing pulse of a spin-echo sequence. These gradients sensitize the MR signal to the presence of random motion, as first shown by Stejskal and Tanner [4]. The amount of diffusion weighting depends on the shape, strength and timing of these gradients and is commonly summarized as the b-value, first introduced as the 'b factor' by Le Bihan [5]. For two rectangular diffusion sensitizing gradients, the b-value is given by [6]

$$b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3). \tag{1.1}$$

Here, γ is the gyromagnetic ratio, G is the gradient amplitude, δ is the gradient duration, and Δ is the time between the start of the first- and the start of the second gradient. A larger b-value means more of the signal is suppressed due to diffusion as is shown in Figure 1a. The MR signal depends on the diffusion of the sample and the b-value following a mono-exponential decay

$$S_b = S_0 e^{-b\,ADC} \tag{1.2}$$

By fitting this equation to the signal intensity measured in a single voxel different b-values, a voxel-wise quantitative map reflecting the apparent diffusion coefficient

(ADC map) can be calculated. An example of such a fit is shown in Figure 1b for a single voxel and measurements at four different b-values. An example of a resulting ADC map can be seen in Figure 1c.

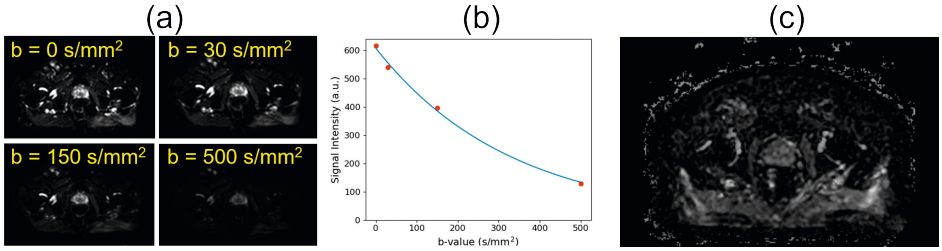


Figure 1.1. ADC mapping as an example of a quantitative technique. (a) DWIs with different b-values. The image contrast is decreased with an increasing b-value. (b) Example of the signal intensity in a single voxel at the b-values of (a), shown as red dots. The blue line shows the least squares fit of Equation 1.2. (c) The resulting ADC map.

Intravoxel Incoherent Motion (IVIM)

IVIM is an extension to conventional ADC mapping. In a homogeneous liquid, a single diffusion coefficient can be measured using DWI. In a biological environment, there are multiple sources of signal attenuation as all 'incoherent motions' result in phase dispersion. LeBihan noticed in 1988 that blood flow in the capillary bed also mimics incoherent motion and by including low b-values (typically below 100 s/mm²), a bi-exponential model can be used to separate diffusion from perfusion [7]:

$$S_b = S_0(fe^{-b\,D^*} + (1 - f)\,e^{-b\,D}), \tag{1.3}$$

where f is the perfusion fraction, D* is the pseudo-diffusion coefficient, and D is the diffusion coefficient.

Relaxometry

The measurement of relaxation times such as T_1 and T_2 using MRI is called relaxometry. These relaxation times depend on the tissue, the magnetic field strength, and temperature. In 1946, Bloch developed equations that describe the relaxation behavior of the net magnetization using classical mechanics [8]. From these "Bloch equations", the relaxation time parameters T_1 and T_2 can be derived. For the magnetization in the z-direction (along the external magnetic field), it can be shown that after a 90° radio frequency (RF) pulse, the z-component of the magnetization relaxes towards its equilibrium value following

$$M_z(t) = M_0(1 - e^{-t/T_1}), \tag{1.4}$$

where M_0 is the net magnetization at equilibrium and T_1 is the longitudinal (or spin-lattice) relaxation time constant. The gold standard measurement for the T_1 relaxation time is called inversion recovery. With this method, a 180° RF pulse is applied first after which the magnetization relaxes toward equilibrium along the z-axis following Equation 1.4. After waiting for a certain amount of time, called the inversion time (TI), the acquisition of the signal is started. The contrast of this image depends on the TI. By acquiring multiple images with multiple TIs, Equation 1.4 can be fitted to these images to determine the T_1 on a voxel-by-voxel basis.

