



Universiteit  
Leiden  
The Netherlands

## Daily quantitative MRI for radiotherapy response monitoring

Kooreman, E.S.

### Citation

Kooreman, E. S. (2023, June 13). *Daily quantitative MRI for radiotherapy response monitoring*. Retrieved from <https://hdl.handle.net/1887/3620208>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3620208>

**Note:** To cite this publication please use the final published version (if applicable).

# **Daily Quantitative MRI for Radiotherapy Response Monitoring**

**ERNST SEBASTIAAN KOOREMAN**

## **Daily Quantitative MRI for Radiotherapy Response Monitoring**

Copyright 2023 © Ernst Sebastiaan Kooreman

All rights reserved. No part of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without permission of the author.

**ISBN** 978-94-6483-146-7

**Printing:** Ridderprint, [www.ridderprint.nl](http://www.ridderprint.nl)

**Layout and design:** Harma Makken, [persoonlijkproefschrift.nl](http://persoonlijkproefschrift.nl)

The work described in this thesis was performed at The Netherlands Cancer Institute, and was funded by ITEA3 project 16016 “STARLIT”.

# Daily Quantitative MRI for Radiotherapy Response Monitoring

Proefschrift

ter verkrijging van  
de graad van doctor aan de Universiteit Leiden,  
op gezag van rector magnificus prof.dr.ir. H. Bijl,  
volgens besluit van het college voor promoties  
te verdedigen op dinsdag 13 juni 2023  
klokke 16.15 uur  
door

Ernst Sebastiaan Kooreman  
geboren te Zeist  
in 1990

Promotor

prof. dr. U.A. van der Heide

Co-promotor

dr. P.J. van Houdt

The Netherlands Cancer Institute

Promotiecommissie

prof. dr. A.G. Webb

prof. dr. ir. M.J.P. van Osch

prof. dr. D. Thorwarth

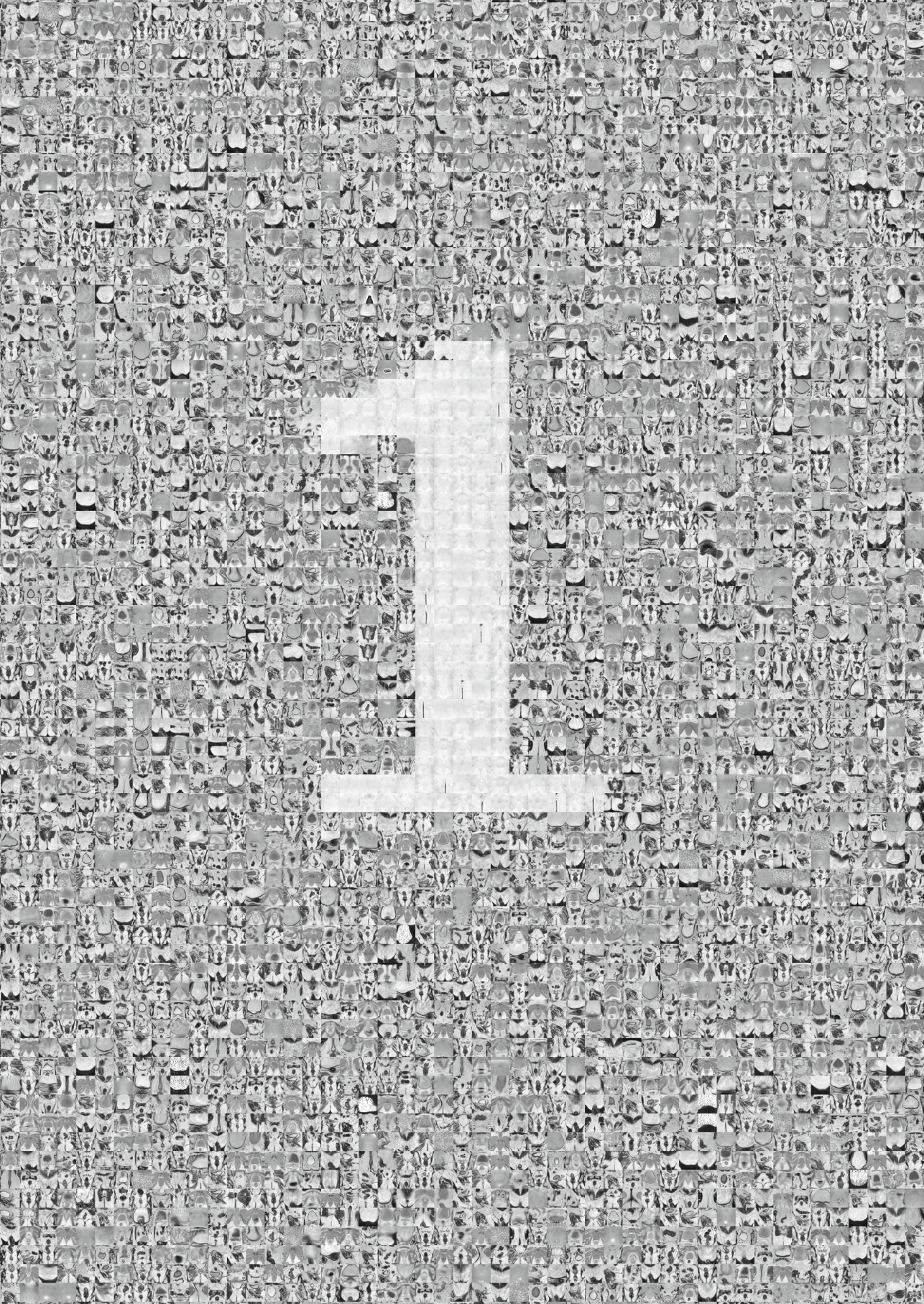
prof. dr. A. Bel

University of Tübingen

University of Amsterdam

## CONTENTS

1	Introduction	6
2	Feasibility and accuracy of quantitative imaging on a 1.5 T MR-linear accelerator	18
3	ADC measurements on the Unity MR-linac – A recommendation on behalf of the Elekta Unity MR-linac consortium	38
4	Daily intravoxel incoherent motion (IVIM) in prostate cancer patients during MR-guided radiotherapy – a multicenter study	58
5	Longitudinal Correlations Between Intravoxel Incoherent Motion (IVIM) and Dynamic Contrast-Enhanced (DCE) MRI During Radiotherapy in Prostate Cancer Patients	72
6	$T_{1\rho}$ for Radiotherapy Treatment Response Monitoring in Rectal Cancer Patients: A Pilot Study	90
7	Discussion	104
	Appendices	114
	References	116
	Summary	130
	Samenvatting	133
	List of Publications	136
	Acknowledgements	139
	Curriculum Vitae	140



# 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  $90^\circ$  excitation pulse and again after the  $180^\circ$  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]

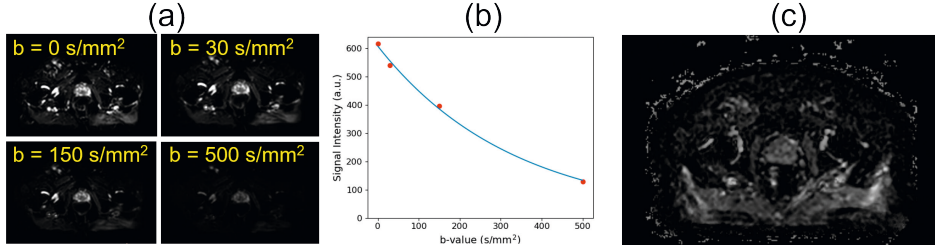
$$b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3). \quad (1.1)$$

Here,  $\gamma$  is the gyromagnetic ratio,  $G$  is the gradient amplitude,  $\delta$  is the gradient duration, and  $\Delta$  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

$$S_b = S_0 e^{-b \text{ADC}} \quad (1.2)$$

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]:

$$S_b = S_0 (f e^{-bD^*} + (1-f) e^{-bD}), \quad (1.3)$$

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

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

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  $180^\circ$  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

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

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  $T_{1\rho}$  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

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

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]:

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

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





# FEASIBILITY AND ACCURACY OF QUANTITATIVE IMAGING ON A 1.5 T MR-LINEAR ACCELERATOR

Ernst S. Kooreman, Petra J. van Houdt, Marlies E. Nowee, Vivian W.J. van Pelt, Rob H.N. Tijssen, Eric S. Paulson, Oliver J. Gurney-Champion, Jihong Wang, Folkert Koetsveld, Laurens D. van Buuren, Leon C. ter Beek, Uulke A. van der Heide

*Radiotherapy & Oncology 133, 156–162 (2019)*

## ABSTRACT

*Purpose:* Systems for magnetic resonance (MR-) guided radiotherapy enable daily MR imaging of cancer patients during treatment, which is of interest for treatment response monitoring and biomarker discovery using quantitative MRI (qMRI). Here, the performance of a 1.5 T MR-linac regarding qMRI was assessed on phantoms. Additionally, we show the feasibility of qMRI in a prostate cancer patient on this system for the first time.

*Materials and methods:* Four 1.5 T MR-linac systems from four institutes were included in this study.  $T_1$  and  $T_2$  relaxation times, and apparent diffusion coefficient (ADC) maps, as well as dynamic contrast enhanced (DCE) images were acquired. Bland-Altman statistics were used, and accuracy, repeatability, and reproducibility were determined.

*Results:* Median accuracy for  $T_1$  ranged over the four systems from 2.7 to 14.3 %, for  $T_2$  from 10.4 to 14.1 %, and for ADC from 1.9 to 2.7 %. For DCE images, the accuracy ranged from 12.8 to 35.8 % for a gadolinium concentration of 0.5 mM and deteriorated for higher concentrations. Median short-term repeatability for  $T_1$  ranged from 0.6 to 5.1 %, for  $T_2$  from 0.4 to 1.2 %, and for ADC from 1.3 to 2.2 %. DCE acquisitions showed a coefficient of variation of 0.1 to 0.6 % in the signal intensity. Long-term repeatability was 1.8 % for  $T_1$ , 1.4 % for  $T_2$ , 1.7 % for ADC, and 17.9 % for DCE. Reproducibility was 11.2 % for  $T_1$ , 2.9 % for  $T_2$ , 2.2 % for ADC, and 18.4 % for DCE.

*Conclusion:* These results indicate that qMRI on the Unity MR-linac is feasible, accurate, and repeatable which is promising for treatment response monitoring and treatment plan adaptation based on daily qMRI.

## INTRODUCTION

Biomarkers derived from quantitative Magnetic Resonance Imaging (qMRI) are promising for oncology, where they provide functional information for treatment response monitoring and prediction. Dynamic contrast-enhanced (DCE-) MRI, which measures perfusion and permeability, and diffusion weighted imaging (DWI) MRI have shown to provide valuable information for different types of cancer [27,28]. Moreover, these modalities are increasingly recommended in guidelines for tumor staging and treatment response monitoring [29,30].

The use of qMRI biomarkers in clinical practice is currently limited [12]. Differences between systems, sequences, image reconstruction algorithms, and data processing methods all influence the results of biomarker studies and complicate the comparison of studies from different centers [14,31]. Moreover, studies investigating the use of qMRI biomarkers as surrogate endpoints have been limited to small patient groups. Larger cohort studies relating qMRI biomarkers to clinical outcome are yet to be done.

An additional challenge for treatment response monitoring, is that imaging studies often are performed at different time points [32–35], and evidence about the optimal timing is lacking. To our knowledge, only one study managed to accomplish daily MRI of patients during radiation treatment of brain metastases [36]. Such studies are challenging to perform, both logistically and in terms of patient burden, but are necessary to acquire more information about MRI related changes due to treatment response.

The recently introduced linear accelerators that are integrated with an MRI scanner (MR-linac), create the possibility of daily imaging during treatment with limited increase of patient burden [37]. Treatments commonly take place over the course of weeks, where patients are daily positioned inside the MR-linac. This makes frequent MR imaging for treatment response monitoring and biomarker studies feasible. However, the design of the MRI scanner of an MR-linac is different from conventional diagnostic systems [1,2,16,38,39], which may influence the quality of the qMRI data. Therefore, before the start of treatment response monitoring and biomarker studies on an MR-linac, a thorough performance assessment is needed. Feasibility of qMRI has been shown previously on a hybrid MR-radiation therapy system that uses a combination of Cobalt-60 sources and a 0.35 T MRI, both qualitatively [40] as quantitatively using DWI [17,18]. These kinds of assessments are important in order to evaluate newly introduced MR-linac systems.

As a first step, this study aims to determine the accuracy, repeatability, and reproducibility of  $T_1$  mapping,  $T_2$  mapping, DWI, and DCE-MRI in phantoms on the Unity MR-linac (Elekta AB, Stockholm, Sweden), which is equipped with a modified Philips 1.5

T MRI system (Philips Healthcare, Best, The Netherlands) [1]. We also show the feasibility of qMRI *in vivo* by acquiring quantitative maps in a prostate cancer patient.

## METHODS AND MATERIALS

### Study Setup

Four qMRI acquisitions were performed:  $T_1$  mapping,  $T_2$  mapping, DWI for apparent diffusion coefficient (ADC) mapping, and DCE-MRI. We used the recommendations established by the quantitative imaging biomarker alliance (QIBA) DWI- and DCE-MRI profiles [41,42] for assessment of the performance of DWI and DCE-MRI data. For  $T_1$  mapping and  $T_2$  mapping we applied similar approaches as explained in detail below.

Phantom measurements were performed on four Unity MR-linac systems across four institutes. These systems are designated here as MR-linac A, B, C, and D. Each institute used their own copy of the phantoms described below, except institutes B and D, which used the same phantom for DWI.

On all systems, accuracy and short-term repeatability was assessed. The measurements were repeated within the scanning sessions for assessment of the short-term repeatability. To assess long-term repeatability, the measurements were repeated on system A after five months. Reproducibility among the four systems was determined as well. All image analysis and curve fitting were done using MATLAB (Release 2017b, MathWorks, Natick, MA).

In a prospective feasibility study, all qMRI acquisitions were performed on one patient with histologically proven prostate cancer. All protocols were approved by the medical ethics committee of The Netherlands Cancer Institute, and written informed consent was obtained.

### $T_1$ mapping

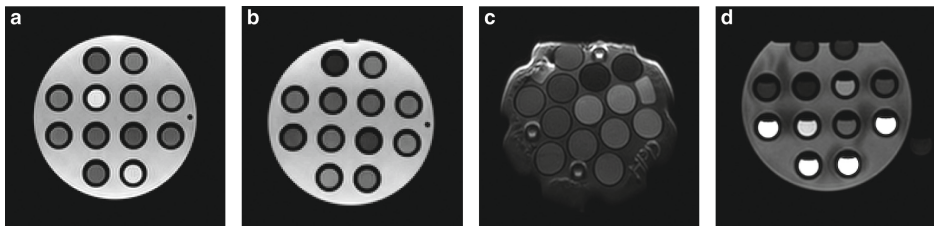
The Eurospin TO5 phantom (Diagnostic Sonar, Livingston, Scotland) was used to evaluate the performance of  $T_1$  mapping (see Figure 1a). Twelve gel samples were chosen with  $T_1$  relaxation times between 329 and 1603 ms at 296 K at 1.5T. The variable flip angle (VFA) method was applied [43], using a spoiled gradient echo sequence with flip angles of 3, 6, 10, 20, and 30°. This method was chosen because it provides a fast way to map a 3D volume and is therefore often used clinically. The remaining acquisition parameters can be found in Table S2.1.

$T_1$  values were estimated from the mean values of regions of interest (ROIs) with a diameter of 12 mm per tube by a linear least-squares method [44]. Two institutes

switched one tube with a tube of the DCE phantom for baseline  $T_1$  measurements, so this tube was omitted in the analysis for all institutes. The temperature was measured before and after each acquisition in a tube with water that was kept near the phantom during all experiments. The average temperature was used to correct estimated  $T_1$  values to the reference value at 296 K. For each tube, this reference value was provided by the phantom manufacturer at a field strength of 1.5T and a temperature of 296 K.

### $T_2$ mapping

The Eurospin TO5 phantom was used for  $T_2$  mapping as well, but with a different set of gel samples (see Figure 1b). The  $T_2$  values ranged between 49 and 212 ms at 296 K at 1.5 T. For acquisition, an accelerated multi-echo spin echo sequence was used. Acquisition parameters differed slightly for each system and can be found Table S2.2. Average decay curves were constructed for ROIs with a diameter of 12 mm within each tube. The  $T_2$  values were estimated by fitting a mono-exponential decay function with a nonlinear least-squares method. To avoid bias from stimulated echoes, the first echo was discarded for analysis which was achieved by skipping the acquisition of the first echo during scanning [45].



**Figure 2.1.** MR images of the  $T_1$ ,  $T_2$ , ADC, and DCE phantoms. (a)  $T_1$ -weighted VFA image with flip angle of 10 degrees of the Eurospin TO5. (b)  $T_2$ -weighted image with a TE of 70 ms. (c) DWI image ( $b = 0$  s/mm<sup>2</sup>) of the QIBA phantom at the iso-center. (d)  $T_1$ -weighted image of the DCE phantom.

As with the  $T_1$  mapping, the temperature was monitored before and after each experiment, and the average temperature was used to correct estimated  $T_2$  values to a reference value at 296 K.

### DWI

A diffusion phantom (High Precision Devices, Inc, Boulder, Colorado) recommended by the QIBA DWI profile was used for DWI measurements (see Figure 1c) [41]. This phantom contains 13 separate vials with aqueous solutions of 0 – 50 % w/w polyvinylpyrrolidone. The vials were surrounded by ice water to ensure measurements at 0 °C. The phantom was placed such that the central tube was in the iso-center of the system. Each institute scanned the QIBA recommended calibration protocol, which uses a spin-echo echo-



planar imaging acquisition. Diffusion weighing was achieved using Stejskal-Tanner diffusion gradients with four b-values: 0, 500, 900, and 2000 s/mm<sup>2</sup>. Other acquisition parameters can be found in Table S2.3. Additionally, a clinical protocol with larger voxels and b-values of 0, 200, and 800 was scanned on systems A and D (Table S2.7). Apparent diffusion coefficient (ADC) maps were calculated offline by fitting a linear function to the log of the signal decay versus b-values using linear least-squares. Mean values of the tubes were determined in a ROI with a diameter of 13 mm and compared to values provided by the phantom manufacturer.

## DCE

The QIBA DCE profile [42] proposes to test three aspects of the acquisition: the accuracy of  $T_1$  estimation, the stability of the signal during acquisition, and the linearity between signal intensity and concentration of the contrast agent. The first aspect was tested with the  $T_1$  mapping of the Eurospin TO5 phantom. For the latter two, a phantom was created consisting of ten tubes with different concentrations of gadolinium between 0 and 9.8 mM (Dotarem, Guerbet, France,  $T_1$  relaxivity  $3.9 \text{ s}^{-1} \text{ mM}^{-1}$ ) dissolved in a stock solution of water and 0.045 mM manganese chloride. These tubes were inserted in the Eurospin TO5 holder for image acquisition (see Figure 1d).

A spoiled gradient-echo sequence was repeated for 4:39–5:17 minutes with a temporal resolution of 4.1–4.7 s, for a total of 65–75 scans. Additional acquisition parameters can be found in Table S2.4. Both the stability of the signal over all scans, and the relation between signal intensity and concentration was assessed. The signal intensity was measured as the mean of a ROI with a diameter of 5 mm in each tube. To ensure a steady state, the first two dynamics were discarded. The median value of the remaining dynamics was used for analysis.

Although the QIBA DCE profile suggests to assess the linearity of the signal intensities over the tubes with a range of contrast agent concentrations, we converted the signal intensities to concentration values to be able to compare the results between systems [46]. For this, a baseline  $T_1$  value is needed, which was represented by the tube with 0 mM gadolinium. The  $T_1$  value of this tube was determined separately using an inversion recovery (IR) series with inversion times of 30 – 4000 ms on each MR-linac individually. Acquisition parameters of the IR series are given in Table S2.5. The IR method was used because it is regarded the gold standard [47,48], so the influence of possible inaccuracies of the clinical  $T_1$  mapping method was minimized allowing for a better assessment of the spoiled gradient-echo sequence. The calculated concentration values were compared to the known gadolinium concentrations.

## Patient data

All quantitative measurements were obtained *in vivo* in a single patient, with similar settings as described above. Details of the sequence parameters are given in Table S2.6.  $T_2$  - and ADC maps were calculated on the system. For DCE imaging, a  $T_1$  map was estimated offline based on a VFA series [46]. The Tofts model was then applied to estimate  $K^{trans}$  [49], using an arterial input function with parameters derived from an in-house study population of prostate cancer patients [50].

## Statistical Analysis

All statistical analysis was done in R (v 3.4.3). Bland-Altman statistics were used to describe the bias and limits of agreement (LoA) of the accuracy and short-term repeatability for  $T_1$ ,  $T_2$ , and ADC. Kendall's Tau (two-sided) was used to identify dependencies of the variation on the mean value, with a significance level of  $\alpha = 0.05$ . For parameters with significant dependencies, the relative percent ratio instead of the differences was used for the y-axis of the Bland-Altman plots [51,52]. For accuracy, the difference between the measured and the reference values was plotted as function of the reference value.

Additionally, the accuracy was calculated for each individual tube as the absolute percentage:

$$\text{Accuracy} = \frac{|\text{Measured} - \text{Reference}|}{\text{Reference}} \cdot 100\%, \quad (2.1)$$

with the reference values specified by the phantom manufacturers. For short-term repeatability, Bland-Altman plots were produced by plotting the difference between the first and second measurements (short-term) as a function of the mean of these two values.

Additionally, the repeatability for each individual phantom tube was calculated as follows:

$$\text{Repeatability} = \frac{|\text{Measurement 2} - \text{Measurement 1}|}{\text{Mean}(\text{Measurement 1}, \text{Measurement 2})} \cdot 100\%. \quad (2.2)$$

Short-term repeatability was calculated using the repeated measurements of a single session. Long-term repeatability was calculated using Eq. 2.2 with repeat measurements on MR-linac A with five months in between.

Reproducibility (variation across systems) was quantified for  $T_1$ ,  $T_2$ , and ADC mapping with the % coefficient of variation (%CV):

$$\%CV = \frac{SD(\text{system A,B,C,D})}{\text{Mean}(\text{system A,B,C,D})} \cdot 100\%, \quad (2.3)$$

where the standard deviation (SD) and mean of the first measurements on each separate system were used.

Additional measures were calculated to enable comparison with previous studies. For T<sub>2</sub>, the %CV was calculated for both short- and long-term repeatability, using the SD and mean of the repeat measurements. For the ADC maps, the %CV was calculated by using the SD and mean of the ROI in the tube in the iso-center of the systems. This was determined in the first acquisition of each measurement series. Additionally, diffusion images were analyzed according to the QIBA DWI profile [41].

For DCE, the stability was determined as the %CV, calculated with the ROI means and SDs over the 66–73 remaining scans from the five-minute acquisition.

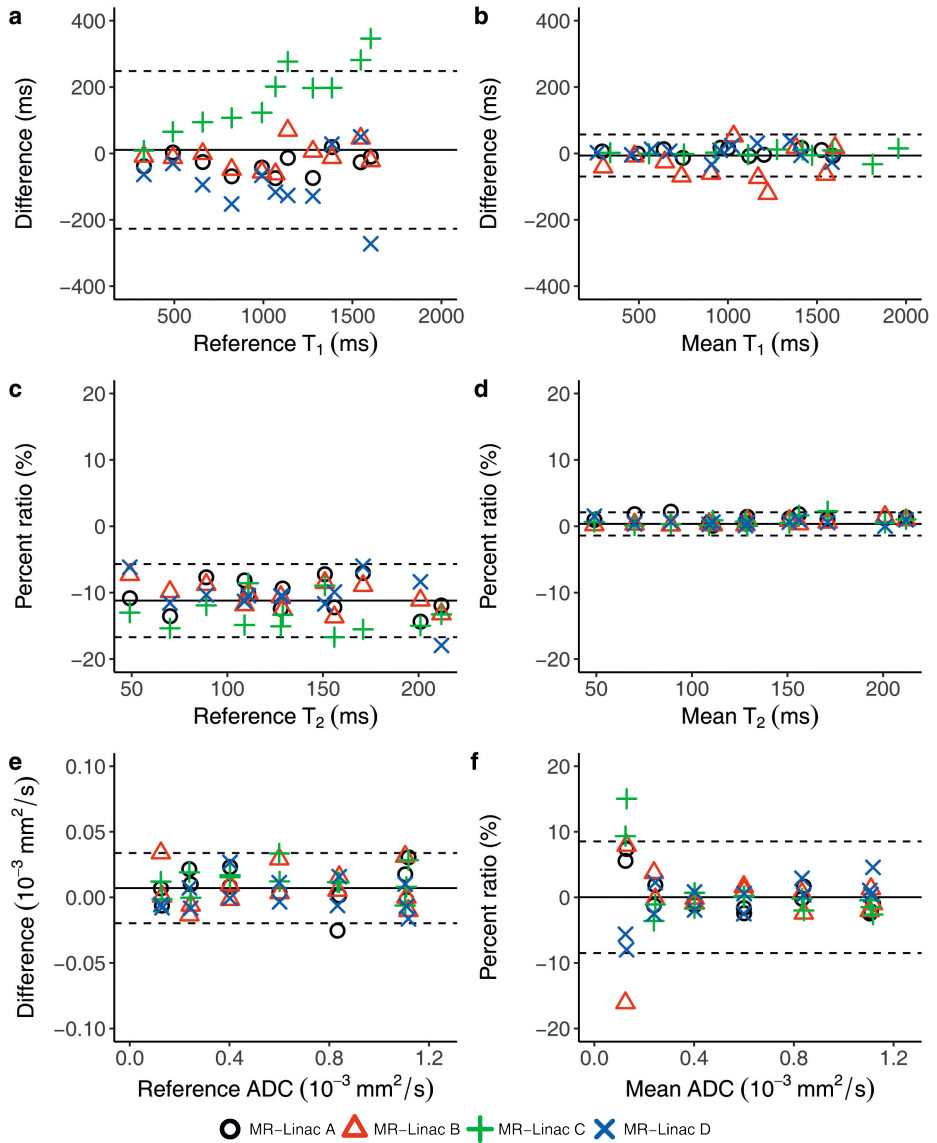
For all acquisitions, only the slice at the center of the phantom was analyzed.

## RESULTS

An overview of the accuracy and short-term repeatability is provided in Figure 2.2. The bias for the accuracy of T<sub>1</sub> (Figure 2.2a) was found to be 11 ms with LoA of ± 238 ms. For the T<sub>1</sub> short-term repeatability (Figure 2.2b), the bias was -6 ms, and the LoA were ± 63 ms. The accuracy, short-term repeatability, long-term repeatability, and reproducibility as calculated according to Eqs. 2.1 – 2.3 are presented in Table 2.1. Except for MR-linac B, the short-term repeatability was found to be lower than the accuracy.

The variation in the accuracy of T<sub>2</sub> (Figure 2.2c) showed a dependence on the T<sub>2</sub> value ( $\tau = .69, p < .001$ ). This was also found for the T<sub>2</sub> repeatability ( $\tau = .44, p < .001$ ) (Figure 2.2d). The bias and LoA, for T<sub>2</sub> were  $-11 \pm 6$  and  $0 \pm 2$  % for accuracy and short-term repeatability, respectively. The individual values for each system (Table 2.1) are comparable among systems for both accuracy and short-term repeatability, although the short-term repeatability is much lower.

The variation in ADC repeatability (Figure 2.2f) was found to depend on the measured ADC value ( $\tau = .20, p = .04$ ), so the ratio is shown. Bias and LoA for the accuracy (Figure 2.2e) were  $0.007 \times 10^{-3} \text{ mm}^2/\text{s}$ , and  $\pm 0.027 \times 10^{-3} \text{ mm}^2/\text{s}$ . For the short-term repeatability of ADC (Figure 2.2f), the bias was 0 and the LoA ± 9 %. Individual values for the systems (Table 2.1) are similar for both accuracy and short-term repeatability, which in turn are comparable to each other.