For the magnetization in the xy- (transverse) plane, it can be shown that after excitation, the net magnetization follows

$$M_{xy}(t) = M_0 e^{-t/T_2}, \tag{1.5}$$

where T_2 is the transverse (or spin-spin) relaxation time constant. For the measurement of T_2, multiple spin-echo images are acquired with a different echo time (TE). The contrast of these images depend on the echo-time according to Equation 1.5 and by fitting this equation to the images on a voxel-by-voxel basis, a T_2 map can be calculated.

In addition to the T_1 and T_2 relaxation times, other relaxation parameters such T_1 relaxation in the rotating frame ($T_{1\rho}$) can be measured in a similar fashion [9]. For $T_{1\rho}$, a magnetization preparation pulse is placed before a conventional acquisition. This causes the magnetization to depend on a time constant called the spin-lock time (TSL) following

$$M_z(t) = M_0 e^{-TSL/T_{1\rho}}. \tag{1.6}$$

As with T_1 and T_2 measurements, multiple scans are acquired with a different TSL, and Equation 1.6 is fitted to the images on a voxel-by-voxel basis to generate a $T_{1\rho}$ map.

Dynamic contrast-enhanced MRI

Dynamic contrast-enhanced (DCE) MRI leverages the use of contrast agent to measure perfusion. A gadolinium-based contrast agent is injected using a power injector while a time series of T_1-weighted images of the target site are acquired. Depending on perfusion of the tissue, and the permeability of the capillary bed, the contrast agent is taken up at a certain rate. The heavy metal inside the contrast agent significantly reduces the T_1 of neighboring tissue which causes hyperintense signal on T_1-weighted scans. To extract quantitative parameters related to perfusion, a model is fit on a voxel-by-voxel basis to the dynamic signal intensity. In this thesis, the Tofts model was used [10]:

$$C_t(t) = K^{trans} \int_0^t C_p(\tau) e^{-k_{ep}(t-\tau)} d\tau \tag{1.7}$$

here, C_t is the concentration of contrast agent in the tissue, K^{trans} is the transfer constant of contrast agent moving from the capillaries to the extracellular extravascular space (EES) and k_{ep} is the transfer constant in the opposite direction.

QUANTITATIVE IMAGING BIOMARKERS

Quantitative imaging biomarkers (QIBs) are quantitative measures derived from a medical image that are an indicator of normal physiology, pathology, or response to treatment [11]. This could be metrics like tumor volume, but also qMRI biomarkers such as the ADC. To be of clinical value, these quantitative values, or changes therein, should be linked to biological processes. As such, they can inform about the presence of a disease, or the progress of a treatment in a quantitative way.

The imaging biomarker roadmap

The journey of an imaging biomarker from inception to clinical application is detailed in a consensus statement by Cancer Research UK and the European Organization for Research and Treatment of Cancer [12]. In this statement, two translational gaps are defined which an imaging biomarker must cross to become clinical practice. To cross these gaps, certain checkpoints need to be reached in three parallel tracks: a technical validation track, a biological/clinical validation track and a cost effectiveness track. In the technical track, the biomarker is assessed in terms of accuracy, repeatability, and reproducibility. In the biological/clinical track, a relationship with an intervention is established. The cost effectiveness regards the cost of a biomarker, which is an important factor in widespread clinical adoption. To cross the first translational gap, an imaging biomarker needs to be assessed pre-clinically or in phantoms in terms of accuracy and repeatability, and some promise of biological linking or clinical relevance needs to be established in small patient cohorts. After crossing the first gap, the imaging biomarker should be assessed in a multi-center setting in large patient cohorts. If the studies confirm the clinical relation of the imaging biomarker to either disease screening, clinical diagnosis, or treatment outcome, the imaging biomarker can cross the second gap to be adopted in clinical practice.