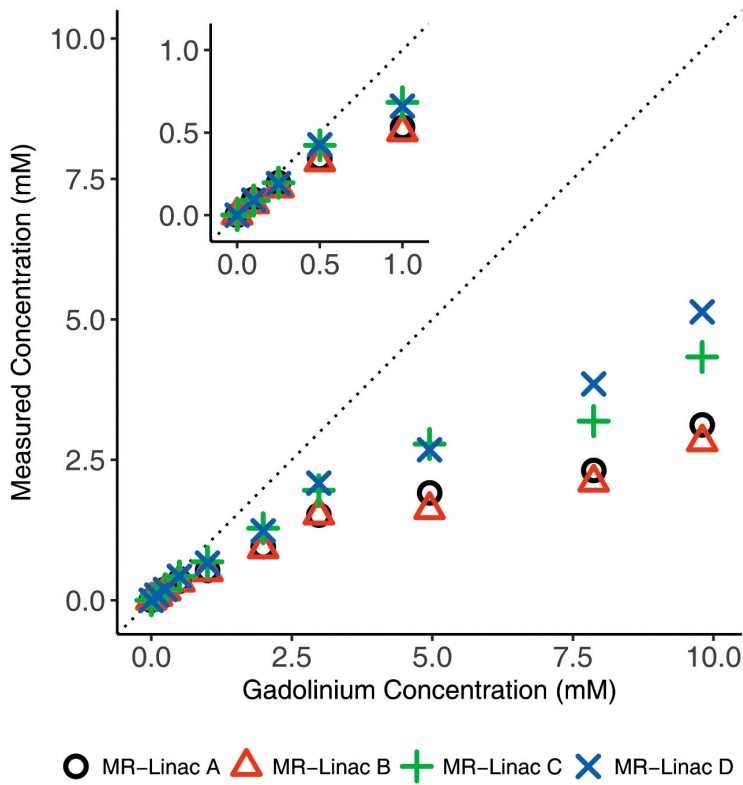
**Figure 2.2.** Bland-Altman plots showing the accuracy and repeatability of  $T_1$ ,  $T_2$ , and ADC mapping on the MR-linacs. The repeatability shown here (b,d,f) is the short-term repeatability, where two measurements were done consecutively. Each marker corresponds to a measurement in a single tube ROI. The solid line corresponds to the bias, and the dotted lines correspond to the LoA. Note that the y-axes of a, b, and e represent the qMRI units, while c, d, and f represent a percent ratio.

The accuracy in the tube at the iso-center, which is also a measure described in the QIBA DWI profile, was found to be 0.2 %, 0.0 %, 0.7 %, and 0.6 % for MR-linac A, B, C, and D, respectively. The %CV based on the ROI mean and SD in the center tube were found to be 5 % for MR-linac A, 9 % for MR-linac B, 7 % for MR-linac C, and 7 % for MR-linac D. A more complete set of the QIBA DWI profile requirement measures can be found in Table S2.8, as well as results from the clinical protocol with larger voxels.

The DCE stability measurements produced a median %CV, which represents the variation in the signal intensity over the five-minute acquisition, of 0.6 (range: 0.2 – 2.0) % over the tubes in MR-linac A. In MR-linac B the median %CV was 0.1 (range: 0.0 – 1.8) %, in MR-linac C 0.1 (range: 0.0 – 2.5) %, and in MR-linac D 0.6 (range: 0.2 – 2.9) %. Figure 2.3 shows an increasing deviation from the reference value with an increase in concentration. For concentrations of 0.5 mM, the median accuracy was 23.5 (range: 14.8 – 35.5) %. For higher concentrations, this increased to a median accuracy of 62.0 (range: 47.6 – 71.2) % for 9.8 mM. Long-term repeatability was found to be 17.9 (median, range: 1.0 – 37.9) %, and reproducibility (%CV) was 16.7 (median, range: 8.0 – 28.3) %, which is high compared to the other modalities (Table 2.1).

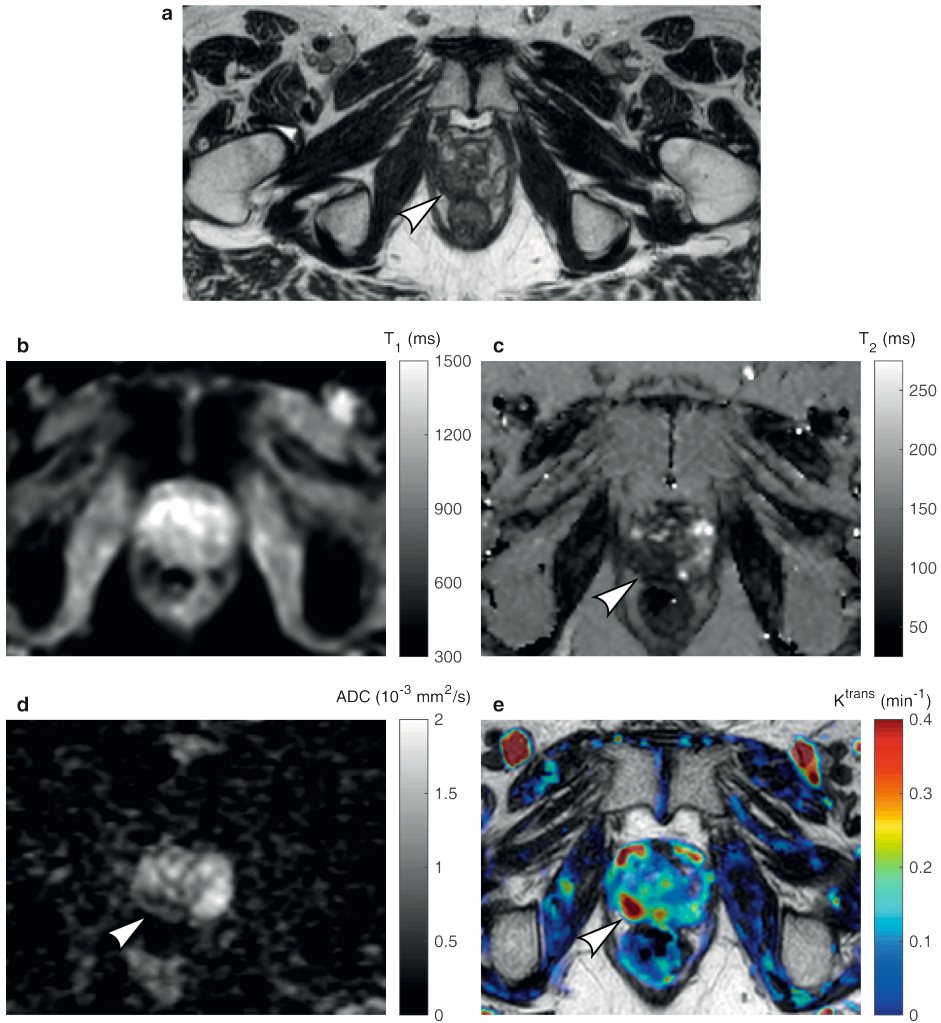
**Table 2.1.** Accuracy and repeatability of  $T_1$ ,  $T_2$ , and ADC mapping. For each system and qMRI parameter, the median (range) of the measured phantom tubes is given.

	$T_1$ (%)	$T_2$ (%)		ADC (%)
<b>Accuracy</b>				
MR-linac A	4.0 (0.6 – 11.8)	10.5 (7.0 – 14.4)		2.7 (0.2 – 9.0)
MR-linac B	2.7 (0.1 – 6.1)	10.4 (7.3 – 13.7)		1.9 (0.0 – 27.1)
MR-linac C	14.3 (2.6 – 24.4)	14.1 (8.6 – 16.7)		2.0 (0.1 – 9.6)
MR-linac D	10.9 (2.0 – 19.2)	10.5 (6.0 – 18.0)		1.9 (0.2 – 6.6)
<b>Short-term repeatability</b>			<b>%CV</b>	
MR-linac A	1.2 (0.2 – 2.1)	1.2 (0.1 – 2.2)	0.8 (0.1 – 1.6)	1.7 (0.2 – 7.1)
MR-linac B	5.1 (1.1 – 13.8)	0.4 (0.2 – 1.5)	0.3 (0.1 – 1.0)	1.3 (0.2 – 17.5)
MR-linac C	0.6 (0.1 – 1.8)	0.6 (0.1 – 2.2)	0.4 (0.0 – 1.6)	1.5 (0.1 – 14.0)
MR-linac D	1.7 (0.4 – 3.7)	0.6 (0.0 – 1.5)	0.4 (0.0 – 1.1)	2.2 (0.4 – 8.3)
<b>Long-term repeatability</b>				
MR-Linac A	1.8 (0.4 – 5.6)	1.4 (0.8 – 2.9)	1.0 (0.6 – 2.1)	1.7 (0.0 – 6.4)
<b>Reproducibility (%CV)</b>				
All systems	11.2 (6.6 – 15.8)	2.9 (0.9 – 4.7)		2.2 (0.6 – 12.0)



**Figure 2.3.** Measured gadolinium concentration plotted against the reference value. The dotted line represents the identity line ( $y = x$ ).

Figure 2.4 shows the quantitative maps from one prostate cancer patient (62 years, initial PSA 37  $\mu\text{g/L}$ , Gleason score 4+5), acquired before the start of radiation treatment. The tumor is clearly visible on the  $T_2$ , ADC, and  $K^{\text{trans}}$  maps (Figure 2.4 b-d), and the images indicate good quality with minimal distortions and no obvious artefacts.



**Figure 2.4.** Quantitative maps of a patient with prostate cancer. (a)  $T_2$ -weighted anatomical image, (b)  $T_1$  map, (c)  $T_2$  map, (d) ADC map, (e)  $K^{\text{trans}}$  map. The white arrows indicate the location of the tumor, visible on the  $T_2$  map, ADC map, and  $K^{\text{trans}}$  map.

## DISCUSSION

MR-linac systems may enable daily qMRI acquisitions for treatment response monitoring, prediction, and biomarker discovery during radiotherapy. In this study, we determined the accuracy, repeatability, and reproducibility of quantitative  $T_1$  mapping,  $T_2$  mapping, ADC mapping, and DCE-MRI on the Unity MR-linac. Additionally, we are the first to demonstrate feasibility of these quantitative acquisitions on the Unity MR-linac in a patient with prostate cancer.

The Bland-Altman plots provide an overview of the four MR-linacs, which means that the reported biases and LoAs summarize the group, and do not represent individual systems. For  $T_1$  and  $T_2$  they however show clearly that repeated measurements on a single system (Figure 2.2b,d) show less variation than measurements between systems (Figure 2.2a,c). The former is interesting for single-center studies, where one system is used to assess variations in individual patients over time, and the latter is interesting for multi-center studies.

The accuracy of the  $T_1$  VFA series (Figure 2.2a) shows great variation between systems, especially for higher  $T_1$  values. This is likely the result of using the VFA method, which is known to depend on several parameters, e.g. spoiling, and is generally known to overestimate  $T_1$  values [47,48]. The variation between systems underwrites that careful validation and optimization is needed if the VFA were to be used in a multi-center study.

For  $T_2$  accuracy, a clear negative bias was found over the entire range of  $T_2$  values, which was also found previously in studies using a multi-echo approach [53–55]. As this bias is over the entire range of  $T_2$  values, this should be of little influence to detect differences in tissues. The LoA indicate a variation of 12 %, which corresponds to 6–25 ms over the range of assessed  $T_2$  values. Therefore, differences between for instance a prostate tumor (80 ms) and healthy prostate tissue (150 ms) [56,57] should be very well detectable. The zero bias and narrow LoA of the repeatability indicate that small changes in  $T_2$  due to radiation treatment should be detectable. Short- and long-term repeatability %CV are comparable to the results of diagnostic systems [57].

For ADC mapping, the QIBA DWI profile presents threshold values for measurements in the iso-center that represent requirements for systems to meet the profile claims about confidence intervals for ADC measurements in patients [41]. For the accuracy, all systems passed the requirement of  $\leq 3.6$  %, as the highest value found was 0.7 %, indicating that ADC measurements in a ROI are accurate. This center tube accuracy also compares to previously found values in diagnostic systems [58,59]. On the other hand, none of the systems met the requirement for the %CV, as this was above the recommended 2 % for all systems. This was also found previously on a 1.5 T system [60]. One reason for the increased ROI %CV values is the reduced signal to noise ratio of the available 8-channel body array [16] compared to a head coil as recommended by the QIBA. The %CV is determined using the SD of the ROI in the central tube and is therefore closely related to the SNR. Indeed, evaluation of the clinical acquisition sequence with larger voxels and lower maximum b-value showed that the %CV requirements were met (Table S2.8). For the other tubes, the accuracy deteriorated up to 27.1 % in MR-linac B for the vial with the lowest ADC ( $0.125 \times 10^{-3}$  mm<sup>2</sup>/s). The repeatability also worsened at these low ADC values. These deviations can partly be explained by lower SNR due



to the receiver coil, but possibly also by gradient nonlinearities which influence ADC values measured away from the iso-center [58]. This influence, as well as system based geometric distortions of the Unity MR-linac should be quantified, especially if the goal for imaging is treatment planning or dose painting [61]. Short- and long-term repeatability of ADC mapping in the iso-center are comparable to previously published results [58].

For DCE, good stability was found, with a %CV in signal intensity below 2 % over all tubes over the course of five minutes. The gadolinium concentrations were determined inaccurately, where concentrations of 0.5 mM could be estimated within a range of ~30 %. For higher gadolinium concentrations, the systems all show a severe underestimation. This should not be a problem for low-perfused organs like the prostate but might result in an under-estimation for well-perfused tissues. Although the errors in repeatability and reproducibility are relatively high, the patient  $K^{\text{trans}}$  image shows the added value on the single patient level, as the tumor is clearly visible (Figure 2.4e).

In conclusion, we assessed the performance of the Unity MR-linac for a range of quantitative MR sequences and showed the feasibility of qMRI in a single patient. The accuracy and repeatability for  $T_1$  and  $T_2$  are similar to literature values from diagnostic systems. ADC mapping is also accurate although larger voxels might be advisable to increase the SNR. DCE acquisitions are decreasingly accurate for increasing contrast agent concentration but are stable and valuable for individual patients. This indicates that the Unity MR-linac performs similar to diagnostic MRI systems and can be used for treatment response monitoring and biomarker discovery studies.

## SUPPLEMENTARY MATERIAL

**Table S2.1:** T<sub>1</sub>-map sequences

System	TR (ms)	TE (ms)	Flip Angle (°)	FOV (mm <sup>3</sup> )	Parallel Imaging	Acquired Voxel Size (Reconstructed) (mm <sup>3</sup> )	Acquisition Time (1 FA) (s)
MR-Linac A	20	4	3, 6, 10, 20, 30	256 x 256 x 20	No	2 x 2 x 4 (2 x 2 x 2)	11
MR-Linac B	20	4	3, 6, 10, 20, 30	256 x 256 x 20	No	2 x 2 x 5 (2 x 2 x 2)	11
MR-Linac C	20	2.4	3, 6, 10, 20, 30	260 x 260 x 150	Yes, 2x	2 x 2 x 10 (1.8 x 1.8 x 5)	19
MR-Linac D	20	4	3, 6, 10, 20, 30	256 x 256 x 20	No	2 x 2 x 4 (2 x 2 x 2)	11

**Table S2.2:** T<sub>2</sub>-map sequences

System	TR (ms)	TE (ms)	TE (number of echoes)	Δ TE	FOV (mm <sup>3</sup> )	Parallel Imaging	Acquired Voxel Size (Reconstructed) (mm <sup>3</sup> )	Acquisition Time (mm:ss)
MR-Linac A	2693	20 – 190 (18)	10	10	256 x 256 x 2	Yes, 2x	2 x 2 x 2 (2 x 2 x 2)	03:33 <sup>a</sup>
MR-Linac B	2693	20 – 190 (18)	10	10	256 x 256 x 2	Yes, 2x	2 x 2 x 2 (2 x 2 x 2)	05:47
MR-Linac C	5000	24 – 252 (20)	12	12	256 x 256 x 2	Yes, 2x	2 x 2 x 2 (2 x 2 x 2)	05:30
MR-Linac D	3439	32 – 192 (10)	16	16	256 x 256 x 90	Yes, 2x	2.3 x 2.3 x 5 (1.6 x 1.6 x 5)	06:15

<sup>a</sup>A k-t undersampling scheme was used for acceleration (f57l)

**Table S2.3:** DWI sequences according to the QIBA profile [41].

System	TR (ms)	TE (ms)	b-values (s/mm <sup>2</sup> )	FOV (mm <sup>3</sup> )	Parallel Imaging	Half scan	Acquired Voxel Size (Reconstructed) (mm <sup>3</sup> )	Acquisition Time (mm:ss)
MR-Linac A	10000	131	0, 500, 900, 2000	220 x 220 x 124	Yes, 2x	0.622	1.7 x 1.7 x 4 (0.8 x 0.8 x 4)	02:00
MR-Linac B	10000	131	0, 500, 900, 2000	220 x 220 x 124	Yes, 2x	0.619	1.7 x 1.7 x 4 (0.8 x 0.8 x 4)	02:00
MR-Linac C	10000	103	0, 500, 900, 2000	220 x 220 x 124	Yes, 2x	0.651	1.7 x 1.8 x 4 (0.9 x 0.9 x 4)	02:00
MR-Linac D	10000	131	0, 500, 900, 2000	220 x 220 x 124	Yes, 2x	0.622	1.7 x 1.7 x 4 (0.8 x 0.8 x 4)	02:00

**Table S2.4:** DCE sequences

System	TR (ms)	TE (ms)	Flip Angle (°)	FOV (mm <sup>3</sup> )	Parallel Imaging	Dynamic scan time (s)	Number of dynamics	Acquired Voxel Size (Reconstructed) (mm <sup>3</sup> )	Acquisition Time (mm:ss)
MR-Linac A	5	1.9	30	220 x 320 x 60	Yes, 2.2x	4.1	75	2.2 x 2.2 x 6 (1.25 x 1.25 x 3)	05:05
MR-Linac B	5	2.4	20	260 x 260 x 100	Yes, 2x	4.7	65	2.5 x 2.5 x 4 (1.35 x 1.35 x 4)	05:17
MR-Linac C	4.4	2.1	25	260 x 260 x 150	Yes, 2x	4.1	69	2 x 2 x 10 (1.8 x 1.8 x 5)	04:39
MR-Linac D	5	1.9	30	220 x 320 x 60	Yes, 2.2x	4.1	75	2.2 x 2.2 x 6 (1.25 x 1.25 x 3)	05:05

**Table S2.5:** Inversion Recovery for DCE baseline

System	TR (ms)	TE (ms)	TSE Factor	Inversion Time (ms)	FOV (mm <sup>3</sup> )	Acquired Voxel Size (Reconstructed) (mm <sup>3</sup> )	Acquisition Time (mm:ss) <sup>a</sup>
MR-Linac A	8000	10	7	30, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000, 1250, 1500, 2000, 4000	256 x 256	2 x 2 x 4 (2 x 2 x 4)	02:40
MR-Linac B	8000	10	7	30, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000, 1250, 1500, 2000, 4000	256 x 256	2 x 2 x 4 (2 x 2 x 4)	02:40
MR-Linac C	5000	7.8	None	50, 75, 100, 150, 200, 300, 400, 500, 750, 1000, 1250, 1500, 2000, 4000	250 x 250	2 x 2 x 4 (1 x 1 x 4)	10:30
MR-Linac D	8000	10	7	30, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, 1000, 1250, 1500, 2000, 4000	256 x 256	2 x 2 x 4 (2 x 2 x 4)	02:40

<sup>a</sup>Acquisition times for one inversion time.

**Table S2.6:** Patient sequences

Sequence	TR (ms)	TE (ms)	Flip Angle	Parallel Imaging	Half- scan	FOV	Acquired Voxel Size (Reconstructed) (mm <sup>3</sup> )	Acquisition Time (mm:ss)
T2w (TSE) <sup>a</sup>	1300	114	90	Yes, 3.5x	0.6	400 x 448 x 250	1.2 x 1.2 x 1.2 (0.5 x 0.5 x 1.2)	06:25 (NSA = 2)
T1 (FFE)	20	4	3, 6, 10, 20, 30	Yes, 3x	No	220 x 251 x 60	2.6 x 2.7 x 7 (1.6 x 1.6 x 3.5)	00:09 (1 FA)
T2 (TSE) <sup>b</sup>	2000	20-190	90	Yes, 2x	0.613	262 x 262 x 60	1.8 x 1.9 x 3 (1.4 x 1.4 x 3)	02:20
DWI (EPI) <sup>c</sup>	3404	61	90	Yes, 2.3x	0.6	430 x 430 x 60	3 x 3 x 3 (1.5 x 1.5 x 3)	04:05
DCE (FFE) <sup>d</sup>	4	1.9	35	Yes, 2x	No	220 x 251 x 60	2.6 x 2.7 x 7 (1.6 x 1.6 x 3.5)	05:05

<sup>a</sup>TSE factor: 110, startup echoes: 4.

<sup>b</sup>Echo times ranged from 20-190 ms with steps of 10 ms for a total of 18.

<sup>c</sup>b-values (averages): 0 (1), 1 (2), 15 (3), 200 (9), 800 (9), only b = 200 and b = 800 were used for ADC reconstruction.

<sup>d</sup>Dynamic scan time = 2.8 s

We suggest in the discussion that the QIBA requirements for the ADC ROI %CV are not met because the system setup requires a head coil which is not available for the Unity MR-linac. Therefore, we acquired additional data on MR-linac A and D using acquisition parameters that increase the SNR (highest b-value 800, shorter TE, increased voxel size, see Table S2.7). The resulting measures show indeed a reduction in the ROI %CV, meeting the requirement on both systems (Table S2.8). Moreover, the requirements for short-term repeatability are also met on both systems.

**Table S2.7:** DWI sequence with higher SNR.

System	TR (ms)	TE (ms)	b-values (s/mm <sup>2</sup> )	FOV (mm <sup>2</sup> )	Parallel Imaging	Half scan	Acquired Voxel Size (Reconstructed) (mm <sup>3</sup> )	Acquisition Time (mm:ss)
MR-Linac A & D	8462	60	0, 200, 800	430 x 430 x 240	Yes, 2.3x	0.6	3 x 3 x 4.8 (1.3 x 1.3 x 4.8)	04:39

**Table S2.8:** Overview of measures calculated according to the QIBA DWI profile [41]. A bold font marks QIBA requirements that are met. Values between parentheses (MR-linac A and D) are derived from the more clinical protocol (Table S2.7).

	QIBA requirement	MR-linac A	MR-linac B <sup>a</sup>	MR-linac C	MR-linac D
ADC ROI bias	≤3.6 %	<b>0.2 %</b>	<b>0.0 %</b>	<b>0.7 %</b>	<b>0.6 %</b>
ADC ROI %CV	≤2 %	5 ( <b>1</b> ) %	9 %	7 %	7 ( <b>2</b> ) %
Short-term ADC repeatability	RC≤1.5 ×10 <sup>-5</sup> mm <sup>2</sup> /s wCV≤0.5 %	RC=3.0 ( <b>1.0</b> ) wCV=1.0 ( <b>0.3</b> )	RC=2.2 wCV=0.9	RC=3.3 wCV=1.1	RC=3.1 ( <b>0.3</b> ) wCV=1.0 ( <b>0.1</b> )
Long-term ADC repeatability	RC≤6.5 ×10 <sup>-5</sup> mm <sup>2</sup> /s wCV≤2.2 %	<b>RC=2.5</b> <b>wCV=0.8</b>			
DWI b=0 SNR	≥50±5	<b>101±11</b>	25±4	<b>52±5</b>	<b>61±7</b>
ADC b-value dependence	≤2 %	<b>0.5 %</b>	<b>0.7 %</b>	<b>0.2 %</b>	<b>0.1 %</b>

<sup>a</sup>Three instead of four intra-session measurements were available from MR-linac B.





ADC MEASUREMENTS ON  
THE UNITY MR-LINAC  
– A RECOMMENDATION  
ON BEHALF OF THE  
ELEKTA UNITY MR-LINAC  
CONSORTIUM

Ernst S. Kooreman, Petra J. van Houdt, Rick Keesman, Floris J. Pos, Vivian W. J. van Pelt, Marlies E. Nowee, Andreas Wetscherek, Rob H. N. Tijssen, Marielle E. P. Philippens, Daniela Thorwarth, Jihong Wang, Amita Shukla-Dave, William A. Hall, Eric S. Paulson, Uulke A. van der Heide

*Radiotherapy & Oncology 153, 106–113 (2020)*



## ABSTRACT

*Background and purpose:* Diffusion-weighted imaging (DWI) for treatment response monitoring is feasible on hybrid magnetic resonance linear accelerator (MR-linac) systems. The MRI scanner of the Elekta Unity system has an adjusted design compared to diagnostic scanners. We investigated its impact on measuring the DWI-derived apparent diffusion coefficient (ADC) regarding three aspects: the choice of b-values, the spatial variation of the ADC, and scanning during radiation treatment. The aim of this study is to give recommendations for accurate ADC measurements on Unity systems.

*Materials and methods:* Signal-to-noise ratio (SNR) measurements with increasing b-values were done to determine the highest b-value that can be measured reliably. The spatial variation of the ADC was assessed on six Unity systems with a cylindrical phantom of 40 cm diameter. The influence of gantry rotation and irradiation was investigated by acquiring DWI images before and during treatment of 11 prostate cancer patients.

*Results:* On the Unity system, a maximum b-value of 500 s/mm<sup>2</sup> should be used for ADC quantification, as a trade-off between SNR and diffusion weighting. Accurate ADC values were obtained within 7 cm from the iso-center, while outside this region ADC values deviated more than 5%. The ADC was not influenced by the rotating linac or irradiation during treatment.

*Conclusion:* We provide Unity system specific recommendations for measuring the ADC. This will increase the consistency of ADC values acquired in different centers on the Unity system, enabling large cohort studies for biomarker discovery and treatment response monitoring.

## INTRODUCTION

Imaging biomarkers are important in oncology as they non-invasively provide information about the tumor and can be used for treatment response monitoring [62]. Quantitative imaging biomarkers (QIBs), are of particular interest because they provide quantitative information about tissue characteristics [63]. Ideally, they facilitate the comparison across different vendors and centers. However, differences in system hardware, acquisition parameters, and image analysis techniques introduce variability of QIB values [12,14]. It is critical to understand these differences and to test and validate QIBs before they can be incorporated in clinical trials [64].

The introduction of hybrid systems, which integrate an MRI with a linear accelerator (MR-linac), presents a unique opportunity for QIB studies. MRI-guided treatments enable QIBs to be acquired daily, which is practically not feasible on diagnostic MRI systems and will provide valuable longitudinal information. However, to maximize the power of these QIB studies, it is critical to harmonize the acquisition protocols across centers.

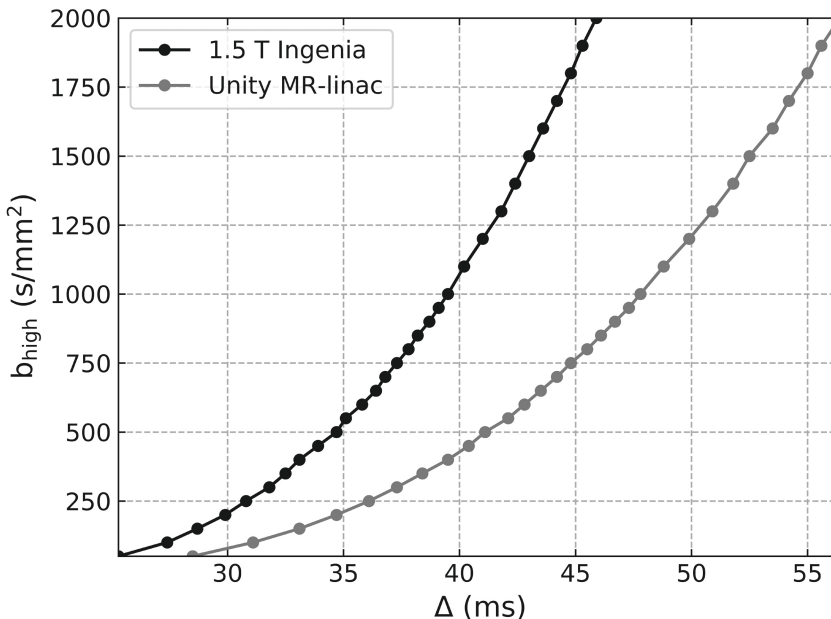
The apparent diffusion coefficient (ADC), derived from diffusion-weighted imaging (DWI), has been shown to change during treatment with radiotherapy [17,36,65–68]. Recently, efforts were made to standardize DWI acquisition. Initiatives like the quantitative imaging biomarkers alliance (QIBA), and other working groups have led to consensus recommendations for DWI in multiple tumor sites based on existing literature [69–71]. However, these recommendations are specifically designed for diagnostic MRI systems, while MR-linacs may warrant additional considerations due to their adjusted design [1,2,38,39].

The Unity MR-linac system (Elekta AB, Stockholm, Sweden), based on a Philips 1.5 T Ingenia MRI, is a hybrid system that can be used for quantitative MRI [72]. To allow accurate radiation treatment, adjustments have been made to the MRI part of the system [1], which necessitate special considerations for DWI with the goal of ADC measurements when compared to diagnostic MRI systems. We have identified three main categories for these adjustments: the choice of b-values, the spatial dependence of the ADC for the Unity system, and the influence of the rotating gantry when acquiring data during patient irradiation.

The choice of b-values is influenced by the performance of the gradient system and the available receive coil. Diffusion weighting on the Unity system is performed using the Stejskal-Tanner pulsed gradient spin echo technique, where two balanced diffusion-sensitizing gradients are placed around a 180° refocusing pulse [4]. Either echo planar imaging (EPI), or turbo spin-echo (TSE) can be used for readout. The amount

of diffusion weighting depends on the strength, slew rate, and timing of the diffusion gradients and is usually represented by a scalar called the b-value ( $\text{s}/\text{mm}^2$ ). The higher the b-value, the stronger the diffusion weighting [4].

The time between the onset of the two diffusion gradients ( $\Delta$ ), sometimes called the diffusion time, is an important parameter for DWI because a longer  $\Delta$  increases the chance that water molecules reach boundaries in a restricted (biological) environment. Therefore, in structured tissue, a different  $\Delta$  can result in different ADC values even when the same b-values are used. Normally,  $\Delta$  is kept as short as possible to minimize T2 decay and thereby maximizing the signal-to-noise ratio (SNR). Under clinical operation, the Unity system uses a maximum gradient strength of 15 mT/m and a slew rate of 65 T/m/s, which is lower compared to values typically used in diagnostic systems. Therefore, to achieve the same b-value on a Unity system, the diffusion gradients are prolonged which results in lower SNR and a longer  $\Delta$ . Figure 3.1 shows the b-value that corresponds to a particular  $\Delta$  for the Unity system and a Philips 1.5 T Ingenua scanner, both using the same acquisition settings (see Supplemental Table S3.1) and both employing their maximum gradient amplitude and slew rate. To achieve the same  $\Delta$ , b-values need to be chosen on the Unity that are about half those of diagnostic 1.5 T systems.



**Figure 3.1.** Comparison between the Unity system and a 1.5 T Philips Ingenua of the time between the start of the two diffusion gradients ( $\Delta$ ). For example, a b-value of 1200  $\text{s}/\text{mm}^2$  at the diagnostic system has a  $\Delta$  of 41 ms. For the Unity system, the b-value with the same  $\Delta$  is less than half: 500  $\text{s}/\text{mm}^2$ .

Another aspect influencing the choice of b-values is the  $2 \times 4$  channel (four anterior, four posterior) receive coil array [16]. In combination with the gradient performance, this results in a lower SNR when compared to diagnostic systems. As a result of the Rician noise distribution of magnitude images [73], a positive bias is introduced to the measured signal intensities which results in an underestimation of the ADC. This effect becomes larger when the SNR is low warranting careful consideration of the choice of b-values.

The gradient coils of the Unity system are physically split, creating a 22 cm gap around the iso-center to allow passage of the treatment beam [74]. This is different from diagnostic systems and might pose problems for images in the iso-center, especially for DWI where large diffusion sensitizing gradients are used. For instance, present day gradient systems are actively shielded to reduce the generation of eddy currents in other conductive structures of the MRI. This works well over the area of the gradient coils, but at the edges there is leakage [75]. In diagnostic systems, these edges are away from the iso-center, but for the Unity system they are close to the iso-center due to its split gradient design. This might increase eddy currents which cause unwanted variation in the magnetic field [76]. For DWI, in addition to geometric distortions, this also affects the measured ADC value spatially. Besides eddy currents, there are other sources that can cause spatial variations of ADC values, such as concomitant fields and gradient non-linearity [58,77,78].

During patient irradiation, a gantry containing the linac rotates around the MRI of the Unity system. Previous studies have shown that the influence of this gantry on the  $B_0$  field is small for both a static gantry at multiple angles [74] and a continuously rotating gantry [79]. It was shown that the inhomogeneities were  $< 500$  nT (or  $< 0.3$  ppm) but varied spatially with the gantry position. It should also be determined if DWI with an EPI readout can be performed during patient irradiation with gantry rotation.

This study is structured around the three categories (choice of b-values, spatial dependence of the ADC, and diffusion measurements during treatment) and the aim is to incorporate system specific considerations regarding these categories in a recommendation for accurate ADC measurement on the Unity system.

## MATERIALS AND METHODS

### Choice of b-values

Assuming a mono-exponential decay of the diffusion signal with increasing b-value, the ADC can be calculated using  $ADC = \ln(S_{low}/S_{high})/(b_{high} - b_{low})$ , where  $b_{low}$  and  $b_{high}$  are two b-values, with  $S_{low}$  and  $S_{high}$  as corresponding signal intensities. The highest SNR

would be obtained when  $b_{\text{low}} = 0 \text{ s/mm}^2$  is used. However, besides diffusion, perfusion also attributes to the signal attenuation of DWI, especially at lower  $b$ -values [5]. When  $b_{\text{low}} = 0 \text{ s/mm}^2$  is used for ADC calculation, the attenuation from perfusion is fully included in the ADC, resulting in an overestimation of the diffusion. Our recommendation focuses on the measurement of water mobility in the extravascular space, which implies that perfusion effects should be excluded from the ADC. To reduce this perfusion bias, a non-zero  $b_{\text{low}}$  should be used [69,80].

There is no straightforward way to determine the exact  $b_{\text{low}}$  that should be used, as this depends on the tissue properties that are measured. As an example here, we take prostate tissue and calculate the contribution of perfusion to the total signal for a set of  $b$ -values, using the intravoxel incoherent motion model and taking differences in tissue and blood relaxation parameters into account [81]. Tissue  $T_1$  and  $T_2$  were set to 1317 and 88 ms using the 1.5 T values reported in [56], and blood  $T_1$  and  $T_2$  were set to 1441 and 290 ms [81]. The ADC, pseudo-diffusion coefficient, and perfusion fraction were  $1.34 \times 10^{-3} \text{ mm}^2/\text{s}$ ,  $21.1 \times 10^{-3} \text{ mm}^2/\text{s}$ , and 0.23 respectively [82]. The echo time (TE) and repetition time (TR) were taken from the Unity MR-linac, with the acquisition settings as shown in Supplemental Table S3.1. To maximize the gradient performance, a setting called ‘gradient overplus’, which employs the simultaneous application of multiple gradients for diffusion weighting, was used. From these simulations, we observed that at a  $b$ -value of 50  $\text{s/mm}^2$ , perfusion constitutes 14% of the total signal, which went down to 5% at 100  $\text{s/mm}^2$  and 2% at 150  $\text{s/mm}^2$ .

To illustrate how the SNR depends on the  $b$ -value on the Unity system, we measured a series of DWIs in a prostate cancer patient, where we increased the  $b$ -value from 50 to 1200  $\text{s/mm}^2$  in steps of 50  $\text{s/mm}^2$ , while keeping the TE and TR to minimal values, reflecting clinical practice. The  $b$ -values were each measured once (no averages). Other imaging parameters can be found in Supplemental Table S3.1. A noise scan with identical imaging and reconstruction parameters as the scan with  $b = 500 \text{ s/mm}^2$ , was acquired by turning the RF pulses off, which is an available research option. The SNR was calculated within the prostate for all  $b$ -values using  $\text{SNR} = S_b/S_n$ , where  $S_b$  is the mean signal in the prostate of the  $b$ -value images, and  $S_n$  is the mean in the prostate of the noise image.

### Spatial dependence of the ADC

To assess the spatial dependence of the ADC, we acquired DWI data using a cylindrical phantom with a 40 cm diameter, containing a copper sulphate solution (Philips Healthcare, Best, The Netherlands) on six Unity systems, and one diagnostic 3T Philips Ingenia for comparison. DWIs were acquired with  $b$ -values of 0, 150, and 500  $\text{s/mm}^2$ , using a turbo-spin echo (TSE) acquisition without averaging. The voxel size was  $4 \times 4 \times 5 \text{ mm}^3$  and scans were made both with and without using gradient overplus

(Supplemental Table S3.2). ADC maps were calculated using the trace images of all b-values, but also from images acquired with employing a single gradient coil (x, y, or z). Here x denotes the left-right direction, y the anterior-posterior direction, and z the cranial-caudal direction for a patient in supine position. All b-values were used because these measurements are done on a phantom where perfusion is not present. To determine the extent to which gradient nonlinearities influence the spatial variation of diffusion weighting, we applied an offline correction using a spherical harmonics expansion that describes the gradients of the Unity system [78]. The scans were repeated with a smaller phantom (15 cm diameter) containing a copper sulphate solution on a single Unity system using the same EPI readout.

### Diffusion measurements during treatment

To determine if acquiring a DWI during treatment influences the ADC values, we imaged 11 prostate cancer patients. On two consecutive treatment days, a  $T_2$ -weighted anatomical image, a  $B_0$  map, and two diffusion-weighted images were acquired, one before and one during irradiation (indicated as 'gantry off' and 'gantry on' below). Acquisition details can be found in Supplemental Table S3.3. For these specific measurements, the DWI acquisition time was kept below 3:30 minutes by reducing the number of averages, so each acquisition would fit within the irradiation time window. For the calculation of the ADC map, the  $b = 150 \text{ s/mm}^2$  and  $b = 500 \text{ s/mm}^2$  images were used.

The prostate and tumors were delineated on the  $T_2$ -weighted image of the first fraction, and the tumor boundaries were checked by a radiation oncologist consulting clinical information including pre-treatment multi-parametric MRI. Some patients had multiple lesions, which were considered individually, giving a total of 11 prostates and 15 lesions. The  $b = 0 \text{ s/mm}^2$  images were registered rigidly to the  $T_2$ -weighted images acquired during the same treatment fraction, and this registration was propagated to the ADC maps. Then, the  $T_2$ -weighted image of the second fraction was registered deformably to the  $T_2$ -weighted image of the first fraction. This registration was then propagated to the ADC maps of the second fraction. After registration, the delineations were propagated to all the separate images, and mean values were used for further analysis.

The mean ADC values of the ROIs were compared using Bland-Altman analysis for four situations: (a) fraction 1 – fraction 2 with gantry off, (b) fraction 1 – fraction 2 with gantry on, (c) gantry on – gantry off during fraction 1, and (d) gantry on – gantry off during fraction 2. The repeatability coefficient (RC) was calculated from the within-subject standard deviation of these four situations by multiplying it by 2.77 [19]. Confidence intervals for these repeatability coefficients were calculated using a  $\chi^2$  distribution with  $n(K-1)$  degrees of freedom, where n is the number of patients and K the number of measurements per patient, according to [83].

To determine the influence of the gantry on the accuracy of the ADC values, we scanned the QIBA recommended diffusion phantom (Diffusion Phantom Model 128, High Precision Devices, Inc, Boulder, Colorado) using acquisition parameters as described in the QIBA diffusion profile for measuring this phantom at 1.5 T [70]. The phantom was scanned three times, once with a static gantry, once while irradiating the phantom with an IMRT plan, and once while moving the gantry continuously. The median accuracy for these situations was calculated as the percent error compared to the known diffusion values of the phantom.

### Patient example

As a patient example of DWI and ADC image quality on the Unity system, we acquired a DWI of a rectal cancer patient following the acquisition settings that are presented as a result of the current study. The acquisition voxel size was  $4 \times 4 \times 5 \text{ mm}^3$ , and b-values of 150 and 500  $\text{s/mm}^2$  were used, with 4 and 16 averages respectively (Supplemental Table S3.4).

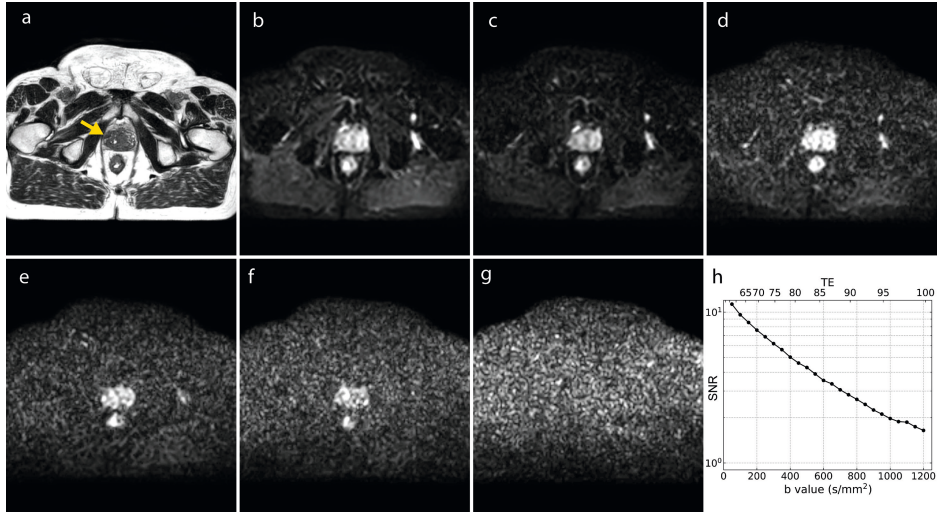
The patient studies were approved by the medical ethics committee of the Netherlands Cancer Institute and written informed consent was obtained.

## RESULTS

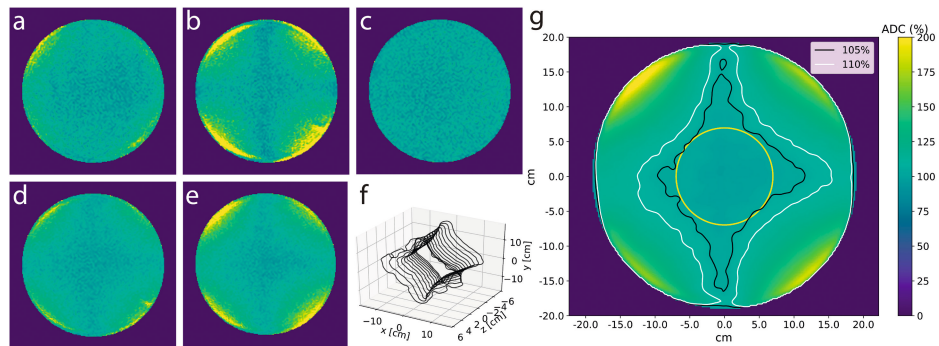
Figure 3.2 shows *in vivo* results of the SNR measurement, including a plot of the SNR as a function of increasing b-value (and TE) in the prostate (Figure 3.2h). The SNR is 11.3 at  $b = 50 \text{ s/mm}^2$  and decreases with an increasing b-value and TE. To illustrate, the SNR at  $b=500 \text{ s/mm}^2$  is 4.3 and at  $b=1000 \text{ s/mm}^2$  is 2.0. ADC values in the prostate, calculated from these scans using their corresponding  $b = 0 \text{ s/mm}^2$  images (with the same TE) were 1.82 and  $1.59 \times 10^{-3} \text{ mm}^2/\text{s}$ .

A distinct spatial pattern appears in the ADC maps of the homogeneous diffusion phantom. Figure 3.3 a-c shows when individual gradient coils are used for diffusion weighting. The most severe increase in ADC was found when using only the y-direction coil (Figure 3.3b). No spatial pattern was present when only using the z-direction gradient coil for diffusion weighting (Figure 3.3c). For both the trace image with and without gradient overplus (Figure 3.3 d and e), the spatial pattern is present. In the most extreme case, the ADC value went up to  $13 \times 10^{-3} \text{ mm}^2/\text{s}$ , which is an increase of more than 600% compared to the mean ADC of a ROI in the iso-center. The extent of the spatial variation is slightly less when gradient overplus is not used. Similar results were found on all six Unity systems (Supplemental Figure S3.1). This spatial pattern was not present on the 3T Philips Ingenia system, which showed a much more homogeneous result, especially after the gradient nonlinearity correction (Supplemental Figure S3.2b

and d). The smaller phantom imaged with an EPI readout showed results consistent with the same region in the TSE scans (Supplemental Figure S3.3).



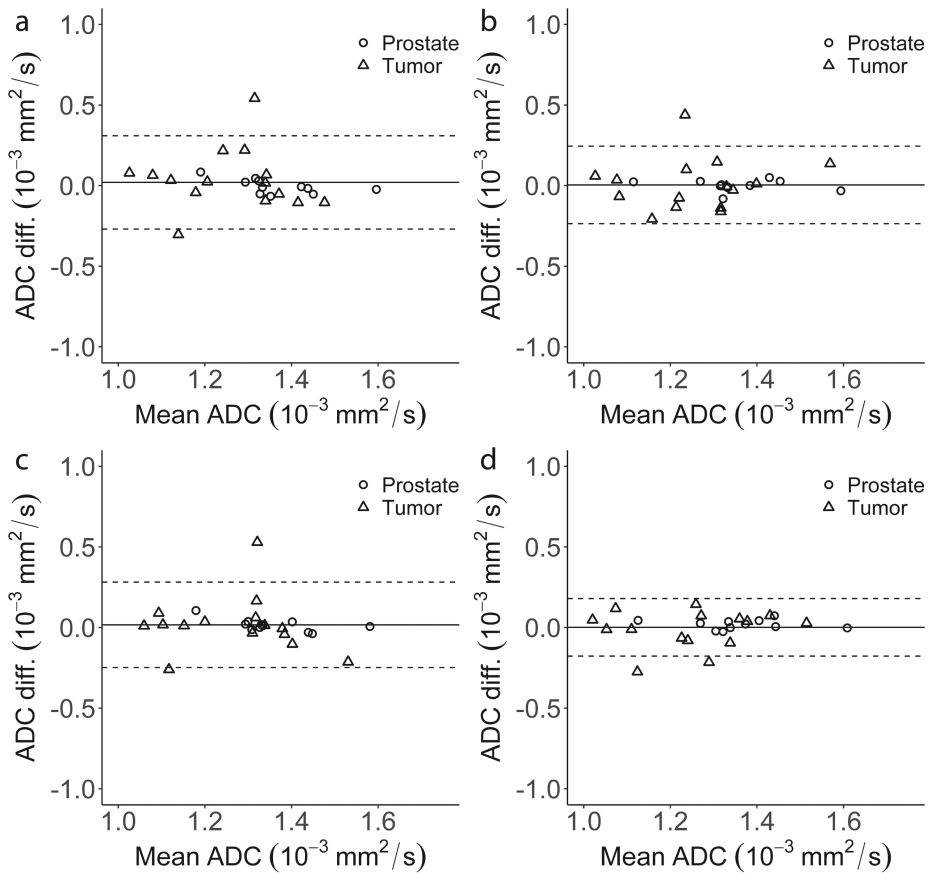
**Figure 3.2.** In vivo data of the prostate cancer patient SNR measurement. (a) T<sub>2</sub>-weighted image, the yellow arrow indicates the location of the tumor. (b–f) DWI with an increasing b-value (100, 250, 500, 750, 1000 s/mm<sup>2</sup>), acquired without averaging. The window/level of these images are decreased to facilitate the reduction in signal. Note that disappearance of surrounding structures with an increasing b-value, indicating a reduction in SNR. (g) Noise map of the patient. (h) SNR of the prostate ROI as a function of b-value (and TE).



**Figure 3.3.** ADC maps of a large cylindrical phantom, acquired with employing different gradient combinations. The phantom is placed axially and the slice in the iso-center ( $z = 0$ ) is shown. (a–c) ADC maps obtained when using single gradient directions, x, y, and z, respectively. (d) Trace ADC map combining (a–c). (e) Trace ADC map acquired while using gradient overplus. (g) A smoothed average image of the acquisitions of all six MR-linacs, with gradient overplus. The contours represent the regions within which the ADC increase is less than 5%, and less than 10% when compared to the ADC value in the center. This results in a diamond shape with a diagonal length of ~20 cm for the 5% contour. The circle in the middle has a radius of 7 cm. (f) 3D representation of the 5% contour for multiple slices of the body phantom, showing a diamond centered at the isocenter, which increases in size when moving outwards in the z-direction.



Bland-Altman plots of the ADC measurements before and during irradiation are shown in Figure 3.4, where mean values from the prostate and lesion delineations are shown. Figure 3.4a indicates the day-to-day variability under normal conditions, with the gantry turned off. The repeatability coefficients for the four situations (Figure 3.4 a-d), including 95% confidence intervals were 0.09 (0.06-0.15), 0.07 (0.05-0.11), 0.08 (0.06-0.13), and 0.07 (0.05-0.11)  $\times 10^{-3} \text{ mm}^2/\text{s}$  for the prostate, showing similar results for cases before and during treatment. For the tumors, the values were 0.37 (0.27-0.57), 0.31 (0.23-0.47), 0.34 (0.24-0.52), and 0.22 (0.16-0.35)  $\times 10^{-3} \text{ mm}^2/\text{s}$ .

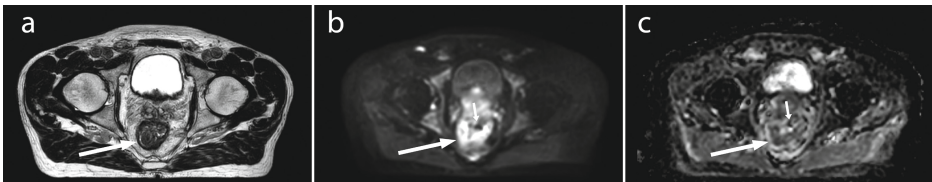


**Figure 3.4.** Bland-Altman plots of the four ADC measurements acquired on consecutive days and before and during treatment. (a) fraction 1 – fraction 2 gantry off, (b) fraction 1 – fraction 2 gantry on, (d) gantry on – gantry off fraction 1, (e) gantry on – gantry off fraction 2. The bias and limits of agreement of a-d are 0.02 (-0.27 to 0.31), 0.00 (-0.24 to 0.25), 0.02 (-0.25 to 0.28), and 0.00 (-0.18 to 0.18)  $\times 10^{-3} \text{ mm}^2/\text{s}$ .

For the accuracy measurements with the diffusion phantom, the median accuracy with a static gantry was 1.1 %, while irradiating the phantom with an IMRT plan was 0.5

%, and with a continuously rotating gantry was 1.3 %. A Bland-Altman plot with the individual data points can be found in Supplemental Figure S3.4.

A representative example of images of a rectal cancer patient is shown in Figure 3.5, including an ADC map in Figure 3.5c.



**Figure 3.5.** Representative example of a rectal cancer patient. (a)  $T_2$ -weighted image, (b) DWI  $b = 500$  s/mm<sup>2</sup> image, and (c) ADC map (scaled between 0 and  $4 \times 10^{-3}$  mm<sup>2</sup>/s). The large white arrows indicate the location of the tumor. On the DWI and ADC images, the tumor shape is slightly deformed due to the presence of air (small arrows).

## DISCUSSION

The adjusted design of the Unity system results in differences in DWI compared to diagnostic MR scanners. For this, we identified three main categories: the choice of b-values, the spatial dependence of the ADC, and diffusion measurements during treatment. Based on our experiments, simulations, and existing literature, we give recommendations for the Unity system.

It is important to realize that the calculated ADC depends on the choice of b-values and therefore should be homogenized among centers to enable comparison. To exclude perfusion from the ADC, a  $b_{\text{low}}$  value of 100 – 150 s/mm<sup>2</sup> should be used, in line with what is recommended in [69]. Ideally this should be tailored to the tissue of interest. We show an example calculation to estimate the perfusion bias which could be used to guide the decision for the  $b_{\text{low}}$  when designing a trial.

The choice to limit the  $b_{\text{high}}$  value is mainly based on the difference in  $\Delta$  as showed in Figure 1. To establish the ADC as a response biomarker, the reproducibility among different systems is important. A sequence with a different  $\Delta$  leads to different ADC values in structured tissues even when the same b-values are used. Rather than measuring at the same b-value, one should make an effort to measure at a similar  $\Delta$ . To facilitate the comparison among different systems, the  $\Delta$  should be reported. Furthermore, trials to compare ADC values from MR-linacs to diagnostic systems should be performed.

**Table 3.1.** Acquisition parameters for ADC measurements on the Unity system.

Topic	Recommendations for ADC measurements at the Unity system	Comments
Highest b-value	500 s/mm <sup>2</sup>	
Spatial dependence of the ADC	< 7 cm of the iso-center in the x and y direction. Outside this volume the ADC depends heavily on position.	Only using the z-direction gradient coils can be considered when ROI is > 7 cm outside iso-center (axially).
Treatment delivery	ADC measurements can be done during treatment delivery	
<b>Example parameters that could serve as a starting point</b>		
Acquisition Sequence	SS-EPI	SS-EPI should be used for its speed and SNR benefit.
Lipid Suppression	On	
Slice thickness (mm)	3-5	Ensure sufficient SNR in the b <sub>high</sub> image.
Gap thickness (mm)	0-1	A gap thickness of 0 should only be used if slices are acquired in an interleaved fashion.
In plane resolution (mm)	1-5	
Phase encoding direction	Choose the direction which causes the least distortions to the tumor site.	
Bandwidth (Hz/pixel)	Max possible in phase encoding direction.	Can be increased by increasing the SENSE factor.
Lowest b-value (other than b=0 s/mm <sup>2</sup> )	100 – 150 s/mm <sup>2</sup>	Depending on target site. The b = 0 s/mm <sup>2</sup> image should not be used for ADC calculation.
Number of signal averages	The highest b-value should be acquired 3-4 times as often as the lowest.	Based on [84–86].
Echo time	Shortest	Gradient overplus should be used to reduce the echo time.
SENSE	0 – 2.5x	Higher SENSE reduces geometric distortions, but also reduces SNR.
Partial Fourier (half-scan)	0.6 - 1	Using half-scan reduces SNR.

Other factors that limit the  $b_{\max}$  are the reduced SNR, the gradient performance, and the available acquisition time. To reduce the bias in the ADC value caused by Rician noise, it is important to verify, for instance in a volunteer, that the  $b_{\text{high}}$  measurement has sufficient SNR. Note here that the SNR in the target area usually also depends on body size. If the SNR needs to be improved, one or a combination of the following parameters should be changed: increasing the acquisition voxel size, lowering the SENSE encoding factor or lowering the partial Fourier imaging factor. These latter two should be considered on a tissue basis as they increase the TE, potentially also decreasing the SNR due to  $T_2$ -decay. While increasing the number of averages for

the b-values yields a more homogeneous image by reducing random signal intensity fluctuations, the ADC bias will not be reduced [87]. The total number of averages that can be acquired depend on the total acquisition time that is available. Once the total number of averages is chosen, the  $b_{\text{high}}$  images should be acquired three to four times as often as the  $b_{\text{low}}$  images [84,86].

Taking the  $\Delta$  and these other factors into consideration, we propose that a maximum  $b_{\text{high}}$  of 500 s/mm<sup>2</sup> should be used for ADC measurements on the Unity system. It is important to note that on the Unity system, the  $\Delta$  of a specific DWI sequence depends on the highest b-value acquired in that sequence. Therefore, adding additional b-values above the recommended 500 s/mm<sup>2</sup> to the sequence will increase the  $\Delta$  and result in a different ADC value even when the ADC is calculated from the  $b = 500$  s/mm<sup>2</sup> image.

ADC maps on the Unity system contain severe spatial variations, even when homogeneous solutions are used for measurements. Although these variations are more present in the trace ADC map acquired with gradient overplus (compared to the trace ADC map without, Figure 3.4e and 3.4d, respectively), we recommend scanning with this setting for its benefit of reduced TE and therefore increased SNR. The spatial dependence of the diffusion weighting is well documented in diagnostic systems where it is largely explained by gradient nonlinearity (see Supplemental Figure S3.2b and S3.2d) [58]. However, applying the offline gradient nonlinearity correction only resulted in a negligible change of the spatial variation (compare Supplemental Figure S3.2a to S3.2c). We therefore hypothesize that on the Unity system, the spatial variation is caused by increased eddy currents resulting from the split gradient coil design. Accurate ADC determination is possible within a radius of 7 cm from the iso-center, depicted as the circle in Figure 3.3g. Further research is needed to investigate if corrections of the distortions away from the iso-center are feasible.