QIB studies are usually performed with small sample sizes [13], making it difficult to produce sufficient evidence that could move QIBs forward along the roadmap. Moreover, differences between systems, sequences, image analysis and other processing software may influence QIBs and complicate comparison of studies [14]. Organizations such as the Quantitative Imaging Biomarker Alliance (QIBA), the Quantitative Imaging Network (QIN), the American Association of Physicists in Medicine (AAPM), and the European

Imaging Biomarker Alliance (EIBALL) have been formed to standardize the acquisition of QIBs among centers and move QIB research forward.

QIBs ON MR-LINACS

When treated with an MR-linac, patients are positioned inside an MRI scanner on each treatment fraction. This opens an opportunity for daily acquisition of qMRI that can potentially be used for treatment response monitoring, without the logistical challenges that would be required to acquire on a diagnostic system. By providing a common platform, the Unity MR-linac reduces much of the sources of variation commonly found in multi-center studies, such as differences between vendors, machines, and acquisition settings.

For the Unity MR-linac, the MRI component of the system is based on a diagnostic Philips Ingenia 1.5 T MRI. Around it, the linear accelerator responsible for generating the radiation used for treatment, is positioned on a large gantry. To minimize the attenuation of the treatment beam by the MRI scanner, it had to be modified. These adaptations include a split gradient coil system, providing a 22 cm gap at the isocenter for the treatment beam to pass through and an adapted radiolucent receive coil array [1,15,16]. This has implications on the ability of the system to perform qMRI measurements. The adapted receive coil array provides a lower signal-to-noise ratio (SNR) when compared to diagnostic receive coils. Also, the maximum gradient strength and slew rate have been limited, which especially influences the acquisition of DWI. Therefore, the capabilities of the system in terms of qMRI should be assessed first.

For the MRIdian MR-linac, some of these capabilities have been assessed. In 2016, Yang *et al.* showed the feasibility of the acquisition of longitudinal DWI measurements in several patients, showing changes in tumor ADC values while values in control tissue remained stable [17]. A year later, the same group presented a sequence to acquire distortion free DWI on the system. This study included phantom measurements using a standardized phantom, showing that phantom ADC values were accurate and repeatable [18].

Accuracy, repeatability, and reproducibility

Common measures that are used to describe the performance of a QIB are the accuracy, repeatability, and reproducibility. The accuracy is defined as the difference between the measurement and the true value. Usually, this can only be assessed using calibrated phantoms, with known values that are determined by well-established (gold standard) techniques. Repeatability, also called precision, is a measure for the variation between repeated measurements under identical experimental conditions.

A common way to express the repeatability is with the repeatability coefficient (RC) or the percent repeatability coefficient (%RC), which are based on the within-subject standard deviation and the within-subject coefficient of variation, respectively [19]. The repeatability depends on the time between the two measurements. When measuring during the same scanning session, the RC represents the intra-session repeatability. When time between measurements is increased, the measured RC will probably increase because biological and scanner changes can add to the variance of the measurements. The RC is important in QIB research especially for treatment response monitoring as it provides limits above or below which a biological change has happened with 95% confidence. To determine the RC in patients, they should be measured under identical conditions, before the start of treatment [20]. This way, a baseline variance of the measurement can be established. Reproducibility describes the agreement between measurements when performed in different institutes under different conditions. This includes measuring the qMRI biomarker on different MRI systems from different vendors and with different field strengths. Ideally, QIBs must have a good reproducibility in order to compare results between institutes, allowing the aggregation of data and move towards widespread adoption.