Eddy current distortions depend on the gradient coils that are used. As visible in Supplemental Figure S3.1, the y-direction gradient has the poorest characteristics for DWI. Also, when only using the z-direction gradient coil (Figure 3.4c, Supplemental Figure S3.1 row 3), no spatial variation was present. Using only this coil could be an option for acquiring DWIs in patients that have tumor sites far outside the iso-center (in the axial plane) and are therefore impacted by the spatial dependence present in the trace images. However, unless the measured diffusion is isotropic, the resulting ADC would differ from the trace ADC.

We were unable to scan the large body phantom with an EPI readout, because in order to position the phantom, the posterior coil must be removed. Therefore, parallel imaging was unavailable resulting in severe distortions due to a low bandwidth. We showed with

a smaller phantom than our recommendation to scan within 7 cm from the iso-center holds for EPI readouts.

Scanning DWI during gantry rotation did not affect the ADC values. The accuracy of the ADC values in the diffusion phantom was comparable between measuring with a static and with a rotating gantry. A similar RC was found when comparing scanning before and during treatment to day-to-day repeatability, and the bias in the Bland-Altman plots for cases c and d (comparing static to moving gantry) were 0.02 and  $0.00 \times 10^{-3} \text{ mm}^2/\text{s}$ , respectively. Hence, DWI could be acquired at any available time during treatment to facilitate an optimal MRI protocol during each fraction.

In summary, our Unity system specific recommendations include limiting the  $b_{\text{high}}$  to  $500 \text{ s}/\text{mm}^2$ , scanning a target within 7 cm of the iso-center, and that measuring the ADC during treatment delivery is possible. Table 3.1 provides an overview of these recommendations, including example acquisition parameters to be used with the goal of measuring the ADC accurately.

The QIBA recommendations for DWI includes the need for test-retest studies in order to calculate the repeatability coefficient [19]. This should be used to establish clinical confidence in the ADC QIB and to help interpret the results from treatment response monitoring and outcome prediction studies. We also emphasize that test-retest studies should be the next step for DWI studies on Unity.

In conclusion, we provide recommendations for DWI on the Unity system, incorporating hardware specific considerations. These recommendations provide a basis for test-retest studies and, when used, will facilitate multi-center biomarker studies and future meta-analyses.

## SUPPLEMENTARY MATERIAL

**Table S3.1:** Series of DWI measurements in prostate cancer patient. The  $b_{\text{high}}$  value is increased from 50 s/mm<sup>2</sup> to 1200 s/mm<sup>2</sup>, with increments of 50 s/mm<sup>2</sup>.

Parameter	DWI (SS-EPI)
Field of view (mm <sup>3</sup> )	430 x 430 x 60
Acquired voxel size (mm <sup>3</sup> )	4 x 4 x 4
TR (ms)	2396 – 3338
TE (ms)	57 – 100
Fat suppression	SPAIR
Phase encoding BW (Hz)	34.5
Parallel imaging factor	2.3
DELTA / delta (ms)	28.5 / 7.4 – 49.9 / 28.8
Acquisition time (mm:ss)	00:14 – 00:20
b-values (averages)	0 (1), 50 – 1200 (1)

**Table S3.2:** DWI acquisition of the Philips body phantom (cylindrical phantom with a diameter of 40 cm). The values in square brackets indicate parameters corresponding to the sequence with gradient overplus.

Parameter	DWI (TSE)
Field of view (mm <sup>3</sup> )	448 x 448 x 125
Acquired voxel size (mm <sup>3</sup> )	4 x 4 x 5
TR (ms)	7500
TE (ms)	106 [89]
DELTA / delta (ms)	31.6 / 24.4
Acquisition time (mm:ss)	02:00
b-values (averages)	0 (1), 150 (1), 500 (1)

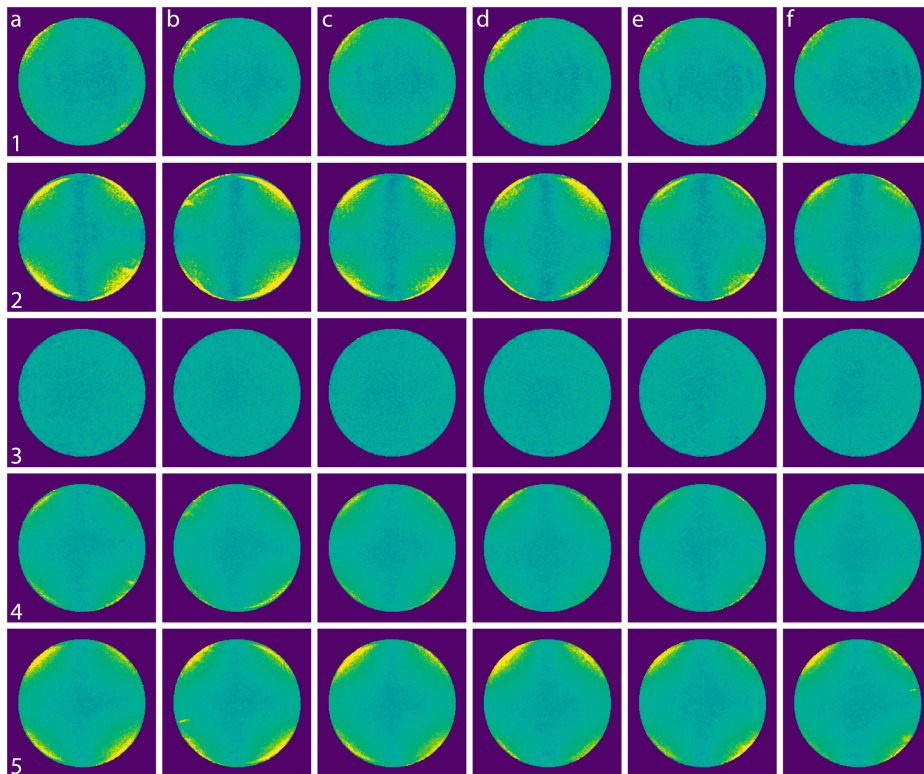
**Table S3.3:** Scan parameters of DWI of patients before and during irradiation.

Parameter	T <sub>2</sub> -weighted (TSE)	B <sub>0</sub> -map	DWI (EPI)
Field of view (mm <sup>3</sup> )	400 x 448 x 250	430 x 430 x 60	430 x 430 x 60
Acquired voxel size (mm <sup>3</sup> )	1.2 x 1.2 x 1.2	3 x 3 x 3	3 x 3 x 3
TR (ms)	1300	11	3831
TE (ms)	114	4.6	69
DELTA / delta (ms)	-	-	34.4 / 22.8
Fat suppression	-	-	SPAIR
Phase encoding BW (Hz)	-	-	20.2
Parallel imaging factor	3.5	-	2.3
Acquisition time (mm:ss)	06:25 (NSA = 2)	01:05	03:23
b-values (averages)	-	-	0 (4), 30 (4), 150 (4), 500 (8)

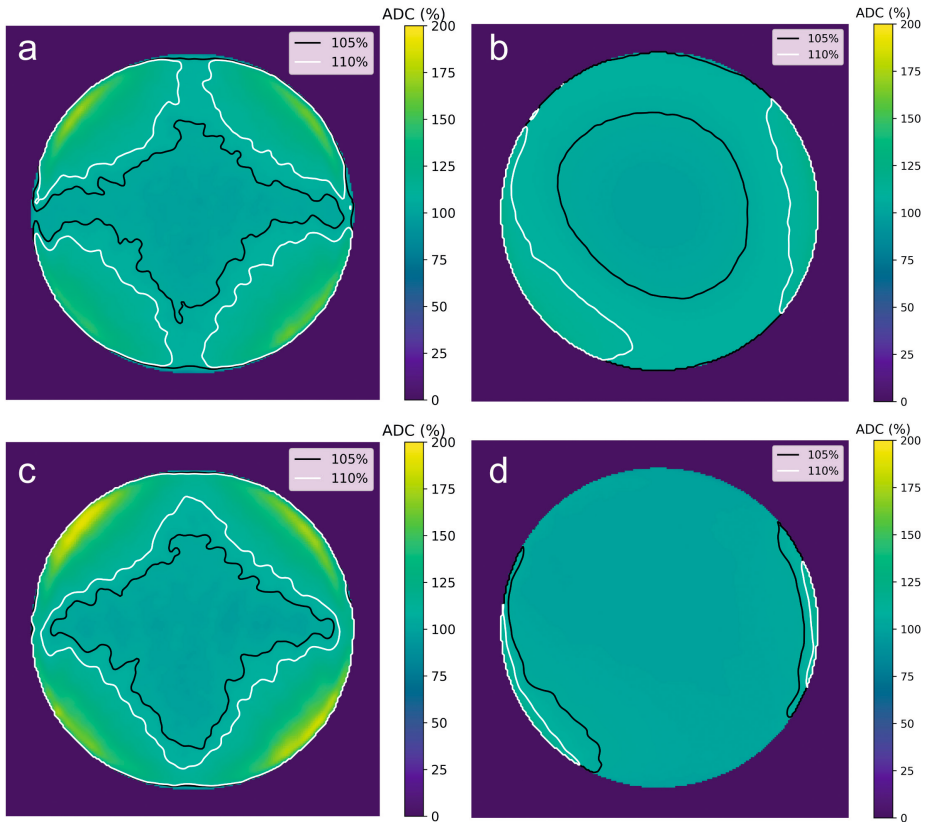
**Table S3.4:** Scan parameters of the rectal cancer patient scan.

Parameter	T <sub>2</sub> -weighted (TSE)	DWI (EPI)
Field of view (mm <sup>3</sup> )	400 x 448 x 250	430 x 430 x 100
Acquired voxel size (mm <sup>3</sup> )	1.2 x 1.2 x 1.2	4 x 4 x 5
TR (ms)	1300	4082
TE (ms)	114	85
DELTA / delta (ms)	-	42.5 / 19.6
Fat suppression	-	SPAIR
Phase encoding BW (Hz)	-	29.2
Parallel imaging factor	3.5	2
Acquisition time (mm:ss)	06:25 (NSA = 2)	04:17
b-values (averages)	-	0 (1), 150 (4), 500 (16)

**Figure S3.1:** ADC maps of a cylindrical phantom with a diameter of 40 cm. The slice in the iso-center is shown. The columns (a-f) correspond to unique Unity systems. The rows (1-5) are ordered by the gradient coils that are used for diffusion weighting. (1) x-gradient only, (2) y-gradient only, (3) z-gradient only, (4) trace of x, y, and z, (5) trace image obtained with gradient overplus. The color represents the ADC (mm<sup>2</sup>/s), scaled from 0 – 200% compared to the average ADC of a ROI in the center.

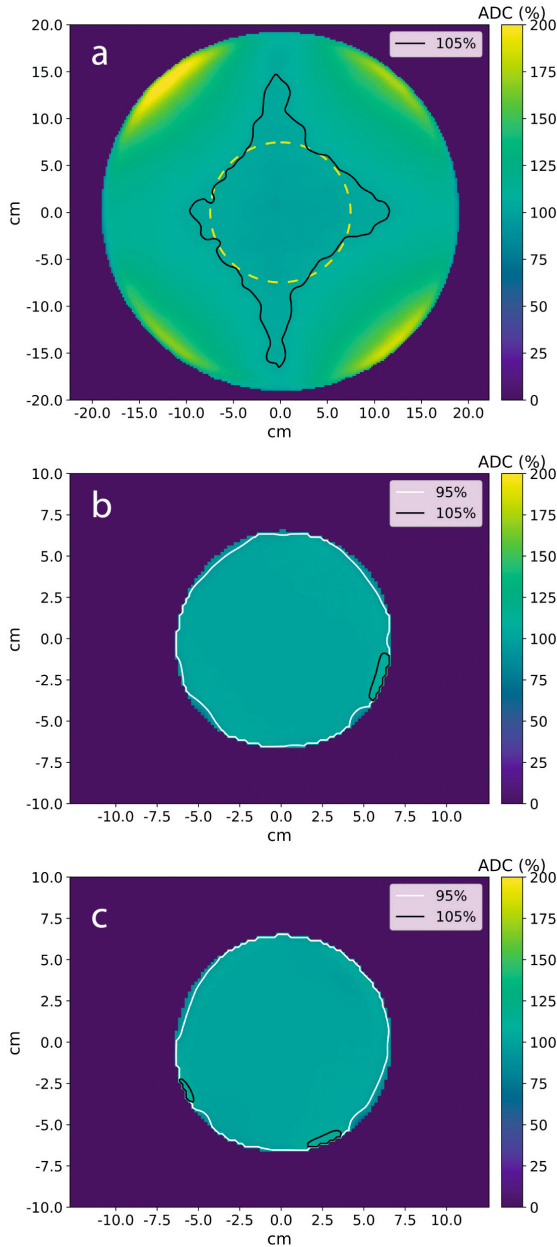


**Figure S3.2:** ADC maps of a cylindrical phantom with a diameter of 40 cm, acquired on the Unity MR-linac and on a 3T Philips Ingenia system. Trace images of x, y, and z-gradients are shown. The images are smoothed with a gaussian filter to provide smoother lines. (a, c) Images acquired on the Unity MR-linac. (b, d) Images acquired on the 3T system. C was obtained after applying an offline gradient nonlinearity correction on image a, and d was obtained after applying an offline gradient nonlinearity correction on image b. The ADC values within the black lines are within 5% of the average of a circular ROI with a diameter of 5 cm at the center of the image. The white lines encapsulate regions where the ADC increase is more than 10%.

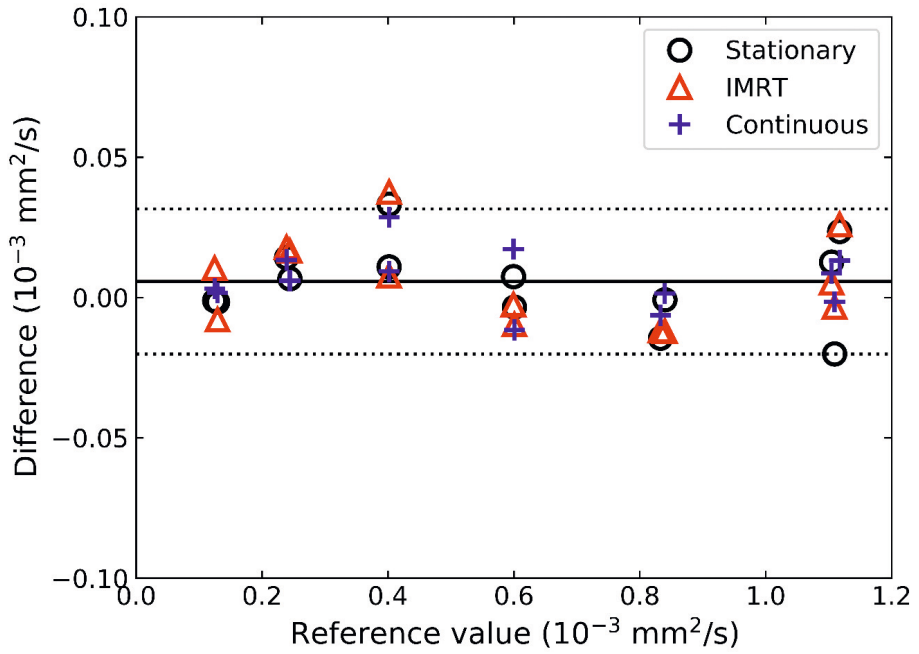




**Figure S3.3:** Comparison of the ADC maps of a phantom with a diameter of 15 cm, imaged using an EPI readout, to the large body phantom imaged with a TSE readout. a) The large body phantom is the same average image as presented in Figure 3.3g. The yellow dashed line shows a circle with a diameter of 15 cm corresponding to the smaller phantom. b) The smaller phantom imaged with an EPI readout, without gradient overplus. c) The smaller phantom imaged with an EPI readout, with gradient overplus. The white lines show the contour where the ADC values are 95% of the average of a ROI in the middle of the slice, and the black lines show the contour where the ADC values are 105%.



**Figure S3.4:** Bland-Altman plot of the QIBA phantom with and without movement of the gantry. The reference values are taken from the QIBA phantom manual. The bias is  $0.006 \times 10^{-3} \text{ mm}^2/\text{s}$ , and the limits of agreement are  $-0.020$  and  $0.032 \times 10^{-3} \text{ mm}^2/\text{s}$ .





DAILY INTRAVOXEL  
INCOHERENT MOTION  
(IVIM) IN PROSTATE CANCER  
PATIENTS DURING MR-GUIDED  
RADIOTHERAPY –  
A MULTICENTER STUDY

Ernst S. Kooreman, Petra J. van Houdt, Rick Keesman, Vivian W. J. van Pelt, Marlies E. Nowee, Floris Pos, Karolina Sikorska, Andreas Wetscherek, Arndt-Christian Müller, Daniela Thorwarth, Alison C. Tree, Uulke A. van der Heide

*Frontiers in Oncology 11, article 705964 (2021)*

## ABSTRACT

*Purpose:* Daily quantitative MR imaging during radiotherapy of cancer patients has become feasible with MRI systems integrated with linear accelerators (MR-linacs). Quantitative images could be used for treatment response monitoring. With intravoxel incoherent motion (IVIM) MRI, it is possible to acquire perfusion information without the use of contrast agents. In this multicenter study, daily IVIM measurements were performed in prostate cancer patients to identify changes which potentially reflect response to treatment.

*Materials and Methods:* Forty-three patients were included which were treated with 20 fractions of 3 Gy on a 1.5 T MR-linac. IVIM measurements were performed on each treatment day. The diffusion coefficient (D), perfusion fraction (f), and pseudo-diffusion coefficient (D\*) were calculated based on the median signal intensities in the non-cancerous prostate and the tumor. Repeatability coefficients (RC) were determined based on the first two treatment fractions. Separate linear mixed-effects models were constructed for the three IVIM parameters..

*Results:* In total, 726 fractions were analyzed. Pre-treatment average values, measured on the first fraction before irradiation, were  $1.46 \times 10^{-3}$  mm<sup>2</sup>/s, 0.086, and  $28.7 \times 10^{-3}$  mm<sup>2</sup>/s in the non-cancerous prostate and  $1.19 \times 10^{-3}$  mm<sup>2</sup>/s, 0.088, and  $28.9 \times 10^{-3}$  mm<sup>2</sup>/s in the tumor, for D, f, and D\*, respectively. The repeatability coefficients for D, f, and D\* in the non-cancerous prostate were  $0.09 \times 10^{-3}$  mm<sup>2</sup>/s, 0.05, and  $15.3 \times 10^{-3}$  mm<sup>2</sup>/s. In the tumor, these values were  $0.44 \times 10^{-3}$  mm<sup>2</sup>/s, 0.16, and  $76.4 \times 10^{-3}$  mm<sup>2</sup>/s. The mixed effects analysis showed an increase in D of the tumors over the course of treatment, while remaining stable in the non-cancerous prostate. The f and D\* increased in both the non-cancerous prostate and tumor.

*Conclusions:* It is feasible to perform daily IVIM measurements on an MR-linac system. Although the repeatability coefficients were high, changes in IVIM perfusion parameters were measured on a group level, indicating that IVIM has potential for measuring treatment response.

## INTRODUCTION

Integrated MR-linac systems combine an MRI scanner with a linear accelerator, allowing acquisition of MRI scans of the patient on each treatment fraction of a radiotherapy (RT) course. On two commercially available systems, acquisition of quantitative MRI was shown to be feasible [17,72]. Daily monitoring of radiotherapy response using quantitative MR imaging biomarkers has become more readily available with the increasing number of MR-linac systems in centers worldwide [88].

Quantitative MRI enables the characterization of tissue properties in a quantitative manner. By measuring this on a daily basis, two exciting ideas for personalized radiotherapy come within reach. The first is to adapt the dose distribution of a treatment plan on a daily basis according to the changing patient biology [89], and the second is to base the total dose that a patient receives on the biological response [90]. For this to become clinical practice, the performance of MR-linacs regarding quantitative MRI first needs to be validated [12]. Furthermore, it needs to be established if daily changes in imaging biomarkers are detectable and if these changes are associated with clinical outcome.

Perfusion is of interest as it is related to tumor hypoxia, which is a prognostic marker for overall survival in a number of tumor sites [91]. An established method for imaging perfusion and permeability in cancer is dynamic contrast-enhanced (DCE) MRI [92]. However, as this requires the injection of an MRI contrast agent, DCE MRI is not suitable for daily treatment response monitoring. An alternative to DCE MRI is intravoxel incoherent motion (IVIM) imaging [7], which is a technique based on diffusion-weighted MRI (DWI). In DWI, MR images are sensitized to random motion by the application of strong diffusion-weighting gradients. The amount of diffusion weighting is expressed with the b-value, where a higher b-value indicates stronger diffusion weighting. Typically two or three images are acquired with a different b-value, from which the apparent diffusion coefficient (ADC) is calculated using a mono-exponential model [5]. With IVIM, additional low b-values are acquired in order to extract information about perfusion [80]. By fitting a bi-exponential model, IVIM allows for the determination of the tissue diffusion coefficient  $D$ , the perfusion- or blood fraction  $f$ , and the pseudo-diffusion coefficient  $D^*$ , thereby separating perfusion and diffusion effects. In prostate cancer,  $D$  (and ADC) parameters were shown to be related to cell density [93,94]. The IVIM parameter  $f$  was shown to correlate with blood vessel density in [93]. Changes in IVIM parameters during treatment might provide valuable information about treatment response [89]. For cervical cancer, early increases in  $f$  have been associated with good response [95,96]. Similarly, in head-and-neck cancer patients, larger reductions in  $f$  and higher  $D$  values were observed in patients with regional failure compared to patients with regional control [97]. In another study

with weekly measurements in head-and-neck cancer patients, a significant increase was found in  $D$  in complete responders, but no significant differences in  $f$  and  $D^*$  were found between responders and non-responders [98]. Daily IVIM measurements in patients with brain metastases showed an increase in  $D$  in responders and a decrease in non-responders [36]. For prostate cancer, only DWI has been investigated as a potential biomarker for treatment response. Two studies have shown an increase in the ADC during radiation treatment [99,100]. Therefore, the aim of this multicenter study was to perform daily IVIM measurements in prostate cancer patients to identify if time trends appear in IVIM parameters which might have potential for treatment response monitoring.

## **MATERIALS AND METHODS**

### **Patients**

Forty-three patients from three institutes with intermediate and high-risk biopsy-proven prostate cancer were included in this study according to the EAU risk classification [101]. Twelve patients were included in the first institute, 8 in the second, and 23 in the third. All patients received the same treatment of 20 fractions of 3 Gy over the course of four to five weeks on a 1.5 T MR-linac system (Unity, Elekta AB, Sweden). In addition, 34 patients also received androgen deprivation therapy (ADT). Patient demographics are presented in Table 4.1. The study was approved by the institutional review boards and written informed consent was obtained from all patients.

### **MRI**

During each treatment fraction, an anatomical  $T_2$ -weighted scan for position verification and an IVIM scan were acquired before the start of irradiation. Thus, the scans on the day of the first fraction provide pre-treatment information. All institutes used the same protocol for the IVIM scan. For the development of the IVIM protocol, previously published guidelines were followed for ADC measurements on the Unity MR-linac [102]. A maximum  $b$ -value of 500 s/mm<sup>2</sup> was recommended to compensate for the limited SNR of the Unity MR-linac and to measure at a diffusion time that is comparable to that of diagnostic systems [102]. An extra  $b$ -value of 30 s/mm<sup>2</sup> was added to be able to measure IVIM parameters. The averages of the  $b = 0$  s/mm<sup>2</sup> image were increased to eight. Sequence parameters can be found in Table 4.2.

**Table 4.1.** Patient demographics. iPSA: initial prostate-specific antigen. ADT: androgen deprivation therapy. ISUP Grade Group: Revised prostate cancer grading system introduced by the International Society of Urological Pathology (ISUP). The median (range) is shown for Age, iPSA, and ADT.

Age	73 (55 – 83)
iPSA (ng/mL)	8.5 (4.4 – 37.6)
ADT (months before start of radiation)	2 (0 – 11)
<b>ISUP Grade Group</b>	
1	4
2	19
3	13
4	5
5	2
<b>T-stage</b>	
T1a	1
T1c	9
T2a	11
T2b	1
T2c	12
T3a	7
T3b	2

### Image registration and delineation

The  $T_2$ -weighted images of each fraction were registered rigidly to the  $T_2$ -weighted image of the first fraction within a box around the prostate using the correlation ratio as a cost function. This rigid registration allowed for translations and rotations. Next, the  $b = 0$  s/mm<sup>2</sup> images were registered to the  $T_2$ -weighted image acquired during the same fraction. All registrations were checked visually and improved manually if required.

The prostate and all visible tumors were delineated on the  $T_2$ -weighted image of the first fraction. The tumors were delineated while consulting diagnostic multi-parametric scans acquired according to the PI-RADS v2.1 guideline [103]. Tumor delineations were excluded from the prostate delineation to obtain the non-cancerous prostate region. Only the tumor focus with the largest volume was used for the analysis in case of multiple foci per patient. All delineations were propagated to the IVIM scans in order to extract quantitative values. Due to the use of an EPI readout, severe susceptibility artifacts could be present in some IVIM images caused by passing air in the rectum. Therefore, the  $b = 500$  s/mm<sup>2</sup> images were checked visually and fractions where air was present inside the propagated contours were excluded. The median values of the signal intensities of the voxels inside the resulting delineations were used for calculation of the IVIM parameters.



**Table 4.2.** Acquisition parameters of the IVIM sequence.

Sequence type	Single-shot echo planar image (ss-EPI)
Field of view (mm <sup>3</sup> )	430 x 430 x 60
Acquired voxel size (mm <sup>3</sup> )	4 x 4 x 4
TR/TE (ms)	2960/82
b-values (averages) (s/mm <sup>2</sup> )	0 (8), 30 (8), 150 (8), 500 (16)
Gradient timings $\Delta/\delta$ (ms)	41/20
Fat suppression	SPAIR
SENSE factor	2.3 (left-right)
Phase encoding bandwidth (Hz/pixel)	32.9
Acquisition time (m:ss)	5:11

### IVIM parameter calculation

The IVIM parameters were calculated by performing a bi-exponential fit in a segmented fashion to increase robustness [104]

$$S_b = S_0(f e^{-b D^*} + (1 - f) e^{-b D}). \quad (4.1)$$

The diffusion coefficient  $D$  was calculated using image intensities at the two highest b-values (150 and 500 s/mm<sup>2</sup>) under the assumption that the contribution of perfusion to the signal at these b-values is negligible [80] using

$$D = \frac{\ln(S_{150}/S_{500})}{(b_{500} - b_{150})}. \quad (4.2)$$

Here,  $S_b$  is the signal intensity in the image acquired at a certain b-value. Next, the perfusion fraction  $f$  was calculated using the previously calculated  $D$  by extrapolating the contribution of the diffusion fraction to  $S_0$  as follows

$$f = 1 - (S_{150}/S_0)e^{(b_{150}-b_0)D}. \quad (4.3)$$

Finally,  $D^*$  was calculated using the obtained values of  $D$  and  $f$  in combination with the signal intensity at the lowest two b-values (0 and 30 s/mm<sup>2</sup>)

$$D^* = -\frac{1}{b_{30}} \ln \left( \frac{S_{30}/S_0 - (1 - f)e^{-b_{30}D}}{f} \right). \quad (4.4)$$

### Statistics

To establish if treatment effects could be found on a population level, for each fraction, the mean and the standard error of the mean of the IVIM parameters of all patients was determined for the tumor and non-cancerous prostate. The difference between pre-

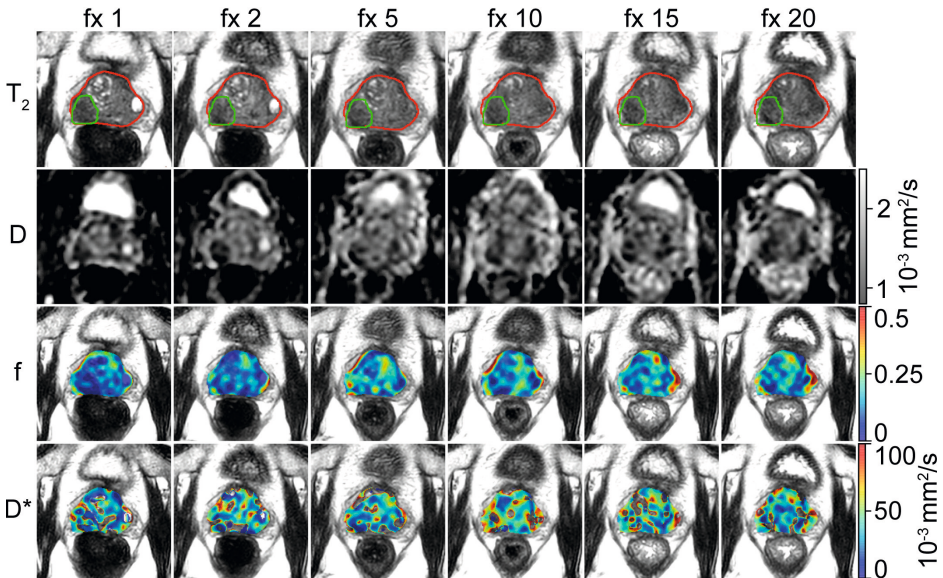
treatment values of tumor and non-cancerous prostate was tested with a two-sided paired t-test with a significance level of  $\alpha = 0.05$ .

To determine which changes in IVIM parameters can be attributed to a treatment effect, the repeatability coefficient (RC) of each IVIM parameter was calculated using  $RC = 1.96\sqrt{2 \text{wVar}}$ , where wVar is the mean within-patient variance [19,105]. The wVar was determined for the non-cancerous prostate and tumor based on the measurements from the first and second treatment fraction, assuming a negligible influence of the single 3 Gy dose that was received in between. The RC values were related to the size of the ROIs.

To analyze the evolution over time, linear mixed effects analysis was performed using R (v3.6.1) and the lme4 package [106]. Separate models were constructed for the D, f, and D\* parameters. Fixed effects were fraction (1-20), ROI (non-cancerous prostate/tumor), ISUP group, ADT, and institute. The ISUP scores were divided into a low (ISUP score 1 and 2) and high (ISUP score 3, 4, and 5) group. For ADT the number of months between the start of ADT and start of radiotherapy was used. Patients were included as a random effect. ROIs were modelled as a random effect nested within the patient. This allows the model intercepts to vary among patients and among ROIs within patients. The three models (for D, f, and D\*) were constructed separately using backwards elimination as implemented by the step function from the lmerTest package [107]. All fixed effects, including their interaction with fraction, were included in the full model. They were then eliminated one at a time based on a significance level of  $\alpha = 0.05$ , where the *p*-value was calculated using an *F*-test based on Satterthwaite's approximation.

## RESULTS

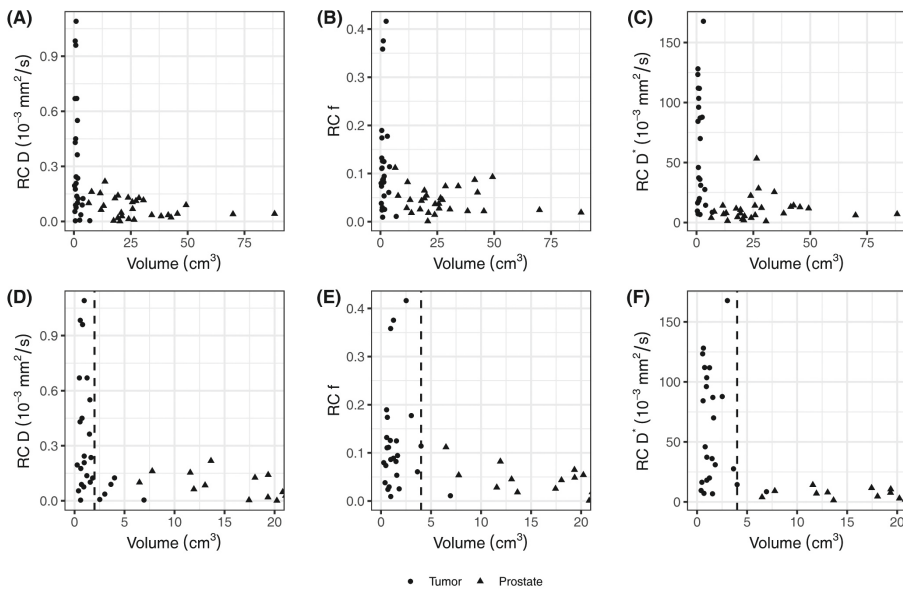
For logistical and technical reasons IVIM scans were missing in 56 out of the total of 860 fractions. From these, 73 were excluded because of anatomical deformations or susceptibility distortions caused by the EPI readout. Five were excluded because the patient moved between the acquisition of different b-values. This left 726 fractions for analysis with a median number of 18 (range 9-20) available fractions per patient. In four patients, a tumor could not be distinguished and was not delineated. For those patients, the entire prostate region was analyzed as non-cancerous. Figure 4.1 shows the IVIM parameter maps for six fractions from a single patient.



**Figure 4.1.** Example of a prostate cancer patient. A voxel-wise map of the IVIM parameters is shown for 6 treatment fractions (fraction 1, 2, 5, 10, 15 and 20). The prostate is delineated in red, and the tumor in green. The images are resampled to the reconstructed voxel sizes of the  $T_2$ -weighted acquisition ( $0.6 \times 0.6 \times 1.2 \text{ mm}^3$ ). Note that for analysis, the signal intensities from inside the non-cancerous prostate (i.e. prostate minus tumor) were used to calculate the IVIM parameters.

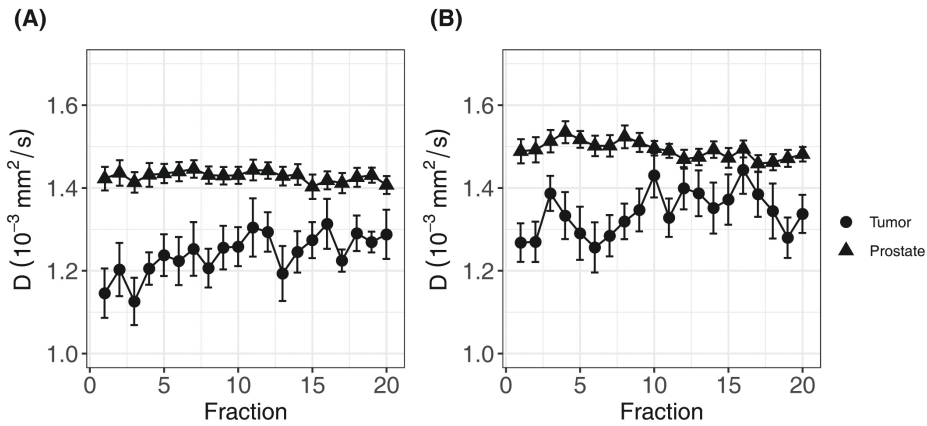
IVIM scans were available for the first fraction in 35 patients. The pre-treatment average and standard error of the mean of  $D$  were  $1.46 \pm 0.02 \times 10^{-3} \text{ mm}^2/\text{s}$  in the non-cancerous prostate, which was significantly higher ( $p < 0.001$ ) than in the tumor ( $1.19 \pm 0.04 \times 10^{-3} \text{ mm}^2/\text{s}$ ). The pre-treatment average and standard error of the mean of  $f$  were  $0.086 \pm 0.005$  in the non-cancerous prostate and  $0.088 \pm 0.01$  in the tumor. The average and standard error of the mean of  $D^*$  were  $28.7 \pm 1.4 \times 10^{-3} \text{ mm}^2/\text{s}$  in the non-cancerous prostate and  $28.9 \pm 5.4 \times 10^{-3} \text{ mm}^2/\text{s}$  in the tumor. The pre-treatment values of  $f$  and  $D^*$  were not significantly different between the non-cancerous prostate and tumor.

The RC in the non-cancerous prostate was  $0.09 \times 10^{-3} \text{ mm}^2/\text{s}$  for  $D$ ,  $0.05$  for  $f$ , and  $15.3 \times 10^{-3} \text{ mm}^2/\text{s}$  for  $D^*$ . In the tumor, the RCs were  $0.44 \times 10^{-3} \text{ mm}^2/\text{s}$ ,  $0.16$ , and  $76.4 \times 10^{-3} \text{ mm}^2/\text{s}$  for  $D$ ,  $f$ , and  $D^*$ , respectively. Figure 4.2 shows that the RC depends on the size of the ROI. The median volume of the non-cancerous prostate delineations was  $24$  (range  $6.5 - 88$ )  $\text{cm}^3$ , whereas the median volume of the tumor delineations was  $1.0$  (range  $0.3 - 6.9$ )  $\text{cm}^3$ . As shown in Figure 4.2D, the RC of  $D$  steeply increases for volumes below  $2 \text{ cm}^3$ , and Figures 4.2E and 4.2F show a similar increase for  $f$  and  $D^*$  for volumes below  $4 \text{ cm}^3$ .



**Figure 4.2.** The repeatability coefficients (RCs) for the IVIM parameters based on the values from the tumor and non-cancerous prostate on the first and second treatment fraction. The top row shows all data points and the bottom row shows the same data but zoomed in on the smaller volumes. In figures (D), (E), and (F), a vertical dashed line indicates the volume below which the RC steeply increases. This value is 2 cm<sup>3</sup> in (D) and 4 cm<sup>3</sup> in (E) and (F).

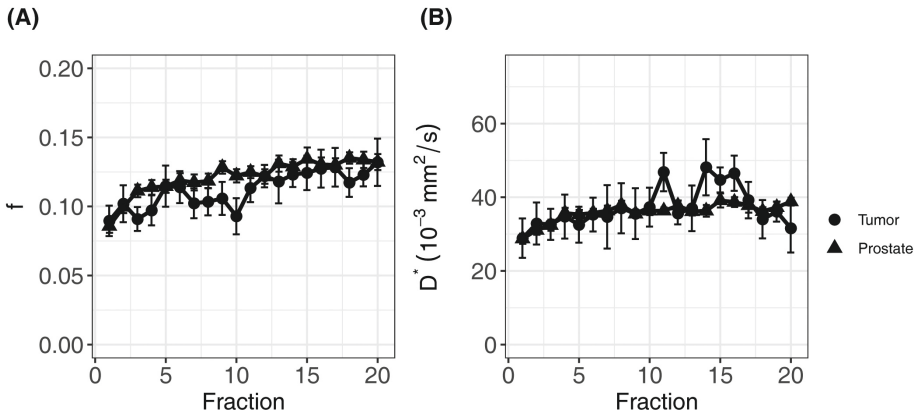
To analyze the evolution over time, mixed effects models were constructed for each IVIM parameter. The set of fixed effects and regression coefficients for each parameter are listed in Table 4.3. For D, these included ISUP groups (low/high,  $p$ -value = 0.003), the ROI (non-cancerous prostate/tumor,  $p$ -value < 0.001), and the fraction (1-20,  $p$ -value < 0.001). Figure 4.3 shows the mean of the D for each fraction for the low and high ISUP groups. The effect size for the difference between the group of patients with a high ISUP score compared to patients with a low score was  $-0.10 \pm 0.03 \times 10^{-3} \text{ mm}^2/\text{s}$ . The D in the non-cancerous prostate was  $0.24 \pm 0.03 \times 10^{-3} \text{ mm}^2/\text{s}$  higher than in the tumor. Both ISUP groups and ROI had an interaction term with the fraction number, meaning that the change in D over the course of treatment was different for these groups. In the tumor, for patients with a low ISUP score, the D increased  $0.005 \pm 0.001 \times 10^{-3} \text{ mm}^2/\text{s}/\text{fraction}$ , whereas for the group with a high ISUP score the increase was  $0.007 \pm 0.001 \times 10^{-3} \text{ mm}^2/\text{s}/\text{fraction}$ . This reduces the difference in the D between these groups over the course of treatment: at the 20<sup>th</sup> fraction, the D as estimated from the model in the tumors of the low ISUP group is increased to  $1.38 \pm 0.02 \times 10^{-3} \text{ mm}^2/\text{s}$ , and in the high ISUP group to  $1.33 \pm 0.03 \times 10^{-3} \text{ mm}^2/\text{s}$ .



**Figure 4.3.** Average of  $D$  of all patients over the course of radiotherapy treatment. (A) shows the average for patients in the low ISUP group, (B) shows the average for patients in the high ISUP group. Error bars indicate the standard error of the mean. As indicated by the result of the mixed effects model, the increase in the high ISUP group (B) is steeper than in the low ISUP group (A).

For the perfusion fraction  $f$ , the significant fixed effects were ROI (non-cancerous prostate/tumor,  $p$ -value = 0.03), the fraction (1-20,  $p$ -value < 0.001), and the interaction between fraction and institute ( $p$ -value = 0.04). As institute is part of the interaction term, it was also added to the model as a fixed effect ( $p$ -value = 0.06) (Table 4.3). The  $f$  values in the non-cancerous prostate were  $0.013 \pm 0.006$  higher than in the tumor. An average increase per treatment fraction of  $0.002 \pm 0.0002$  was found in both the non-cancerous prostate and tumor for institutes 1 and 3. For institute 2, this increase was significantly lower ( $p$ -value = 0.01) at  $0.001 \pm 0.0005$  per treatment fraction, which was the only significant effect containing institute. Figure 4.4A shows the mean  $f$  values per fraction grouped by ROI.

For the pseudo-diffusion coefficient  $D^*$ , the significant fixed effects were the fraction (1-20,  $p$ -value < 0.001) and ADT (months before the start of treatment,  $p$ -value < 0.001) (Table 4.3). The  $D^*$  changed with  $-1.37 \pm 0.35 \text{ mm}^2/\text{s}$  for every month of ADT. The change in  $D^*$  due to ADT was independent of treatment fraction. The  $D^*$  increased with  $0.35 \pm 0.09 \times 10^{-3} \text{ mm}^2/\text{s}$  each fraction, in both the non-cancerous prostate and the tumor. Figure 4.4B shows the mean values per fraction grouped by ROI.



**Figure 4.4.** Average of  $f$  (A) and  $D^*$  (B) of all patients over the course of radiotherapy treatment. Error bars indicate the standard error of the mean. Both  $f$  and  $D^*$  increase over the course of treatment in both the tumor and the non-cancerous prostate.

**Table 4.3.** Model parameters of the mixed effects models for  $D$ ,  $f$ , and  $D^*$ .

	Regression coefficients ( $\beta$ )	Std. error
<b>Model for <math>D</math></b>		
	<b>(<math>10^{-3} \text{ mm}^2/\text{s}</math>)</b>	
Intercept ( $\beta_0$ )	1.284	0.025
Fraction (per one unit)	0.005	0.001
ISUP high (versus ISUP low)	-0.100	0.032
ROI non-cancerous prostate (versus tumor)	0.242	0.026
Interaction Fraction – ISUP high	0.003	0.001
Interaction Fraction – ROI non-cancerous prostate	-0.007	0.001
<b>Model for <math>f</math></b>		
Intercept ( $\beta_0$ )	0.090	0.007
Fraction (per one unit)	0.002	0.0003
ROI non-cancerous prostate (versus tumor)	0.013	0.006
Institute 1 (versus institute 3)	-0.011	0.010
Institute 2 (versus institute 3)	0.021	0.012
Interaction Fraction – Institute 1	-0.0006	0.0004
Interaction Fraction – Institute 2	-0.001	0.0006
<b>Model for <math>D^*</math></b>		
	<b>(<math>10^{-3} \text{ mm}^2/\text{s}</math>)</b>	
Intercept ( $\beta_0$ )	36.6	1.71
Fraction (per one unit)	0.35	0.09
ADT (per one unit)	-1.37	0.35

## DISCUSSION

In this multicenter study we acquired daily IVIM scans of prostate cancer patients during radiotherapy treatment on three 1.5 T MR-linac systems. IVIM parameters were calculated from the median signal intensities of the tumor and non-cancerous prostate. We analyzed the changes in these parameters over the course of the treatment. The diffusion coefficient  $D$  showed an increase in the tumor, while the values in the non-cancerous prostate remained unchanged. The parameters  $f$  and  $D^*$  increased in the tumor as well as in the non-cancerous prostate.

The average pre-treatment  $D$  values in the non-cancerous prostate are in line with values reported in the literature, although the range of the reported values in the literature is large:  $0.16 - 1.78 \times 10^{-3} \text{ mm}^2/\text{s}$  [108]. For the tumor, our average pre-treatment  $D$  ( $1.21 \pm 0.04 \times 10^{-3} \text{ mm}^2/\text{s}$ ) is higher than previously reported (range  $0.13 - 1.06 \times 10^{-3} \text{ mm}^2/\text{s}$ ) [108]. The lower  $D$  in the tumor for the high-ISUP group is consistent with the literature [108].

The pre-treatment  $f$  and  $D^*$  values in our study are within the range that was previously reported in the literature [108]. In their meta-analysis, He et al. found no difference in  $f$  between the tumor and the non-cancerous prostate, which is consistent with our pre-treatment findings [108]. However, in contrast to our findings, they did find a difference in  $D^*$  between the non-cancerous prostate and tumor. A reason for this could be the high variance in  $D^*$  in the current study, in combination with the small standardized mean difference of  $0.29 \times 10^{-3} \text{ mm}^2/\text{s}$  between tumor and non-cancerous prostate reported by He et al. [108].

The RC depended on the size of the ROIs [57,109]. We observed a strong increase in the RC with lower ROI sizes. The mean RC for  $D$  in the tumor corresponded to 36 % of the mean value in the tumor. This means that a change of 36 % would have to occur in order to be significant. While this corresponds to earlier reported values [19], such large changes are not expected in prostate cancer. Van Schie et al. found a change on the group level caused by radiotherapy of 20 % and Foltz et al. of 13 % [99,100]. Other tumor sites may have larger tumors, which would reduce their RC, or exhibit larger changes throughout treatment and hold therefore more potential for treatment response monitoring using DWI or IVIM. The same holds for  $f$  and  $D^*$ , where the RCs in the tumor were even higher.

All IVIM parameters, except for  $D$  of the non-cancerous prostate, increased during treatment. Interestingly, for the high ISUP group,  $D$  increased more during the treatment than for the low ISUP group. This suggests that the cellularity at the end of treatment was similar for both groups. Further work is needed to establish if these

observations are linked to treatment outcome. The  $f$  and  $D^*$  also showed an increase during treatment. For  $D^*$  we saw an effect of hormonal therapy, where  $D^*$  was reduced slightly with an increasing duration of ADT before the start of radiation treatment. This is consistent with a reduction in DCE parameters, which is linked to devascularization in patients that received ADT [110,111]. As the entire prostate gland is irradiated, the overall increase in the  $f$  and  $D^*$  values might be caused by an inflammation response of the prostate [112], obscuring more subtle differences that might be present between the tumor and non-cancerous prostate.

A limitation of this study is the use of rigid registration to match the scans from all fractions to the scan of the first fraction. This type of registration cannot account for anatomical deformations caused by e.g. passing air in the rectum. Moreover, we saw that the contrast of the  $T_2$ -weighted images inside the prostate reduced over the course of treatment, causing the tumor to disappear. This reduction in contrast has been reported before [99,100,113], but made it impossible to check the propagated tumor contours visually in later fractions. Because these tumor volumes are relatively small, a small mismatch could lead to a significant difference in the tumor values. In an effort to reduce the influence of small misregistrations, we calculated the IVIM parameters based on the median values of the signal intensities inside the delineations, thereby reducing the effect of outliers.

As indicated by the RCs, the noise in the IVIM acquisition posed problems for voxel-based analysis, especially for the  $f$  and  $D^*$ . This can also be seen in the voxel-wise maps shown in Figure 4.1, where holes appear in the  $D^*$  maps. This happens when due to noise, the logarithm that is used in Equation 4.4 becomes undefined. By using the median values of the signal intensities for estimation of the parameters, the influence of noise was reduced.

It must be noted that the RC was based on the first two treatment fractions and therefore might include some treatment effect. The RC denotes the smallest significant difference between two measurements taken under identical conditions, with 95% confidence [105]. While it is useful for the comparison of the precision of our measurements to previously reported studies, it might not be the right metric to denote a significant change in a time series. As there are multiple measurements per patient, a small change compared to the pre-treatment value that is consistent over time could be statistically significant even if that change is smaller than the RC.

In conclusion, we have successfully acquired daily IVIM scans in prostate cancer patients on the Unity MR-linac system. On a group level, changes in IVIM parameters caused by radiation treatment were found, indicating that it might be useful for treatment response evaluation.





LONGITUDINAL CORRELATIONS  
BETWEEN INTRAVOXEL  
INCOHERENT MOTION (IVIM)  
AND DYNAMIC CONTRAST-  
ENHANCED (DCE) MRI DURING  
RADIOTHERAPY IN PROSTATE  
CANCER PATIENTS

Ernst S. Kooreman, Vivian W. J. van Pelt, Marlies E. Nowee, Floris Pos, Uulke A. van  
der Heide, Petra J. van Houdt

*Frontiers in Oncology 12, article 897130 (2022)*

## ABSTRACT

*Purpose:* Intravoxel incoherent motion (IVIM) is a promising technique that can acquire perfusion information without the use of contrast agent, contrary to the more established dynamic contrast-enhanced (DCE) technique. This is of interest for treatment response monitoring, where patients can be imaged on each treatment fraction. In this study, longitudinal correlations between IVIM- and DCE parameters were assessed in prostate cancer patients receiving radiation treatment.

*Materials and Methods:* 20 prostate cancer patients were treated on a 1.5T MR-linac with  $20 \times 3$  or 3.1 Gy. Weekly IVIM and DCE scans were acquired. Tumors, the peripheral zone (PZ), and the transition zone (TZ) were delineated on a  $T_2$ -weighted scan acquired on the first fraction. IVIM and DCE scans were registered to this scan and the delineations were propagated. Median values from these delineations were used for further analysis. The IVIM parameters  $D$ ,  $f$ ,  $D^*$  and the product  $fD^*$  were calculated. The Tofts model was used to calculate the DCE parameters  $K_{trans}$ ,  $kep$  and  $v_e$ . Pearson correlations were calculated for the IVIM and DCE parameters on values from the first fraction for each region of interest (ROI). For longitudinal analysis, the repeated measures correlation coefficient was used to determine correlations between IVIM and DCE parameters in each ROI.

*Results:* When averaging over patients, an increase during treatment in all IVIM and DCE parameters was observed in all ROIs, except for  $D$  in the PZ and TZ. No significant Pearson correlations were found between any pair of IVIM and DCE parameters measured on the first fraction. Significant but low longitudinal correlations were found for some combinations of IVIM and DCE parameters in the PZ and TZ, while no significant longitudinal correlations were found in the tumor. Notably in the TZ, for both  $f$  and  $fD^*$ , significant longitudinal correlations with all DCE parameters were found.

*Conclusions:* The increase in IVIM- and DCE parameters when averaging over patients indicates a measurable response to radiation treatment with both techniques. Although low, significant longitudinal correlations were found which suggests that IVIM could potentially be used as an alternative to DCE for treatment response monitoring.

## INTRODUCTION

Non-invasive perfusion imaging is of interest in oncology, as low perfusion is related to hypoxia which holds prognostic value [27,93,114]. A common way to measure perfusion is by using dynamic contrast enhanced (DCE-) MRI [114–116]. In addition to prognosis, DCE has been shown to have value for mid-treatment response assessment in cervix [117], esophageal [118], and head-and-neck cancer [119–121].

Acquiring quantitative MRI (qMRI) images during radiation treatment for the purpose of treatment response monitoring has become feasible with the introduction of MR-guided radiotherapy. Using MR-linacs, which consist of a linear accelerator integrated with an MRI system, qMRI sequences can be acquired on each treatment fraction, without the increase of patient burden [17,20,62,89,122,123].

Although DCE-MRI is a candidate for treatment response monitoring, acquiring a DCE scan during each treatment fraction is undesirable due to the use of contrast agent. Alternative techniques that can provide perfusion information without the use of contrast agent are needed. One such alternative is intravoxel incoherent motion (IVIM), which is an extension to diffusion weighted imaging (DWI) [80]. IVIM parameters provide information about diffusion and perfusion. It is based on the concept that inside a voxel, signal from water flowing in the capillaries can be separated from diffusing water [7]. In addition to the diffusion coefficient ( $D$ ), the perfusion parameters  $f$  (perfusion fraction),  $D^*$  (pseudo-diffusion coefficient), and the product  $fD^*$  can be determined.

Previous studies have investigated correlations between IVIM and DCE-MRI parameters in different tumor sites, with conflicting results [23]. These studies usually determine the correlation between IVIM and DCE parameters on a single time point. For treatment response purposes however, correlations between changes in parameters, induced by radiation treatment, are more relevant. A study performed in 21 liver tumor-bearing rabbits assessed the correlations between IVIM and DCE parameters longitudinally, while the rabbits were treated with a vascular disrupting agent [124]. Interestingly, the authors did not find any significant correlations between IVIM and DCE parameters when assessing the imaging time points separately, but did find a significant longitudinal correlation. This longitudinal correlation is of importance for treatment response monitoring purposes and indicates that IVIM could be a potential substitute for DCE-MRI for this purpose.

In the current study, longitudinal correlations between IVIM- and DCE parameters are assessed in a cohort of prostate cancer patients that were imaged weekly during radiation treatment. Each week a DCE and an IVIM scan were acquired to enable

longitudinal assessment. The aim of this study is to determine whether IVIM and DCE parameters correlate when measured longitudinally and whether there is potential for IVIM to substitute DCE for treatment response monitoring.

## MATERIALS AND METHODS

### Patients

Twenty patients, with a median age of 70.5 (range 53 – 82) years with biopsy proven prostate cancer were included in this study. Only patients with an adequate renal function (glomerular filtration rate  $GFR > 60 \text{ ml/min/1.7m}^2$ ) were included. Thirteen patients were treated with  $20 \times 3 \text{ Gy}$  and due to a change in clinical practice, seven patients were treated with  $20 \times 3.1 \text{ Gy}$ . Treatment took place over the course of five weeks. Patient characteristics are presented in Table 5.1. The study was approved by the local ethics committee and each patient gave written informed consent.

**Table 5.1.** Patient characteristics. iPSA, initial prostate specific antigen; GFR, glomerular filtration rate; ISUP score, prostate cancer grading score.

Patient characteristic	Median (range)
Age (years)	70.5 (53 – 82)
iPSA (ng/ml)	15 (8 – 38)
GFR (ml/min/1.7m <sup>2</sup> )	
Pre-treatment	79 (67 – 107)
Post-treatment	82 (65 – 110)
ISUP	No. of patients
1	3
2	8
3	4
4	3
5	2

### Image Acquisition

All patients were treated on a 1.5 T MR-linac (Unity, Elekta AB, Stockholm, Sweden). This is a hybrid system, where a linear accelerator is integrated with an MRI scanner to enable concurrent patient irradiation and MRI acquisition. The MRI system of the MR-linac is based on a 1.5 T Ingenia system (Philips Healthcare, Best, The Netherlands), with split gradient coils to create a window for the radiotherapy beam [1]. The system uses an 8-channel radio-translucent phased array receive coil [16].

A  $T_2$ -weighted anatomical scan, an IVIM scan and a DCE-MRI scan were acquired weekly over the course of five weeks, starting at the first day of treatment. Scan parameters

can be found in Table 5.2. The IVIM sequence was optimized for the MR-linac system, which has lower gradient performance compared to diagnostic systems and lower SNR due to the simpler receive coil system [102,122]. To compensate this, the highest b-value was limited to 500 s/mm<sup>2</sup>, and a relatively large isotropic acquisition voxel size of 4 mm<sup>3</sup> was used. To calculate contrast agent concentration values, the pre-contrast T<sub>1</sub> was measured using the variable flip angle (VFA) method with a similar sequence as the DCE scan, but with a TR/TE of 20/4 ms and flip angles of 3, 6, 10, 20, and 30 °.