Together with the first institutes that acquired a Unity MR-linac, the manufacturer Elekta has started the MR-linac consortium to coordinate joint efforts in research aimed at an evidence based clinical introduction of the machine [21]. As a part of this, the MOMENTUM study has been set up [22]. For centers participating in the MOMENTUM study, all patients that are scanned on any MR-linac are eligible for inclusion. Patient data, including (quantitative) imaging data are uploaded to a large registry to enable systematic research about the clinical performance of the system. Follow-up of patients is performed up to two years. For QIB research, this registry is an opportunity to store and share large amounts of multi-center patient data allowing the analysis of large numbers of patients and as clinical follow-up data is collected as well, would allow linking changes in QIB to clinical outcome. Realizing that the Unity MR-linac is an excellent platform to acquire QIBs daily in a large group of patients, the imaging biomarker brainstorm group was called to life within the MR-linac consortium to stimulate joint research about this topic. Chapter 3, which contains a recommendation paper about the acquisition of ADC maps on the Unity MR-linac is a result of a collaboration within this group.

THESIS OUTLINE

The work in this thesis is focused on the acquisition of daily QIBs on the novel Unity MR-linac system. These QIBs need to have sufficient accuracy, repeatability, and reproducibility in order to become clinically relevant, with the ultimate purpose of using them for radiation treatment response monitoring.

As a first and necessary step, we assessed the accuracy and feasibility of qMRI on the Unity MR-linac with a series of phantom measurements, which is described in Chapter 2. The results of our measurements are encouraging, showing accuracy and repeatability in phantoms to be comparable to published literature on diagnostic systems. Additionally, we acquired quantitative images in a single prostate cancer patient showing feasibility of qMRI on the system.

In the following chapter, we describe a study about the capabilities of the Unity MR-linac with respect to the acquisition of diffusion weighted images (DWI). We provide a recommendation for the acquisition of ADC maps on the system. This recommendation is intended to serve as a starting point for the community to develop acquisition protocols that produce accurate, repeatable, and reproducible ADC values on the Unity MR-linac.

The ability to measure perfusion during radiotherapy is valuable, as radiotherapy is less effective in low-perfused, hypoxic tissue while hypoxic tumors are typically more aggressive. Usually, perfusion measurements on MRI use a Gadolinium based contrast agent as described in the DCE paragraph above. Because of the use of contrast agent that needs to be injected, it is undesirable to acquire this daily in a treatment response monitoring setting on the MR-linac. Alternatives, such as IVIM, that potentially measure perfusion without the use of contrast agent are therefore of interest [23]. The subject of Chapter 4 is a study in which we acquired daily IVIM scans of 40 prostate cancer patients in a multi-center setting. The goal was to see if changes in IVIM parameters were measurable already during radiotherapy, to determine if they hold potential as a candidate for treatment response monitoring.

Multiple studies have tried to relate IVIM parameters to DCE and dynamic susceptibility contrast (DSC) MRI [23]. They report varying levels of correlations, with the most consistent one between the DSC blood volume parameter and the IVIM perfusion fraction parameter f. Results from correlations between DCE and IVIM are inconsistent [23]. However, most of these studies assess correlations in tumors at a single time point. For treatment response monitoring, it is interesting to investigate the longitudinal correlations between IVIM and DCE parameters, measured during treatment. If significant and high correlations are found, this could mean that IVIM would be a

potential substitute for DCE in terms of treatment response monitoring, avoiding the need for contrast agent. To investigate this, Chapter 5 describes a study in which 20 prostate cancer patients were imaged weekly during treatment with a DCE and IVIM scan and longitudinal correlations between DCE and IVIM parameters were assessed in different ROIs.

$T_{1\rho}$ measurements are commonly used in the field of musculoskeletal MRI for assessing cartilage [9]. They have also been shown to be valuable for the detection of fibrosis and cirrhosis in the liver [24,25]. As a qMRI method, it is interesting to investigate in the context of treatment response monitoring. $T_{1\rho}$ has been investigated for this purpose previously, although pre-clinically and for the purpose of measuring the response of tumors to hyperthermia treatment [26]. In Chapter 6, an investigation into the use of $T_{1\rho}$ as a QIB for treatment response monitoring in rectal cancer patients is described. A custom phantom was developed to characterize the accuracy of the implemented $T_{1\rho}$ sequence on the Unity MR-linac. Additionally, the $T_{1\rho}$ relaxation time was measured during radiation treatment of 10 rectal cancer patients.