**Table 5.2.** MRI sequence parameters

	T2-weighted	IVIM	DCE
Sequence type	3D-TSE	ss-EPI	3D-FFE
Field of view (mm <sup>3</sup> )	400 x 448 x 250	430 x 430 x 60	220 x 251 x 60
Acquired voxel size (mm <sup>3</sup> )	1.2 x 1.2 x 1.2	3.98 x 3.98 x 4.00	2.62 x 2.62 x 7.00
Reconstructed voxel size (mm <sup>3</sup> )	0.57 x 0.57 x 1.2	1.92 x 1.92 x 4.00	1.57 x 1.57 x 3.50
Flip angle (°)	90	90	35
TR/TE (ms)	1300/129	2960/82	4.0/1.9
Fat suppression	-	SPAIR	-
Parallel imaging (SENSE) factor	3.5	2.3	2
Acceleration factor	110	47	-
b-values (averages) (s/mm <sup>2</sup> )	-	0 (8), 30 (8), 150 (8), 500 (16)	-
Phase encoding bandwidth (Hz/pixel)	-	32.9	-
Gradient timings $\Delta/\delta$ (ms)	-	41.1/20.0	-
Dynamic scan time (s)	-	-	2.8
Number of dynamics	-	-	110
NSA	2	1	1
Acquisition time (m:ss)	5:48	5:11	5:04

For the DCE scan, during the fifth dynamic, 15 mmol gadoteric acid (Dotarem, Geurbet, France) was injected at a rate of 3 mL/s using a power injector followed by a 30 ml saline flush. While a study by Wang et al. demonstrated no significant effect of radiation on the chemical composition of Gadolinium based contrast agents [125], DCE scans were acquired after the radiation treatment, without repositioning of the patient to avoid interactions of the contrast agent with radiation.

### Image Registration

Tumor, peripheral zone (PZ), and transition zone (TZ) were delineated on the T<sub>2</sub>-weighted scans of the first fraction. Of three patients, who received a trans-urethral resection of the prostate (TURP), the TURP cavity was delineated to be excluded from

analysis. Tumors were delineated while consulting biopsy results and diagnostic images, following the PI-RADS V2.1 criteria [103].

The IVIM and DCE images were registered separately to the  $T_2$ -weighted scan of the first fraction which contained the delineations using rigid registration allowing rotations and translations. For IVIM, the  $b = 0$  s/mm<sup>2</sup> image was used as this contains the most anatomical information. For DCE, the 100<sup>th</sup> dynamic was used as a scan with relatively high enhancement in the prostate signal. All registrations were checked visually and corrected manually when needed. After registration, the delineations were propagated to the IVIM and DCE scans, where only voxels that were fully inside the propagated delineation were included for further analysis.

IVIM scans were excluded when susceptibility artifacts were present inside any of the delineations, or when movement between b-values was present. DCE scans were excluded if patient movement occurred during the scan.

The volume of the structures was calculated by multiplying the number of voxels completely inside the delineation by the voxel size of the  $T_2$ -weighted scan they were delineated on.

## Image processing

### IVIM

The bi-exponential IVIM model,  $S_{(b)} / S_0 = fe^{-bD^*} + (1 - f)e^{-bD}$ , was fitted using a segmented approach [104]. Using the median signal intensity values from the delineations, the tissue diffusion coefficient (D) was determined first using the two highest b-values (150 and 500 s/mm<sup>2</sup>). Next the perfusion fraction (f) was calculated using this D and the  $b = 0$  s/mm<sup>2</sup> signal intensity. Both D and f were then used in combination with the signal intensities from the lowest two b-value images (0 and 30 s/mm<sup>2</sup>) to calculate the pseudo-diffusion coefficient  $D^*$ . The parameter  $fD^*$  was calculated by multiplying f with  $D^*$ .

### DCE

To extract an arterial input function (AIF), external iliac artery was delineated on all DCE scans of all patients. Due to slight variations in the  $B_1$  field (see Supplementary Figure S5.1), only the left external iliac artery was used. Signal intensities were converted to concentration time curves using the spoiled gradient echo equation following Schabel and Parker [46] assuming a  $T_1$  value of 1429 ms for blood at 1.5 T [126] and a contrast agent relaxivity of 3.6 L mM<sup>-1</sup> s<sup>-1</sup> [127]. Following Georgiou et al. the maximum relative change in concentration during the DCE scan was determined for all voxels inside this delineation [128]. The voxels between the 50<sup>th</sup> and 95<sup>th</sup> percentile of this relative change were averaged to obtain an AIF for each treatment fraction. Per

patient, the median AIF of all five measurements, based on peak height, was used for all tracer kinetic modeling for that patient. Supplementary Figure S5.2 shows all AIFs of all patients.

A voxel-wise  $T_1$ -map was calculated from the VFA series using a linear implementation [129]. The  $T_1$  map was used to convert signal intensity to concentration values using the method of Schabel and Parker [46]. The bolus arrival time was estimated for each voxel using an automated method [130]. The volume transfer constant ( $K^{\text{trans}}$ ) and the rate constant ( $k_{\text{ep}}$ ) from the standard Tofts model [131] were calculated on a voxel-basis following the approach developed by Murase [49] using and median AIF as input. The extracellular extravascular space volume fraction ( $v_e$ ) was then calculated on a voxel basis using ( $K^{\text{trans}}/k_{\text{ep}}$ ).

### Statistics

Baseline values from the IVIM and DCE parameters were taken from the scans of the first fraction. To check for differences in parameters between ROIs, a one-way analysis of variance (ANOVA) was performed for each parameter, with ROI as the independent variable. ANOVA results are presented with their  $F$ -statistic including within- and between group degrees of freedom, and  $p$ -value. Pearson correlation coefficients were calculated between IVIM and DCE parameters of the first fraction for each ROI.

To determine longitudinal correlations between the IVIM and DCE parameters, the *rmcorr* package in R was used [132]. The *rmcorr* package provides a repeated measures correlation ( $r_{rm}$ ), which takes into account the non-independence of repeated measures. To do so, the relationship between two continuous variables (in this case the IVIM and DCE parameters) is determined while controlling for between-patient variance. Specifically, separate parallel lines are fitted to the data of each patient using a common slope but allowing the intercept to vary per patient [132]. The  $r_{rm}$  is then calculated from the sum of squares values for the measure and the error as follows

$$r_{rm} = \sqrt{\frac{SS_{\text{Measure}}}{SS_{\text{Measure}} + SS_{\text{Error}}}}. \quad (5.1)$$

The sign of  $r_{rm}$  is taken from the sign of the common slope. The degrees of freedom are calculated using  $N(k-1)-1$ , where  $k$  is the (average) number of repeated measures per participant and  $N$  is the total number of participants [132].

As IVIM and DCE measure different biological properties which are both related to perfusion, it is possible that their correlation depends on the particular tissue measured. Therefore,  $r_{rm}$  was calculated separately for each ROI. It can be interpreted as the intra-patient correlation between IVIM and DCE parameters during radiation treatment for a

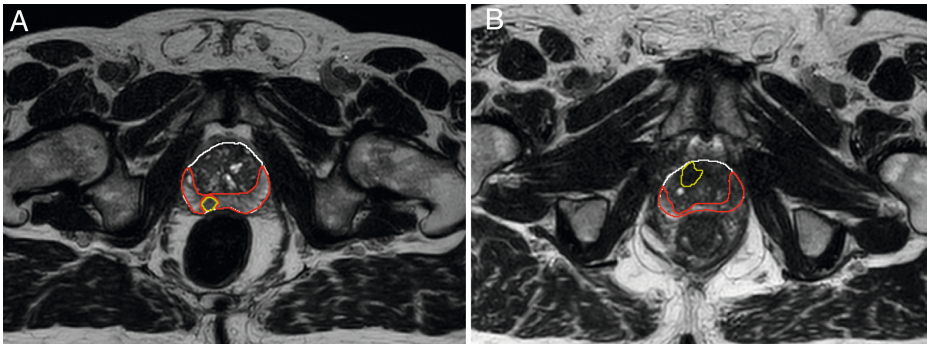


given ROI. Repeated measures correlation results are presented as  $r_{rm}$  (error degrees of freedom),  $p$ -value, and a 95% confidence interval calculated using bootstrapping with 10.000 resamples. Statistical significance was assumed for all tests when  $p < 0.05$ .

## RESULTS

Imaging data was acquired on five fractions for 19/20 patients and one patient was imaged four times. This resulted in a total of 99 fractions with IVIM and DCE scans. Two DCE scans were excluded due to movement during acquisition, both from the same patient. Seven IVIM scans were excluded due to susceptibility artifacts causing deformations within the delineations and two IVIM scans were excluded because the patient moved between the acquisition of images with a different  $b$ -value, leaving 97 DCE acquisitions and 90 IVIM acquisitions for further analysis.

In two of the patients, no tumor was visible on the diagnostic scans and therefore not delineated. Of one patient with a TURP all remaining tissue was treated as tumor. The median (range) volume of the ROIs were 0.9 (0.1 – 14)  $\text{cm}^3$  for the tumor, 8.9 (5.0 – 26)  $\text{cm}^3$  for the PZ, and 20 (7.2 – 66)  $\text{cm}^3$  for the TZ. An example of the delineations in two different patients is shown in Figure 5.1.



**Figure 5.1.** Example of delineations of the different prostate zones in two different patients (A, B). Delineations were made on  $T_2$ -weighted scans from the first treatment fraction. The entire prostate is shown in white, the peripheral zone (PZ) in red, and the tumor in yellow. The transition zone (TZ) was extracted in post processing by subtracting the PZ from the prostate delineation. Tumor voxels were excluded from all other zones during analysis.

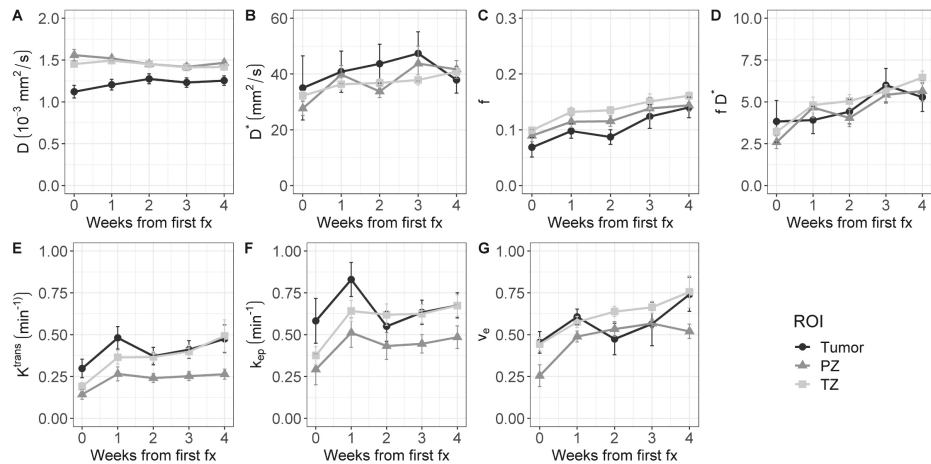
Baseline mean values with the standard error of the mean (SEM) are presented in Table 5.3. These are based on the IVIM scans acquired before the patients received any radiation and the DCE scans acquired directly after a single dose of 3 or 3.1 Gy. One-way ANOVA revealed a statistically significant difference between the tumor, PZ, and TZ for  $D$  ( $F_{2,44} = 15$ ,  $p < 0.001$ ),  $K^{\text{trans}}$  ( $F_{2,53} = 4.3$ ,  $p = 0.02$ ) and  $v_e$  ( $F_{2,53} = 3.9$ ,  $p = 0.03$ ). No

statistically significant correlations were found between IVIM and DCE parameters when using values from the first fraction only.

**Table 5.3.** Pre-treatment values of the IVIM and DCE parameters. The IVIM parameters were acquired before irradiation, the DCE parameters were acquired directly after receiving the first treatment fraction.

	Tumor	PZ	TZ
D (10 <sup>-3</sup> s/mm <sup>2</sup> )	1.12 ± 0.08	1.56 ± 0.07	1.45 ± 0.02
f	0.07 ± 0.02	0.09 ± 0.01	0.10 ± 0.01
D* (10 <sup>-3</sup> s/mm <sup>2</sup> )	35 ± 12	28 ± 3	32 ± 2
fD* (10 <sup>-3</sup> s/mm <sup>2</sup> )	3.9 ± 1.3	2.7 ± 0.4	3.2 ± 0.3
K <sub>trans</sub> (min <sup>-1</sup> )	0.30 ± 0.04	0.14 ± 0.02	0.19 ± 0.02
kep (min <sup>-1</sup> )	0.58 ± 0.09	0.29 ± 0.07	0.38 ± 0.05
ve	0.45 ± 0.08	0.25 ± 0.08	0.44 ± 0.05

Figure 5.2 shows the average time trends over all patients of the IVIM and DCE parameters. All IVIM and DCE parameters increase in all ROIs over the weeks, except for D in the PZ and TZ. The IVIM perfusion parameters increase steadily over the weeks. The DCE parameters steeply increase from the first to the second week and stabilize or slightly increase after that.



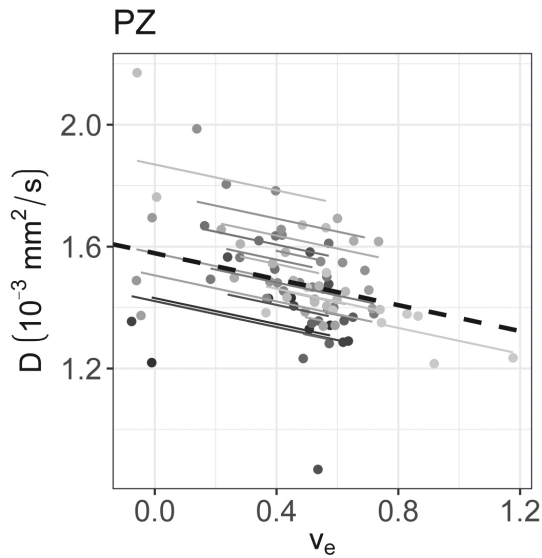
**Figure 5.2.** Evolution of intravoxel incoherent motion (IVIM, A-D) and dynamic contrast-enhanced (DCE, E-G) parameters during radiation treatment. The average value of all patients is shown for the tumor, PZ, and TZ. Error bars indicate the standard error of the mean (SEM).

The  $r_{rm}$  calculated on the longitudinal data are presented in Table 5.4. No statistically significant correlations were found between any IVIM and DCE parameter in the tumor.

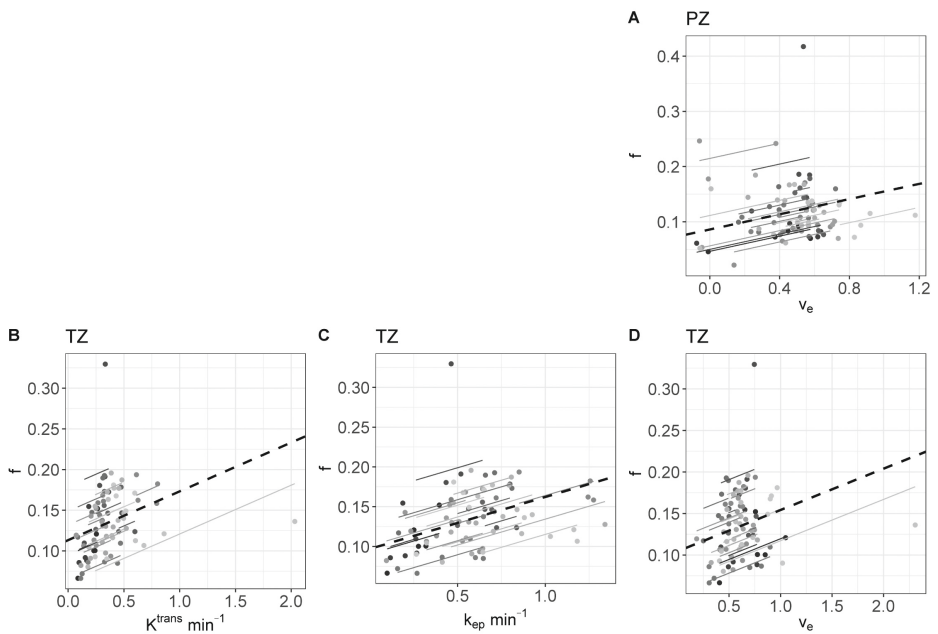
**Table 5.4.** Repeated measures correlations between IVIM and DCE parameters, separately presented for each ROI. The degrees of freedom are shown between parentheses and the 95% confidence interval of the repeated measures correlation is shown between brackets. Bold values show significant correlations ( $p < 0.05$ ).

	$K^{\text{trans}}$	$k_{\text{ep}}$	$v_e$
Tumor	D $r_{(60)} = 0.04 [-0.13, 0.24]$ , $p = 0.74$	$r_{(60)} = -0.08 [-0.33, 0.19]$ , $p = 0.55$	$r_{(60)} = 0.19 [-0.05, 0.41]$ , $p = 0.15$
	f $r_{(60)} = 0.09 [-0.09, 0.32]$ , $p = 0.48$	$r_{(60)} = 0.02 [-0.19, 0.27]$ , $p = 0.86$	$r_{(60)} = -0.12 [-0.43, 0.24]$ , $p = 0.34$
	D* $r_{(54)} = -0.02 [-0.33, 0.25]$ , $p = 0.89$	$r_{(54)} = -0.13 [-0.35, 0.13]$ , $p = 0.34$	$r_{(54)} = 0.02 [-0.23, 0.27]$ , $p = 0.90$
	fD* $r_{(54)} = 0.03 [-0.24, 0.27]$ , $p = 0.82$	$r_{(54)} = -0.09 [-0.33, 0.19]$ , $p = 0.53$	$r_{(54)} = -0.08 [-0.35, 0.22]$ , $p = 0.58$
PZ	D $r_{(63)} = -0.21 [-0.39, -0.02]$ , $p = 0.09$	$r_{(63)} = -0.06 [-0.27, 0.15]$ , $p = 0.64$	<b><math>r_{(63)} = -0.33 [-0.54, -0.11]</math>, <math>p &lt; 0.01</math></b>
	f $r_{(63)} = 0.21 [0.01, 0.47]$ , $p = 0.10$	$r_{(63)} = 0.07 [-0.15, 0.39]$ , $p = 0.56$	<b><math>r_{(63)} = 0.33 [0.14, 0.57]</math>, <math>p &lt; 0.01</math></b>
	D* $r_{(63)} = 0.16 [-0.04, 0.36]$ , $p = 0.19$	$r_{(63)} = 0.12 [-0.06, 0.32]$ , $p = 0.34$	$r_{(63)} = 0.04 [-0.16, 0.27]$ , $p = 0.75$
	fD* $r_{(63)} = 0.23 [-0.03, 0.47]$ , $p = 0.07$	$r_{(63)} = 0.13 [-0.12, 0.43]$ , $p = 0.29$	$r_{(63)} = 0.20 [-0.01, 0.40]$ , $p = 0.11$
TZ	D $r_{(63)} = -0.01 [-0.17, 0.25]$ , $p = 0.94$	$r_{(63)} = 0.16 [-0.01, 0.35]$ , $p = 0.21$	$r_{(63)} = -0.13 [-0.29, 0.06]$ , $p = 0.29$
	f <b><math>r_{(63)} = 0.38 [0.28, 0.64]</math>, <math>p &lt; 0.01</math></b>	<b><math>r_{(63)} = 0.39 [0.19, 0.60]</math>, <math>p &lt; 0.01</math></b>	<b><math>r_{(63)} = 0.37 [0.28, 0.62]</math>, <math>p &lt; 0.01</math></b>
	D* $r_{(63)} = 0.21 [0.03, 0.52]$ , $p = 0.09$	<b><math>r_{(63)} = 0.35 [0.16, 0.54]</math>, <math>p &lt; 0.01</math></b>	$r_{(63)} = 0.19 [0.02, 0.53]$ , $p = 0.12$
	fD* <b><math>r_{(63)} = 0.39 [0.26, 0.66]</math>, <math>p &lt; 0.01</math></b>	<b><math>r_{(63)} = 0.48 [0.27, 0.66]</math>, <math>p &lt; 0.001</math></b>	<b><math>r_{(63)} = 0.37 [0.24, 0.63]</math>, <math>p &lt; 0.01</math></b>

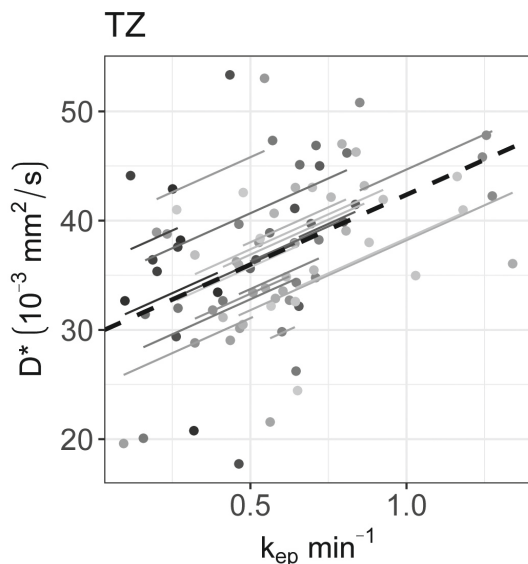
In the PZ, statistically significant correlations were found only between D and  $v_e$  and between f and  $v_e$ . In the TZ, statistically significant correlations were found between f and  $K^{\text{trans}}$ , f and  $k_{\text{ep}}$ , and f and  $v_e$ . D\* correlated significantly only with  $k_{\text{ep}}$ , while the product fD\* did so with all DCE parameters. Graphs showing the common slope and the slope per patient of the significant within-subject longitudinal correlations are presented in Figure 5.3 for D, Figure 5.4 for f, Figure 5.5 for D\* and Figure 5.6 for fD\*.



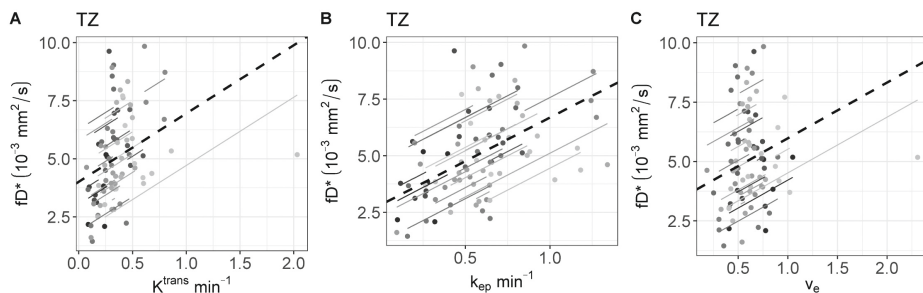
**Figure 5.3.** The significant repeated measures correlation of  $D$  with the DCE parameters are shown.  $D$  only correlated significantly with  $v_e$  in the PZ. Each line shows the fit for a single patient and the dashed black line shows the overall common slope.



**Figure 5.4.** The significant repeated measures correlations of  $f$  with the DCE parameters are shown. (A) is the correlation between  $f$  and  $v_e$  in the PZ, (B-D) are values from the TZ. Each line shows the fit for a single patient and the dashed black line shows the overall common slope.



**Figure 5.5.** The significant repeated measures correlation of  $D^*$  with the DCE parameters are shown.  $D^*$  only correlated significantly with  $k_{ep}$  in the TZ. Each line shows the fit for a single patient and the dashed black line shows the overall common slope.



**Figure 5.6.** The significant repeated measures correlation of  $fD^*$  with the DCE parameters are shown.  $fD^*$  only correlated significantly with DCE parameters in the TZ (A-C). Each line shows the fit for a single patient and the dashed black line shows the overall common slope.

## DISCUSSION

In this study, the longitudinal correlations between IVIM and DCE parameters in different ROIs of prostate cancer patients were assessed during radiation treatment. Weekly IVIM and DCE scans were performed and resulting correlations were tested taking into account the non-independence of repeated measurements on the same patients.

Baseline  $f$  and  $D^*$  values of the IVIM parameters corresponded to values found in the literature, although the reported range is large. The baseline tumor  $D$  values found in this study ( $1.12 \pm 0.08 \cdot 10^{-3} \text{ mm}^2/\text{s}$ ) were higher than previously found (reported range:

0.13 – 1.06  $10^{-3}$  mm<sup>2</sup>/s) [108]. Baseline  $K^{\text{trans}}$  and  $v_e$  values were consistent with those found in the literature, while the  $k_{\text{ep}}$  values were relatively low [133].

When averaging over patients, an increase in all perfusion parameters over the course of radiation treatment can be seen. In the DCE parameter values, this increase was the largest between week 0 and week 1, after which the values seemed to stabilize. This trend is also visible in the IVIM parameters  $D^*$  and  $fD^*$ . The similar behavior on the group level suggests that there is an overall biological response to radiation that can be measured similarly with both techniques. Previous results comparing DCE parameters before treatment to values acquired at a minimum of two years after treatment showed a decrease in  $K^{\text{trans}}$  and  $k_{\text{ep}}$  in the PZ and TZ [113]. Taken together with the current results, this could indicate that perfusion is increased during treatment, followed by a decline longer after treatment. The discrepancy between short-term and long-term differences highlights the importance of determining the optimal measurement time for treatment response purposes.

A possible explanation for the early increase of the perfusion parameters in all prostate zones could be an inflammatory response to the radiation treatment in the entire prostate, similar to what was found previously in cervix patients [117]. Such an overall response could limit the predictive value of early perfusion for treatment response in prostate cancer patients as it could obscure more subtle changes related to outcome. To investigate this, early changes in perfusion parameters should be related to clinical outcome data. However, these data were not available yet for the current study population.

Although comparison with histology has shown that IVIM parameters provide perfusion information, the specific interpretation of IVIM parameters and their relation to DCE parameters remains unclear [23,134]. Correlations between IVIM and DCE parameters should be carefully interpreted based on the context. When the goal is to assess the ability of both techniques to differentiate between tumor and benign tissue, as done in Pang et al. for prostate cancer [135], it is appropriate to use values from both ROIs combined to determine the correlation. In that case, the correlation reflects how differences between ROIs in IVIM parameters correlate with differences between ROIs in DCE parameters. A ROI effect is clearly visible in the scatterplots presented by Pang et al. [135]. However, when investigating longitudinal data, we are interested in the correlation of changes within the ROIs over time. This within-ROI correlation could be different for different ROIs, and theoretically even have an opposite sign compared to the between-ROI correlation. This effect is known as Simpsons paradox [136].

In the current study, the focus is on treatment response monitoring. To measure the longitudinal correlations the  $r_{\text{rm}}$  was used on data from each ROI separately. The  $r_{\text{rm}}$  can

in this case be interpreted for each ROI as the intra-patient correlation between IVIM and DCE parameters while measuring during treatment, indicating the degree to which both parameters reflect the same time trends induced by irradiation.

No significant correlations were found in the tumors. A reason for this could be a low precision of the IVIM and DCE parameters as acquired in the current study. Median values were calculated per ROI, and the variance of these median values scale with  $1/n$ , where  $n$  is the number of voxels. As prostate tumors are relatively small, the variance of the median values is relatively high. We showed previously that the test-retest repeatability coefficient of IVIM parameters in prostate tumors is high for the current imaging sequence and analysis:  $0.44 \cdot 10^{-3} \text{ mm}^2/\text{s}$ ,  $0.16$  and  $76.4 \text{ mm}^2/\text{s}$  for  $D$ ,  $f$ , and  $D^*$  [122]. Additionally, DCE parameters are known to have poor repeatability [19,137,138]. Within-patient coefficients of variation reported previously in prostate tumors measured on a 1.5 T system were around 20% for  $K^{\text{trans}}$ , 15% for  $v_e$  and 30% for  $k_{ep}$  [139]. Poor repeatability in both IVIM and DCE parameters can attenuate the correlation coefficients [140]. In the TZ, which is the ROI with the largest volume, significant positive correlations were found, although all were low ( $< 0.5$ ).

In order to test the correlations between IVIM and DCE parameters in different ROIs, 36 statistical tests were performed with a significance threshold of  $\alpha = 0.05$ . This means that the chance of finding at least one false positive result is 84%, because multiple testing inflates the type 1 error rate. However, since this is the first study to test the longitudinal correlation between IVIM and DCE parameters in humans undergoing radiation treatment, type 1 error rate is less of a concern. These correlations can be used as a direction for future studies.

In conclusion, when assessing changes in group averages over time, a clear increase in IVIM perfusion parameters was found. This increase was also present in all DCE parameters. Although low, it is encouraging that significant longitudinal correlations were found between IVIM- and DCE parameters, suggesting that IVIM could potentially be used as an alternative to DCE for treatment response monitoring purposes, in particular when repeated DCE-MRI is not feasible.

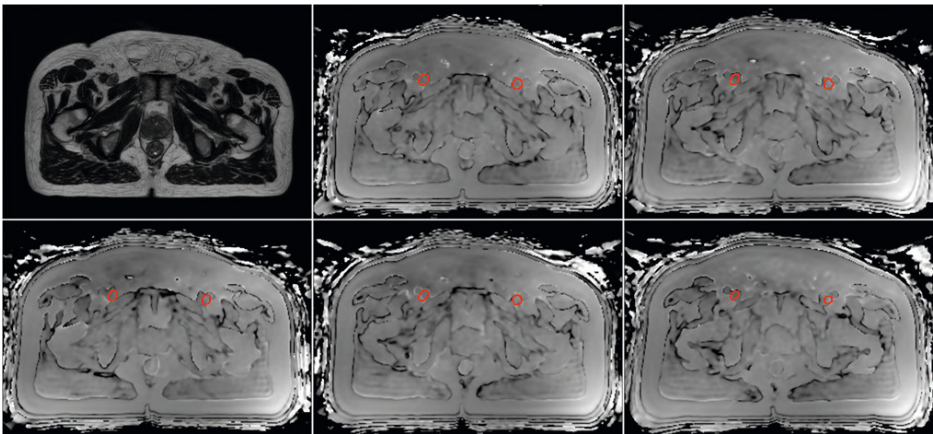
## SUPPLEMENTARY MATERIAL

### $B_1$ inhomogeneity

$B_1$  inhomogeneities influence the conversion of signal intensity to concentration values. [46] To check the variation of the  $B_1$  field, we measured  $B_1$  maps in 3 patients during 5 treatment sessions (four fractions for one patient, and five for the other two).

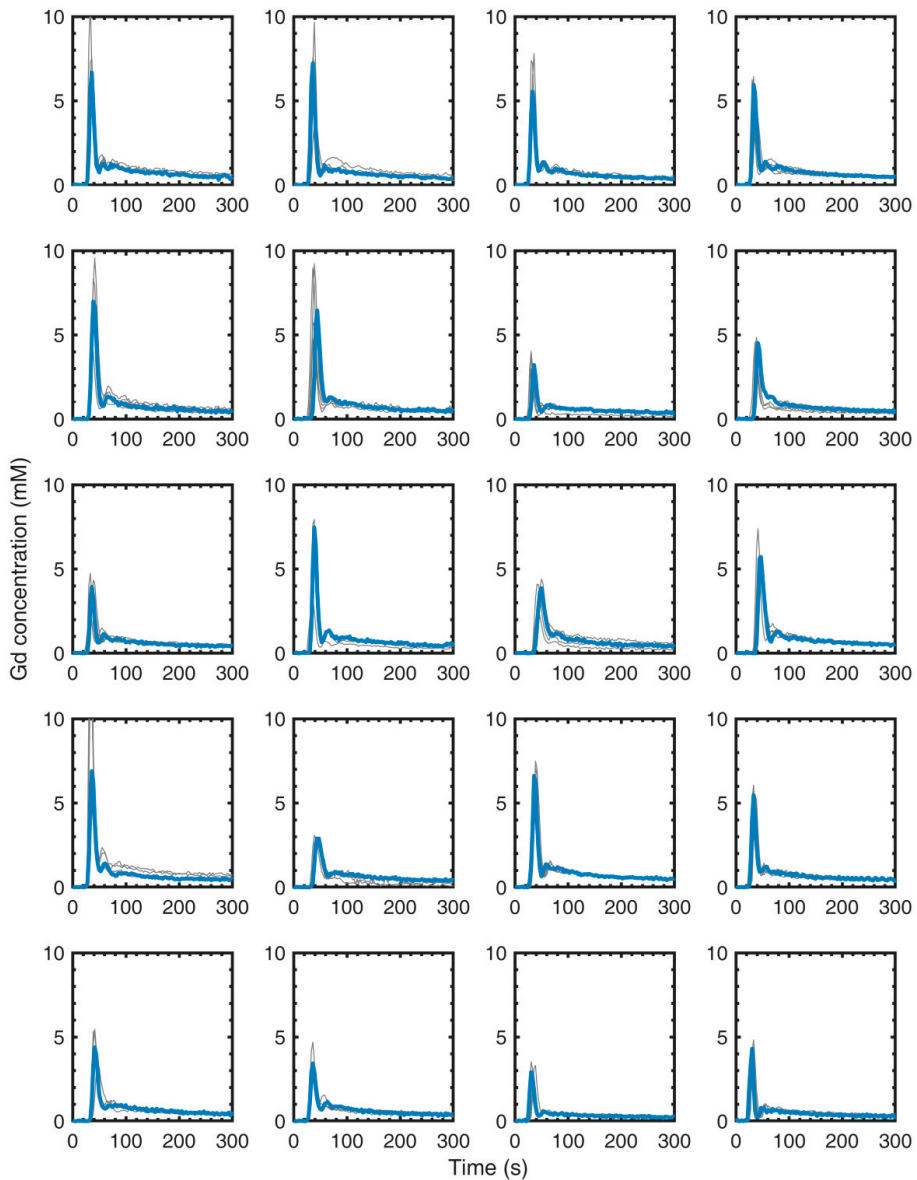
A dual TR 3D FFE scan was used with a flip angle of  $45^\circ$ . The first TR was 20 ms, the second 100 ms, and the TE was 2.7 ms. Acquired voxel size was  $4 \times 4 \times 12 \text{ mm}^3$  with a total FOV of  $430 \times 430 \times 60 \text{ mm}^3$ . The NSA was set to 3 for a total scan time of 1:29 min.

A consistent pattern was seen in all patients on all fractions with a diagonal band that includes the prostate and left iliac artery, see Figure S1. The  $B_1$  value in this band is close to 100% and drops off to the (diagonal) sides. The iliac arteries were delineated and the median values from the delineations were calculated. The average value of these patients in all fractions was  $103 \pm 5 \%$  in the left iliac artery and  $80 \pm 3 \%$  in the right iliac artery. Therefore we only used the AIF from the left iliac artery in this study.



**Figure S5.1.**  $B_1$  maps of a single patient, acquired during five separate treatment fractions. The delineations show the left and right iliac arteries.



**Arterial input functions**

**Figure S5.2.** Each panel shows the AIFs of a single patient. The AIF was determined for each DCE scan separately (5 scans for 19 patients, 4 scans for 1 patient). The median AIF, which was used for all DCE analyses of that patient, is shown in blue.





**T<sub>1ρ</sub> FOR RADIOTHERAPY  
TREATMENT RESPONSE  
MONITORING IN RECTAL  
CANCER PATIENTS:  
A PILOT STUDY**

Ernst S. Kooreman, Max Tanaka, Leon C. ter Beek, Femke P. Peters, Corrie A. M.  
Marijnen, Uulke A. van der Heide, Petra J. van Houdt

*Journal of Clinical Medicine* 11 (7), article 1998 (2022)

## ABSTRACT

Quantitative MRI has the potential to produce imaging biomarkers for the prediction of early response to radiotherapy treatment. In this pilot study, a potential imaging biomarker, the  $T_{1\rho}$  relaxation time, is assessed for this purpose. A  $T_{1\rho}$  sequence was implemented on a 1.5 T MR-linac system, a system that combines an MRI with a linear accelerator for radiation treatment. An agar phantom with concentrations of 1–4 % w/w was constructed for technical validation of the sequence. Phantom images were assessed in terms of short-term repeatability and signal-to-noise ratio. Twelve rectal cancer patients, who were treated with  $5 \times 5$  Gy, were imaged on each treatment fraction. Individual changes in the  $T_{1\rho}$  values of the gross tumor volume (GTV) showed an increase for most patients, although a paired *t*-test comparing values in the GTV from the first to the last treatment fraction showed no statistically significant difference. The phantom measurements showed excellent short-term repeatability (0.5–1.5 ms), and phantom  $T_{1\rho}$  values corresponded to the literature values.  $T_{1\rho}$  imaging was implemented successfully on the MR-linac, with a repeatability comparable to diagnostic systems, although clinical benefit in terms of treatment response monitoring remains to be demonstrated.

## INTRODUCTION

Conventional assessment of treatment response to radiation therapy involves re-evaluating the tumor using MRI or CT, and is largely based on morphological change. However, changes in the tumor microenvironment, such as changes in protein concentration already happen directly after irradiation, at a much shorter time scale than morphological changes [141]. This tumor microenvironment may be imaged using quantitative MRI (qMRI), and thus early changes in qMRI metrics could potentially be used as quantitative imaging biomarkers (QIBs) for early treatment response assessment [13]. With the introduction of MR-linacs, hybrid machines that combine an MRI with a linear accelerator for radiation treatment, the acquisition of qMRI on each treatment fraction is possible. This creates a platform where potential novel qMRI biomarkers can be searched for and evaluated with limited increase of patient burden.

One such potential biomarker is called  $T_{1\rho}$ , which stands for  $T_1$  relaxation in the rotating frame.  $T_{1\rho}$  is the relaxation time of spins while under the influence of a continuous RF pulse, which is mainly influenced by protein-water interactions and therefore sensitive to the presence protein molecules in tissue [9,142,143]. Some preliminary studies have shown potential for  $T_{1\rho}$  as a QIB for treatment response monitoring of different kinds of cancer treatment. In a preclinical study, Hectors et al. demonstrated changes only 3 days after treatment, as a result of high-intensity focused ultrasound treatment in mice with murine colon carcinoma [26]. In humans,  $T_{1\rho}$  was demonstrated to be able to distinguish between tumor and the peripheral zone in prostate cancer [144], and between tumor, fat, and fibrosis in freshly excised breast tissue [145]. In terms of treatment response monitoring, an increase in  $T_{1\rho}$  was found in healthy parotids during radiotherapy treatment of nasopharyngeal cancer patients [146].

The aim of the current study is to explore  $T_{1\rho}$  as a potential QIB for treatment response monitoring. In order to do this, a phantom was constructed and measured for technical validation purposes, and a  $T_{1\rho}$  sequence was scanned in rectal cancer patients to show clinical feasibility.

## MATERIALS AND METHODS

### MRI sequence

In this study, the Unity MR-linac (Elekta AB, Stockholm, Sweden) was used. This system integrates a 1.5 T MRI with a linear accelerator. The MRI system is based on a Philips Ingenia (Philips Healthcare, Best, The Netherlands), with adaptations made to allow for patient irradiation [1]. The gradient coils of the MRI are physically split to allow

the radiation to pass through, and the system uses an 8-channel radio-translucent phased array receive coil [16].

With  $T_{1\rho}$ , after excitation, a continuous RF pulse (the spin-lock pulse) is applied on-resonance along the magnetization vector in the transverse plane. This spin-lock pulse is a weak magnetic field (in the  $\mu\text{T}$  range) that rotates with the spins at the Larmor frequency. The spins then relax towards a new equilibrium state associated with the spin-lock pulse, with the  $T_{1\rho}$  relaxation time constant.

A  $T_{1\rho}$  sequence consisting of a spin-lock pulse cluster followed by a TSE readout was implemented [147]. The  $\Delta B_0$  and  $B_1$  insensitive spin-lock cluster as described by Witschey et al. was used [148], with a spin-lock amplitude of 400 Hz. Six images were acquired with spin lock times (TSL) of 0, 5, 10, 20, 40, and 60 ms. Due to software limitations, short gaps of 0.6 ms had to be introduced after the excitation and around the refocusing pulses, and spin lock pulses exceeding 10 ms were interrupted with these gaps every 10 ms. The  $T_{1\rho}$  pre-pulse was followed by a crusher gradient and a single-shot TSE sequence was used for readout. The FOV was  $420 \times 420 \times 104 \text{ mm}^3$ , with acquisition voxel sizes of  $3 \times 3 \times 5 \text{ mm}^3$ . Partial Fourier was used with a factor of 0.6, resulting in an echo train length of 84.

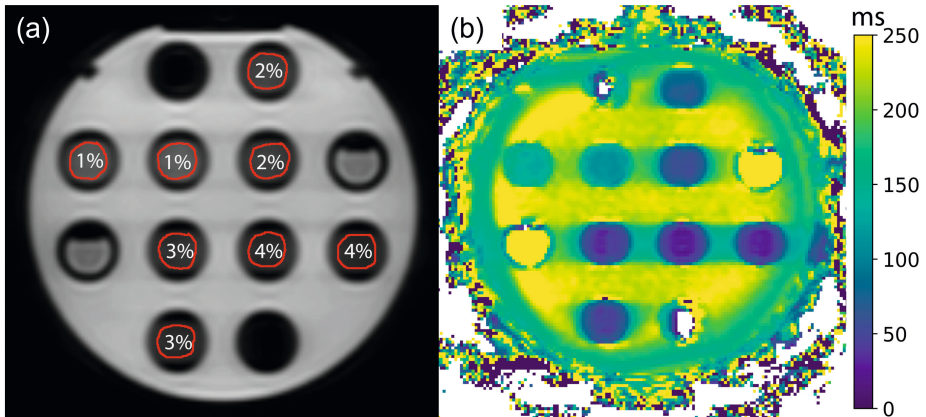
The number of acquired slices was 19 with a gap of 0.5 mm between each slice. TR/TE were 3000/4.2 ms and the scan time was 57 s per spin-lock time. The total scan time for the complete  $T_{1\rho}$  scan was 5 min and 42 s. To calculate a  $T_{1\rho}$  map, a straight line was fitted to the logarithm of the signal intensity values using weighted least squares on a voxel-by-voxel basis.

## Phantom

An agar phantom was created in a similar fashion to Buck et al [149]. Agar powder was diluted in hot distilled water to create agarose gel stock solutions with concentrations of 1, 2, 3, and 4 % w/w. The mixtures were poured into 30 mL tubes when still warm. Two tubes were created from the same stock solution for each concentration. These tubes were placed in a custom made holder filled with a copper-sulfate solution. Figure 6.1 shows the placement of the tubes. The phantom was scanned at room temperature, while placed in the iso-center of the bore.

As, to date, there are no guidelines for  $T_{1\rho}$  validation with phantoms, we followed the phantom validation framework as described in the diffusion profile of the Quantitative Imaging Biomarkers Alliance (QIBA) [41]. The agar phantom was scanned four times, consecutively, in one scan session. This allowed for the determination of short-term phantom repeatability and the signal-to-noise ratio (SNR), based on the variation in signal intensities between the four consecutive scans. For analysis, a ROI was delineated

for each tube on the center slice. As there is no gold standard for measuring  $T_{1\rho}$ , the accuracy of the sequence could not be determined. The repeatability coefficient (RC) was determined as  $RC = 2.77\sigma_w$  and the within-subject (phantom tube) coefficient of variation (wCV) as  $wCV = 100\%\sigma_w/\mu$ . Here,  $\sigma_w$  is the within-subject standard deviation and  $\mu$  the mean, both calculated from the mean values of the tube ROIs of the repeated measurements.



**Figure 6.1.** Example phantom measurement. (a) The phantom is shown as measured with a spin-lock time of 0 ms, and the percentages indicate the agar concentration in the tubes. The unmarked tubes contain distilled water. (b)  $T_{1\rho}$  map of the phantom shown in (a).

For the SNR of each tube, voxel-wise standard deviation (SD) maps and mean maps were calculated from the repeated measurements to make a temporal noise map and a temporal mean map. The SNR was then calculated for each tube by dividing the ROI means of the temporal mean map by the ROI means of the temporal noise map. The 95% confidence intervals for these SNR estimates were calculated as  $\pm 1.96 \sigma_{SNR} / \sqrt{N}$ , where  $N$  is the number of voxels in the ROI, and  $\sigma_{SNR} = SNR \sqrt{mCV^2 + nCV^2}$ . Here SNR is the SNR of the tube, and mCV and nCV are the coefficients of variation (sd/mean) of each ROI in the temporal mean map and temporal noise map, respectively [41].

To see if there are differences between tubes with the same agar concentration but a different location in the phantom, a two-sided  $t$ -test was performed per agar concentration on the voxel values inside the ROIs from the first acquisition.

## Patients

Twelve intermediate-risk rectal cancer patients who received  $5 \times 5$  Gy external beam radiation therapy during one week were included in this study. Target volumes included the mesorectum and elective lymph nodes. All patients were treated on a Unity MR-



linac, allowing them to be imaged at each treatment fraction. The study was approved by the local ethics committee and all patients gave written informed consent.

A 3D-TSE  $T_2$  weighted anatomical scan was acquired right before the  $T_{1p}$  scan for delineation purposes. Scan parameters included TR/TE = 1300/123 ms, a FOV of  $400 \times 449 \times 250 \text{ mm}^3$ , and acquisition voxel sizes of  $1.8 \times 1.8 \times 1.8 \text{ mm}^3$ . A SENSE factor of 4 was used, as well as partial Fourier with a factor of 0.6 in one direction, and 0.7 in a second direction for a scan time of 1 min 58 s.

Patient follow-up was available for 10 to 15 months after treatment. Patients were classified as complete responders or incomplete responders. Complete responders were either patients with a pathological complete response after surgery (ypT0N0), or patients with a sustained clinical complete response (cCR) during the available follow up time. cCR is defined as no or minimal residual tumor after neoadjuvant therapy, based on digital rectal examination, endoscopy, and MRI [150]. Patients with ypT1-T4 after surgery were classified as incomplete responders. Patient characteristics are given in Table 6.1.

**Table 6.1.** Patient age and T-stage, sorted by response status. Age is presented as median (range).

Complete Response ( $n = 5$ )	
Age	61 (52–72)
T-stage	
T2N0	2
T3aN0	1
T3aN2	1
T3bN1	1
Incomplete Response ( $n = 5$ )	
Age	53 (34–64)
T-stage	
T2N0	2
T3bN0	1
T3cN1	1
T4bN1	1
Response Unknown ( $n = 2$ )	
Age	61, 73
T-stage	
T3bN0	1
T3bN1	1

For all patients, the gross tumor volume (GTV), as visible on MRI, was delineated at each treatment fraction. As a control, a region of the mesorectum close to the tumor was delineated, as were both femoral heads. These delineations were propagated to

the  $T_{1p}$  maps, and median values from the delineations were used for further analysis. The RC was calculated in the femoral heads, using the  $\sigma_w$  from all treatment fractions.

## RESULTS

### Phantom measurements

An image of the  $T_{1p}$  map of the first phantom measurement is shown in Figure 6.1.

Mean values from the ROIs of the center slice, the repeatability measures, and the SNR are presented in Table 6.2. The average RC was 0.9 ms, and the average wCV was 0.5 %. All outer tubes, positioned further away from the iso-center, show consistently higher  $T_{1p}$  values than their inner counterparts. Two-sided  $t$ -tests show that these differences are significant (all  $p < 0.001$ ).

**Table 6.2.** Metrics derived from the phantom measurements: Inside indicates a tube positioned in the inner four spaces of the phantom holder and outside indicates tubes placed close to the edge of the phantom. The mean and SD were calculated from the voxels inside the ROIs in the first scan. RC = repeatability coefficient, wCV = within-tube coefficient of variation, SNR = signal-to-noise ratio.

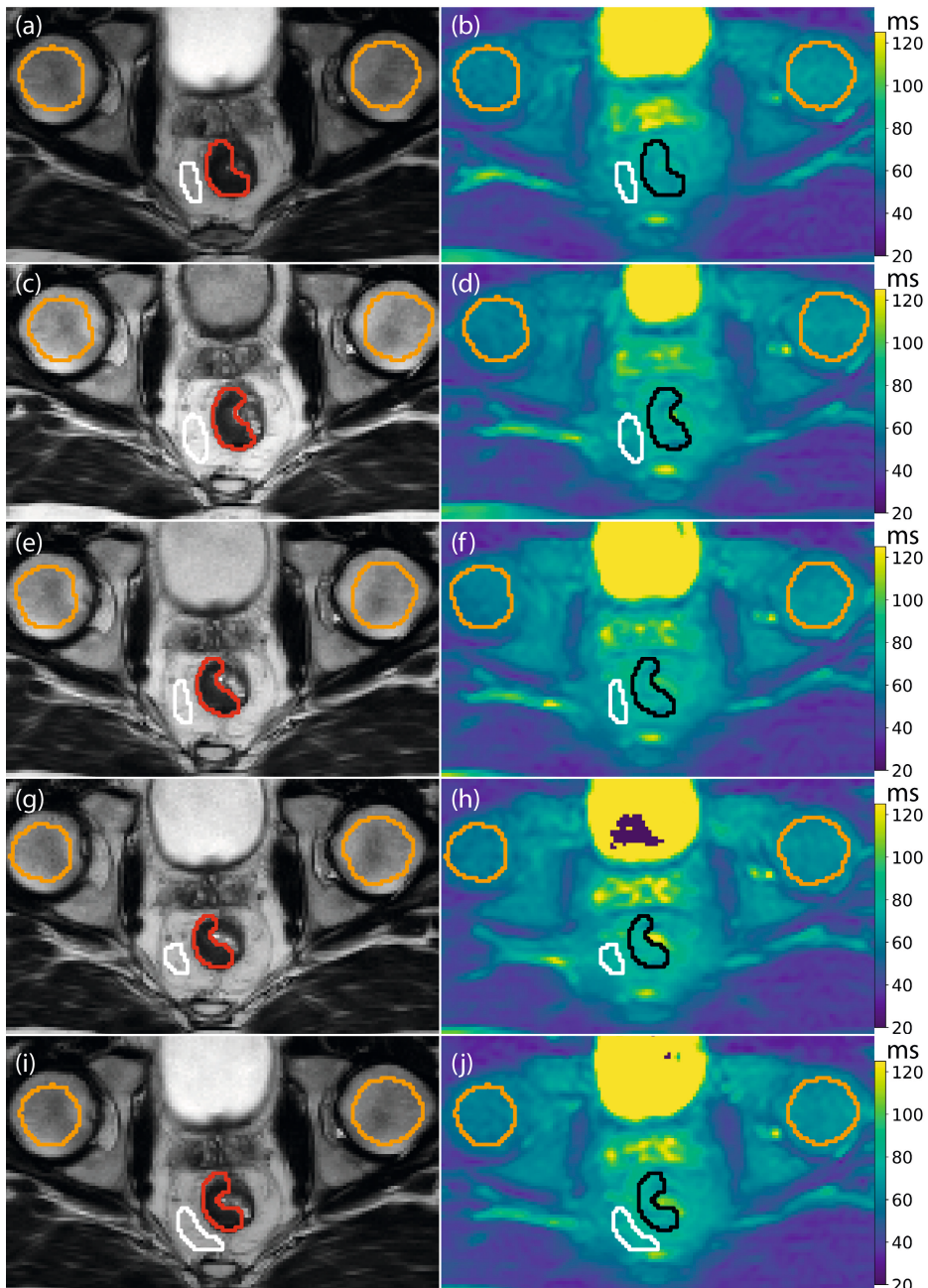
Tube	Mean $\pm$ SD (ms)	RC (ms)	wCV (%)	SNR $\pm$ 95% CI
1% inside	113 $\pm$ 2	0.5	0.1	96 $\pm$ 11
1% outside	136 $\pm$ 5	1.5	0.4	73 $\pm$ 9
2% inside	59 $\pm$ 3	0.6	0.4	83 $\pm$ 10
2% outside	76 $\pm$ 6	1.3	0.6	101 $\pm$ 11
3% inside	39 $\pm$ 5	0.8	0.7	37 $\pm$ 4
3% outside	48 $\pm$ 6	1.4	1.1	45 $\pm$ 6
4% inside	28 $\pm$ 4	0.7	0.9	36 $\pm$ 7
4% outside	32 $\pm$ 5	0.8	0.9	36 $\pm$ 5

To test for differences in  $T_{1p}$  values of the GTV between complete- and incomplete responders, a  $t$ -test was used on the data from the first fraction, and separately on the data from the last fraction. To test for differences between values from all ROIs between the first- and last fractions, a paired  $t$ -test was used.

All statistical analysis was done in R (v3.6.1), and statistical significance was assumed when  $\alpha < 0.05$ .

### Patients

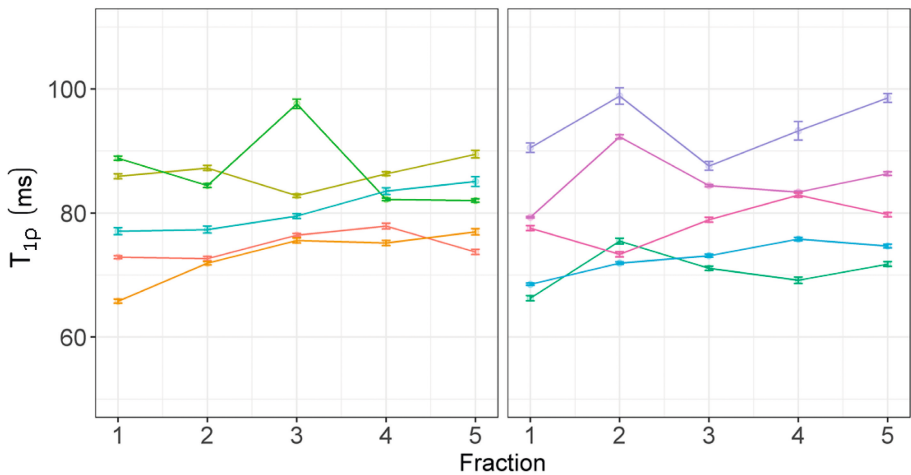
All 12 patients were successfully scanned at each treatment fraction for a total of 60 fractions. An example  $T_{1p}$  map from all fractions of a single patient is shown in Figure 6.2.



**Figure 6.2.** Example  $T_{1\rho}$  maps of a single patient. Each row corresponds to a different treatment fraction. On the left, (a, c, e, g, i) the  $T_2$ -weighted images used for delineation is shown, with the corresponding  $T_{1\rho}$  maps on the right (b, d, f, h, j). The gross tumor volume (GTV) is shown in red on the  $T_2$ -weighted image and in black on the  $T_{1\rho}$  image. The mesorectum region of interest (ROI) is shown in white, and the femoral heads are shown in orange. The  $T_{1\rho}$  maps are shown on a color scale from 20 to 120 ms.

Baseline  $T_{1\rho}$  values, measured on the first fraction before receiving the first radiation dose, were  $77 \pm 8$  ms (mean  $\pm$  SD of all patients) for the GTV,  $73 \pm 11$  ms for the mesorectum,  $64 \pm 4$  ms for the left femoral head, and  $59 \pm 5$  ms for the right femoral head. The RC in both femoral heads was 4 ms.

Figure 6.3 shows the change in  $T_{1\rho}$  values during treatment from the GTVs of the complete- (a) and incomplete responders (b). The median number of voxels in these GTVs was 567 (range: 122 – 3215). Most patients show an increasing trend in  $T_{1\rho}$  values. In addition, individual histograms of the voxels from the GTVs of all patients on all treatment fractions can be found in Supplemental Figure S6.1.



**Figure 6.3.**  $T_{1\rho}$  relaxation times from individual patients for each fraction. (a) Complete responders, (b) incomplete responders. The mean  $\pm$  standard error of the mean is shown from all voxels inside the GTV. Each line (and color) shows a different patient.

Table 6.3 contains averaged values over all patients for each ROI on each treatment fraction. The GTV is additionally split up between complete- and incomplete responders. Values in the femoral heads are relatively stable over the course of treatment. There is a difference between the left and right femoral heads, where the  $T_{1\rho}$  in the left femoral head is consistently higher. In the GTV, a slight increase in the average can be seen, and the  $T_{1\rho}$  relaxation time in the GTV is consistently higher than in the nearby mesorectum. The change in  $T_{1\rho}$  values between the first and the last treatment fraction is given in the last column, and paired  $t$ -tests show a statistically significant difference for the GTV values of the incomplete responders ( $p = 0.02$ ). A  $t$ -test for the difference in the  $T_{1\rho}$  values from the GTV between complete- and incomplete responders on the first treatment fraction showed no statistically significant difference.

**Table 6.3.** Group mean  $T_{1\rho}$  values for each ROI and treatment fraction. The values are presented in ms as standard  $\pm$  standard error of the mean (SEM). The last column shows the mean  $\pm$  SEM of the paired differences between fraction 1 and fraction 5. The  $p$ -values for the difference between the first and last fraction were calculated using a paired  $t$ -test.

ROI	Fraction 1	Fraction 2	Fraction 3	Fraction 4	Fraction 5	Fraction 5 - Fraction 1
GTV						
All ( $n = 12$ )	77 $\pm$ 2	79 $\pm$ 2	80 $\pm$ 2	80 $\pm$ 2	81 $\pm$ 2	4 $\pm$ 1 ( $p = 0.13$ )
Complete responders ( $n = 5$ )	77 $\pm$ 4	77 $\pm$ 3	81 $\pm$ 3	80 $\pm$ 2	80 $\pm$ 2	3 $\pm$ 3 ( $p = 0.44$ )
Incomplete responders ( $n = 5$ )	76 $\pm$ 4	79 $\pm$ 4	78 $\pm$ 3	80 $\pm$ 4	81 $\pm$ 5	5 $\pm$ 1 ( $p = 0.02$ )
Mesorectum	73 $\pm$ 3	72 $\pm$ 3	75 $\pm$ 3	76 $\pm$ 3	75 $\pm$ 3	2 $\pm$ 1 ( $p = 0.24$ )
Femoral head left	64 $\pm$ 1	65 $\pm$ 2	65 $\pm$ 1	66 $\pm$ 2	65 $\pm$ 1	1 $\pm$ 1 ( $p = 0.46$ )
Femoral head right	59 $\pm$ 1	59 $\pm$ 1	59 $\pm$ 1	60 $\pm$ 2	59 $\pm$ 2	1 $\pm$ 1 ( $p = 0.10$ )

## DISCUSSION

In this pilot study, we investigated  $T_{1\rho}$  as a potential QIB for treatment response monitoring. A  $\Delta B_0$  and  $B_1$  insensitive spin lock pulse was adapted slightly for use on the Unity MR-linac system. With the agar phantom measurements, it was shown that  $T_{1\rho}$  relaxation times can be quantified on the MR-linac, and feasibility of acquiring a  $T_{1\rho}$  map on each treatment fraction was demonstrated in twelve rectal cancer patients.

The first step in technical validation is to assess the accuracy and repeatability of the qMRI measurements in phantoms. The  $T_{1\rho}$  sequence was adapted by inserting short interruptions in the normally continuous spin-lock pulse. It is conceivable that interruptions influence the  $T_{1\rho}$  relaxation time, or cause a loss of spin-lock. In this sense, it is encouraging that the  $T_{1\rho}$  values of our phantom measurements decrease with an increasing agar concentration, as shown in previous studies. In their study, Buck et al. found a decrease from 55 to 29 ms in an agar phantom for concentrations of 2-4% [151], compared to 59 to 28 ms for the same concentrations in the inner tubes in this study.

To assess if such a new technique has potential for treatment response measurements, the next step is to determine whether changes during treatment occur. In this study, some increase in the  $T_{1\rho}$  values can be observed in the GTVs of individual patients, although the change between the first and last treatment fraction on a group level was not statistically significant. Although a statistically significant change was found between the  $T_{1\rho}$  values from the GTVs of incomplete responders from the first fraction versus the last fraction, this should be carefully interpreted as multiple tests were performed and the number of patients in this group is low. More patients would need to be assessed to determine if the increase in  $T_{1\rho}$  indeed holds.

Other uses of  $T_{1\rho}$  are possible. For instance,  $T_{1\rho}$  has been shown to be valuable for fibrosis detection in the liver, parotid glands, and breast [145,146,152]. Fibrosis is also an important factor in the response evaluation of rectal cancer, and especially differentiating between tumor and fibrosis remains difficult [153]. In this regard, it might be interesting to explore  $T_{1\rho}$  voxel maps.

The  $T_{1\rho}$  relaxation time depends on the strength of the spin-lock pulse and by using different strengths, the  $T_{1\rho}$  relaxation time reflects different parts of the biological microenvironment. It is probable that the 400 Hz used in this study is not optimal for treatment response purposes in rectal cancer patients.

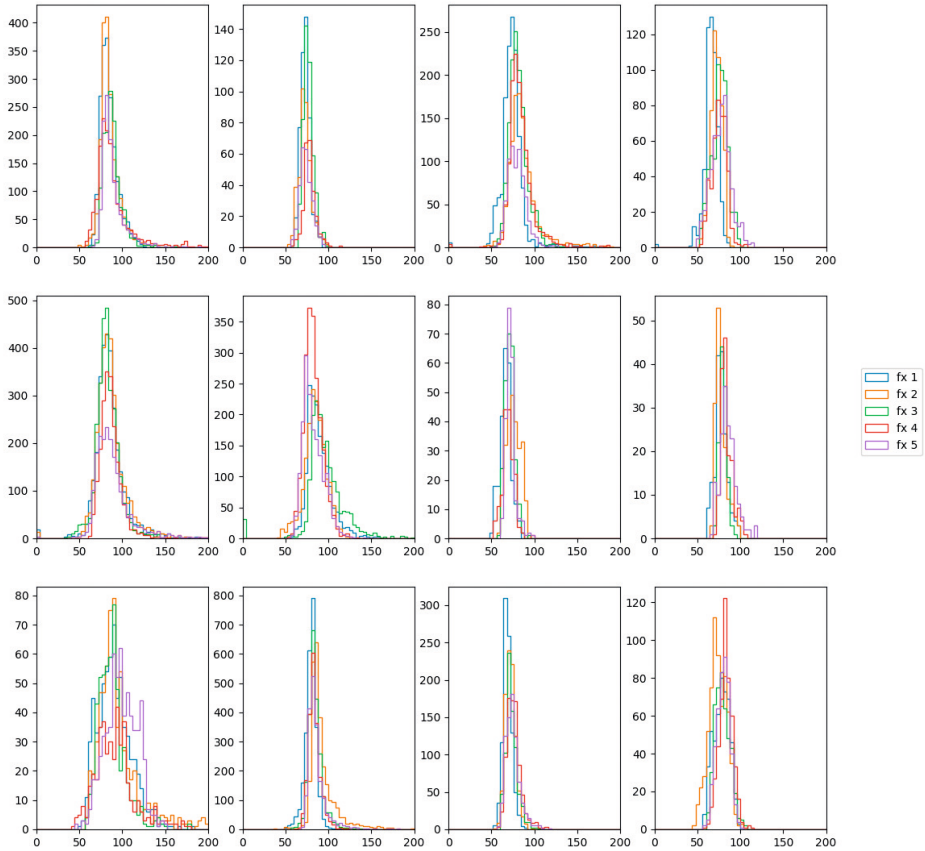
The RC from femoral heads corresponds to previously reported values, although in other tissues. In the head and neck, an RC of 2.3–5.2 ms was found for different organs [154], and a study in the prostate reported values between 13 and 23 % relative to the median value of different prostate regions (which would be 7% for both femoral heads in this study) [144]. Values from cartilage studies reported RCs of 0 – 18% [154]. The RCs reported here therefore indicate that  $T_{1\rho}$  measurements can be acquired as reliable on the Unity system as on diagnostic MRI systems.

A consistent difference between the left and the right femoral heads was found, and also between the inner and outer tubes containing the same solution of agar in the phantom. This indicates that there might be an influence of the spatial location inside the MR-linac on the  $T_{1\rho}$  measurement, similar to diffusion MRI [102]. This could be related to magnetic field inhomogeneities, and would be worth investigating in the future.

In conclusion, in this pilot study a  $T_{1\rho}$  sequence was implemented and evaluated on an MR-linac system. The phantom measurements showed high repeatability and  $T_{1\rho}$  values corresponded to literature, although spatial variation was present. Additionally,  $T_{1\rho}$  maps of rectal cancer patients were successfully acquired on each treatment fraction, with a repeatability comparable to diagnostic systems. Validation in a larger cohort is desired to establish if  $T_{1\rho}$  could be of clinical benefit in terms of treatment response monitoring.

## SUPPLEMENTARY MATERIAL

**Figure S6.1:** Histograms of the GTVs of all 12 patients and all treatment fractions are shown.









# DISCUSSION

Hybrid MR-linac systems create an opportunity for daily non-invasive treatment response monitoring using qMRI without an increase in patient burden. However, from an MRI perspective, these machines are adjusted to allow the treatment beam to pass through. Because these adjustments might influence the qMRI measurements compared to diagnostic systems, their capabilities in terms of qMRI need to be evaluated.

As mentioned in the introduction, the imaging biomarker roadmap presented by O'Connor *et al.* identifies three tracks that are followed in parallel by a biomarker on its way from inception to clinical adoption [12]. When a new machine is introduced, like the Unity MR-linac in our hospital in 2016, their impact on the acquisition of QIBs should be validated. Accuracy, repeatability, and reproducibility are influenced by the machine that the QIB is measured on and these measures need to be determined for each QIB on a new machine.

## PHANTOM MEASUREMENTS

The main value of phantom measurements in QIB studies is that they allow the assessment of the accuracy of QIBs, which is influenced the specific machine, acquisition settings and data analysis. Also, repeated phantom measurements are useful for monitoring the longitudinal stability of the system [155]. This can be done directly in qMRI phantoms, but other phantoms are also used to measure more generic characteristics of a system, such as the  $B_0$  and  $B_1$  field [74,156,157]. Monitoring these over time could prevent changes in these fields to unknowingly influence longitudinal qMRI measurements. For instance, Subashi *et al.* detected a deterioration of the SNR of their system which they were able to identify by consistently performing QA measurements in phantoms [156]. They identified table electronics as the source of additional noise, and after replacing these electronic components, the SNR was improved. In the same study, they assessed the stability of  $T_1$ ,  $T_2$ , and ADC mapping by repeatedly measuring qMRI phantoms over the course of four weeks. They found excellent repeatability with low %CV, consistent with our findings from Chapter 2.

Phantom measurements can also be used to compare acquisition sequences. In a recent study by McDonald *et al.* three DWI sequences were compared with the help of a diffusion phantom, and performance of the MR-linac was compared to the performance of a 1.5 T diagnostic system [158]. The sequences they tested included non-EPI sequences, which are important for tumor sites with a high variation in the magnetic susceptibility such as the head and neck, where transitions between air and human tissue cause large deformations in images acquired with EPI sequences.

In the studies described in Chapter 2, 3, and 6, phantoms measurements were essential. In Chapter 2, they are used as a tool for the initial validation of qMRI capabilities on the Unity MR-linac for a selection of sequences. This allowed us to demonstrate the feasibility of qMRI measurements on the Unity MR-linac, showing accurate results. For patient studies, similar phantom measurements should be performed for each specific sequence or modification to a sequence, to reassess the accuracy, repeatability and reproducibility [2]. In the study presented in Chapter 3, phantoms were used for further specific characterization of the Unity MR-linac with respect to DWI. By using a large phantom filled with a homogeneous liquid, we were able to demonstrate a large spatial variation in the ADC value when measuring away from the iso-center. Another use of a phantom is described in Chapter 6, where we implemented an acquisition sequence to measure  $T_{1\rho}$  relaxation times. By producing a phantom with different  $T_{1\rho}$  relaxation times, we could compare our measurements to previously published results, validating that we were able to accurately measure  $T_{1\rho}$ .

## TREATMENT RESPONSE MONITORING ON THE UNITY MR-LINAC

As the MR-linac enables the acquisition of daily qMRI without an increase in patient burden, this opens opportunities to measure changes with a much higher temporal resolution than before. Previously, studies assessing QIBs for radiation treatment response purposes typically measure at three timepoints: before treatment, once during treatment, and after treatment [89]. A QIB has the most value in terms of treatment response monitoring when it can predict treatment outcome before, or early during treatment. Potentially, this would enable treatment adaptation, or in a different scenario the QIB could serve as a biological target for stopping the treatment [90]. For this, QIBs that change over time because of treatment or disease progression are valuable. For instance, a recent meta-analysis showed no predictive value of pre-treatment tumor ADC values for treatment outcome in cervical cancer, but a meta-analysis of the early change in the ADC values during treatment showed potential as a predictive QIB [159,160]. This thesis contains studies that investigated if changes were measurable during treatment for IVIM and DCE parameters in prostate cancer patients (Chapters 4 and 5), and  $T_{1\rho}$  in rectal cancer patients (Chapter 6). Chapter 4 contains encouraging results, where no change was measured in the D values of the healthy prostate, but D values increased in the tumor. However, while results from all these studies indicate that changes are measurable, this is on the group level. To move further towards personalized treatment, it is important to reduce the noise in the measured QIBs, either by improvements of acquisition sequences, or in post-processing.

## TECHNICAL CHALLENGES

### Signal to noise ratio

One of the most important aspects that influence the ability to measure changes over time is the SNR of the system, which influences the random variation of qMRI parameters. In general, the visual quality of qMRI maps is lower than qualitative maps, because multiple images need to be acquired in a short time. To compensate, this usually necessitates larger voxels, producing unclear images. Additionally, motion of the patient between multiple images complicates fitting of the quantitative models especially when done in a voxel-by-voxel manner to create quantitative maps. Therefore, often some kind of image registration is required before fitting the model. The Unity MR-linac differs substantially from diagnostic MRI systems, and as shown in Chapters 2-5 the SNR is an issue that keeps resurfacing.

In Chapter 2, we were unable to meet all the requirements of the QIBA tests for the DWI phantom, specifically requirements related to SNR [70]. When adjusting the sequence to enable higher SNR using larger voxels and a restricted maximum b-value, the requirements were met. In Chapter 3, considerations for the recommendations were also based largely on SNR limitations. Here, mainly the effect of the software-limited performance of the gradient system, which causes an increased TE in our DWI sequence underlies the recommendations regarding limiting the maximum b-value.

As described in the introduction, the receive coil consists of an 8-channel phased array coil, which is rigid. For certain tumor sites, such as head and neck, it is difficult to minimize the distance between the coil and the target. Improvements of the receive coils are challenging because the treatment beam imposes additional constraints to the design. Recently, Zijlema *et al.* developed a 32-channel flexible receive coil using high-impedance coil loops [161]. They demonstrated an improved SNR performance over the commercially available receive coil, especially with accelerated imaging.

For DWI, the gradient performance of the system is an important factor. The maximum gradient strength and slew rate of the Unity MR-linac system are restricted to 15 mT/m and 65 T/m/s, respectively. This is low compared to the diagnostic system it is based on, which has a maximum gradient strength of 40 mT/m and a maximum slew rate of 200 T/m/s. For DWI this means that to achieve the same b-value, the diffusion sensitizing gradients need to be extended, increasing the TE and reducing the SNR. Therefore, we recommend limiting the maximum b-value to 500 s/mm<sup>2</sup> for ADC mapping on the Unity system (Chapter 3).

The SNR of the system influences the repeatability, which is important for longitudinal monitoring of patients such as is the case for treatment response monitoring. In Chapter

4, the RC of IVIM parameters was determined, however it should be noted that the patients were not scanned twice before treatment in a test-retest setting, but rather the values of the first and second fraction of treatment were used. The RCs were high for the perfusion parameters of the IVIM model, meaning that the repeatability of these parameters is poor. This is in accordance with other studies, which also showed very high RCs for  $f$  and especially for  $D^*$ . In a repeatability study on a 3 T system, Sun *et al.* reported a short-term %RC of 88% for  $D^*$  in the prostate [162]. In a comparison between fit algorithms, Gurney-Champion *et al.* found %RCs for  $D^*$  of 67% for the best performing algorithm up to 44.2% of the worst performing one in pancreatic tumors [163].

One of the reasons for the low repeatability could be that the IVIM sequence used in Chapter 4 was not optimized yet for measuring the perfusion parameters on the MR-linac. Our sequence contained only one specific point to measure  $D^*$  (the  $b$ -value of 30 s/mm<sup>2</sup>), while studies looking to optimize the  $b$ -value distribution for IVIM acquisitions typically include more  $b$ -values [164,165]. By measuring only four  $b$ -values, the calculation of IVIM parameters was straight forward and fast as no optimizer was needed. Currently, efforts are undertaken by Wetscherek *et al.* to specifically optimize IVIM sequences for the Unity MR-linac for multiple target sites [166]. This abstract takes a data-driven approach, where Monte-Carlo simulations are used to find a set of  $b$ -values and averages that minimize the average relative error in the  $D$ ,  $f$ , and  $D^*$  parameter maps. The abstract provides optimal acquisition parameters for a measurement time of 5 minutes for head and neck, brain, prostate, and rectum tumors. By using ROI averages for the fit, a better coefficient of variation is achieved than for voxel wise fits.

Novel deep learning based algorithms are promising techniques that can reduce the RC in quantitative parameter maps. In a study by Kaandorp *et al.*, a neural network outperformed conventional least squares based and Bayesian algorithms in terms of accuracy and repeatability of the IVIM parameters in a cohort of pancreatic cancer patients, although accuracy was only assessed on simulated data [167].

Noisy parameter maps also played a role in Chapter 5. While we did find some statistically significant longitudinal correlations between IVIM and DCE parameters, these correlations were low and mainly present in the transition zone, the largest ROI. The correlations could be reduced by the noisiness of the parameter maps on the patient level, which has the least influence on the largest ROI as median values were used. Future efforts should include improving the repeatability of both IVIM and DCE on the Unity MR-linac. The improvements mentioned above, such as a better receive coil, improved fitting methods, and optimized acquisition parameters would also be beneficial for both techniques.

### **Spatial inhomogeneity**

An aspect of ADC mapping on the Unity system that deserves further attention is the spatial dependence of the ADC value. The current recommendation from Chapter 3 includes the advice to be careful when scanning away from the iso-center in the xy-plane. Within a radius of 7 cm the variation of the ADC that we found was < 5% of the mean value in the iso-center, while outside of that the ADC value can steeply increase up to 600% of the value in the iso-center. This might be difficult to adhere to in practice for certain tumor positions, for instance in liver cancer patients. As the pattern seems consistent over multiple measurements on multiple machines, longitudinal measurements might still be useful with careful patient positioning. We hypothesize that this spatial variation is caused by eddy currents which are increased in the Unity MR-linac because of the split gradient coil. The influence of eddy currents on DWI using EPI sequences and correction methods have been proposed before [168–170]. However, these techniques require the acquisition of additional scans, advanced sequence design, or the use of field probes. As eddy current properties are system specific, such measurements would have to be performed on each system separately. Additionally, extensive knowledge about reconstruction algorithms is needed to implement these corrections.

## **TUMOR SITES**

Clinically, our results suggest that there is a limited value for prostate cancer patients in terms of treatment response monitoring using qMRI, because the measured changes during treatment are low compared to the RCs. The largest change in D from our mixed effects model over 20 fractions on the group level was found for patients with a high ISUP score and constituted an increase of about 14 % compared to the mean value on the first fraction. This corresponds to the 14 % increase Foltz *et al.* found in the tumor after six weeks of treatment [99], however it is well below our estimate of the RC for the D, which was about 36 % of the mean value from the tumors on the first fraction. Another factor limiting the value in the case of prostate cancer, is the time frame in which the QIBs are measured. Currently, most low- and intermediate-risk prostate cancer patients in our clinic are treated with a schedule of  $5 \times 7.25$  Gy and trials investigating ultra-hypofractionation with only a single fraction are currently recruiting [171]. This is too short for treatment adaption based on QIBs although monitoring can also take place after treatment.

Prostate cancer patients were selected as first subjects for qMRI on the MR-linac for multiple reasons. Multiparametric protocols, including DWI and DCE are well established [172,173], the organ is positioned in the center of the body and close to the iso-center of the system, the movement is relatively small, and at our center a large

number of these patients were eligible for treatment on the MR-linac. This allows for the assessment of important aspects of QIBs that can reveal challenges also applicable to other tumor sites. For instance, the repeatability we found can be expected to be worse in other tumor sites were e.g. movement due to breathing or bowel motion, or tumor regression can complicate things.

Other groups have investigated changes in QIBs during radiotherapy in different organ sites. Yang *et al.* measured three sarcoma and three head and neck cancer patients on a 0.35 T MRI integrated with three cobalt-60 heads [17,18]. They analyzed their patients individually, and report varying results. Most notably, they found a drop of 33 % in the ADC value of a head and neck cancer patient, who was later shown by biopsy to have a necrotic core. On the Unity MR-linac, Thorwarth *et al.* showed preliminary results in a single head and neck cancer patient, where they found a doubling of the ADC value during treatment (from about 0.75 to  $1.5 \times 10^{-3} \text{ mm}^2/\text{s}$ ) [174]. Such large changes on an individual basis are encouraging results for treatment response potential in these tumor sites.

## OTHER QUANTITATIVE TECHNIQUES

Chan *et al.* showed the feasibility of chemical exchange saturation transfer (CEST) MRI on the Unity MR-linac [175], validating results in a phantom, and measuring differences between high- and low-grade tumors of central nervous system (CNS) cancer patients. They demonstrated changes in CEST parameters on the cohort level in 54 patients, but also in individual patients. The same group assessed DWI in CNS in 24 patients on the MR-linac, and compared results to measurements of 20 patients (six patients were scanned on both systems) on a 1.5 T diagnostic system [123]. They found a negative bias in ADC values from the MR-linac, which increased with an increasing ADC value. In the 20 patients scanned on the MR-linac, they were able to identify individual patients based on an increase, decrease, both, or no change of ADC values in the GTV. The repeatability of the MR-linac with the 8-channel receive array they found was similar to a diagnostic 1.5 T system using a 16-channel birdcage coil, and better than we found in the prostate (Chapter 4). For white matter, the wSD was  $0.0097 \times 10^{-3} \text{ mm}^2/\text{s}$  compared to  $0.033 \times 10^{-3} \text{ mm}^2/\text{s}$  in the prostate [123]. Having a better repeatability in CNS cancer patients is encouraging for QIBs for treatment response monitoring in this group of patients.

Another interesting technique called MR fingerprinting [176], which can produce  $T_1$  and  $T_2$  maps in an extremely short amount of time (about 5 seconds for a single slice) has been shown to be feasible on the MRIdian system and the Unity system [177–179].



## BIOLOGICAL LINK

Biological links for the ADC, DCE, and IVIM parameters are relatively well established [23]. The ADC corresponds to cell density and DCE and IVIM parameters correspond to perfusion characteristics. Currently, a link between the  $T_{1\rho}$  relaxation time and radiation induced changes is not clear.  $T_{1\rho}$  has been investigated in relation to measuring damage in OARs in head and neck cancer, specifically in the parotid glands which are sensitive to radiation [146]. In this study they found an increase in  $T_{1\rho}$  values of the parotids during radiation treatment. Citing animal studies, the authors hypothesized that an early build-up of collagen, eventually leading to fibrosis could be the underlying biological mechanism. Fibrosis is also common in rectal cancer, and differentiating between fibrosis and residual tumor using conventional MRI techniques is difficult [153]. It is possible that the early onset of the formation of fibrosis causes an increase in  $T_{1\rho}$ . Establishing such a biological link would be an important next step in the development of  $T_{1\rho}$  as a QIB.

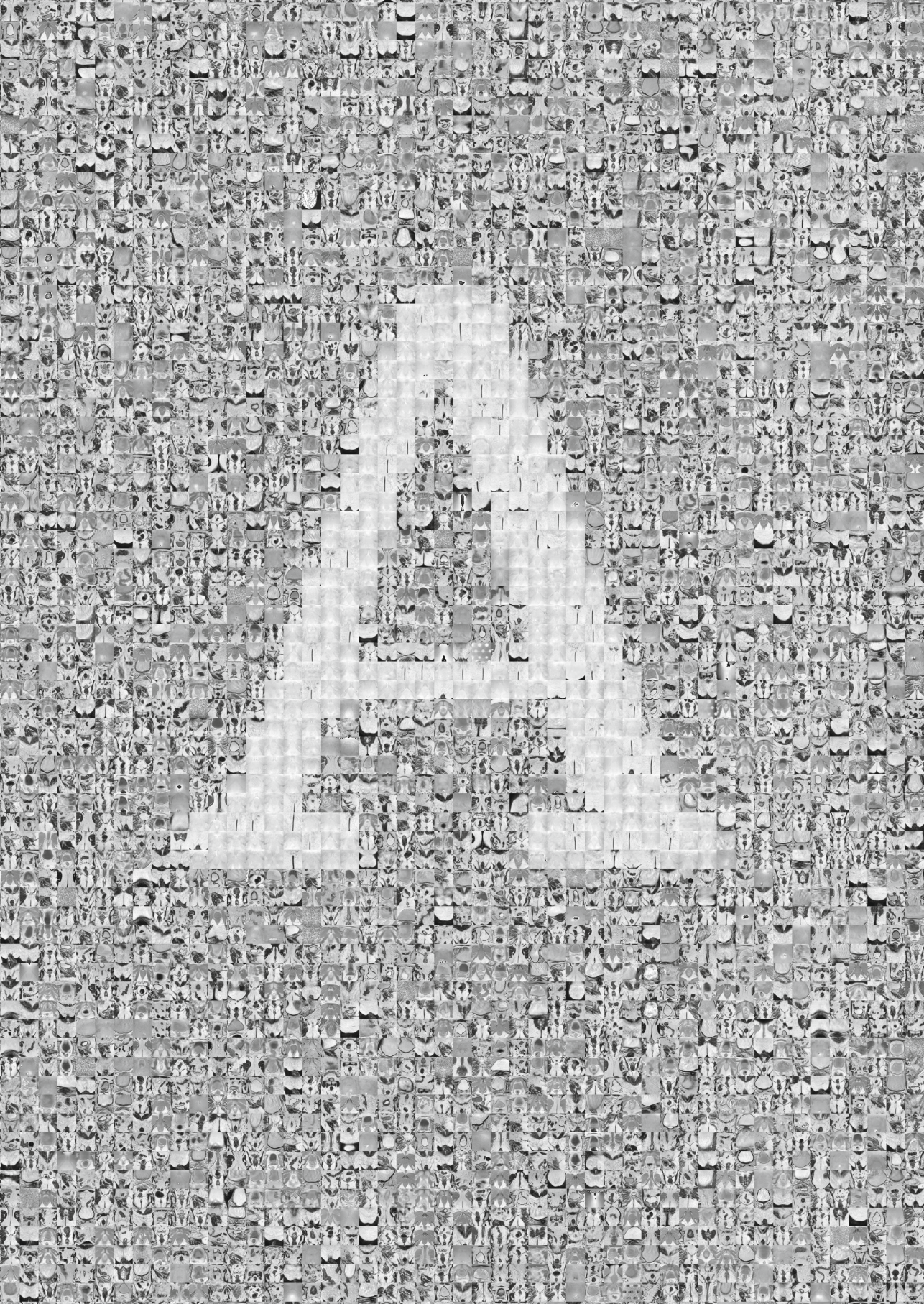
## DATA ANALYSIS

A challenging aspect of the daily acquisition of qMRI scans is the large amount of data that is generated for each patient. For instance, 43 patients were included in the study described in Chapter 3, who received 20 fractions of radiotherapy. These 43 patients would get 860 scanning sessions in total, with multiple scans per session. In our workflow, these scans were all registered to the first fraction, and delineations for each patient on the first fraction were needed. The IVIM scans consisted of 4 b-values, and scans were checked for motion between b-values. This is feasible to do in a research setting but would require substantial additional work on the daily in the clinic. A proper infrastructure is needed to handle this amount of data, and automation tools should be developed to lower the workload and improve the reproducibility. Auto-registration, auto-segmentation, and automated fitting procedures should be developed before the use of daily QIB informed therapy can become reality.

The statistical analysis of longitudinal imaging data should be considered carefully, as multiple data points from the same patient are correlated. Depending on the research question, proper statistical tests should be selected to account for this, such as the mixed effects models used in Chapter 4 or the repeated measures correlation in Chapter 5 of this thesis. As stated in Simpson's paradox, ignoring clustering when determining correlations can potentially lead to correlation coefficients with an opposite sign as when you take the clustering into account [136].

## **CLINICAL IMPLEMENTATION OF QIBs**

The adaptation of qMRI and QIBs into the clinic is a long and slow process with many challenges, and groups such as QIBA, QIN, AAPM, and EIBALL have been formed with the goal to move QIB research further towards clinical adoption by specifically focusing on education, standardization of acquisition and software, and improvement of accuracy, repeatability, and reproducibility [12,19,62,63,89,180–183]. The MR-linac provides an opportunity to simplify some of these difficulties by providing an identical platform to multiple centers, thereby allowing sharing of MRI protocols and therefore easier pooling of patients.



# APPENDICES

REFERENCES

SUMMARY

SAMENVATTING

LIST OF PUBLICATIONS

ACKNOWLEDGEMENTS

CURRICULUM VITAE

## REFERENCES

- [1] Raaymakers BW, Lagendijk JJW, Overweg J, Kok JGM, Raaijmakers AJE, Kerkhof EM, et al. Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept. *Phys Med Biol* 2009;54:N229–37. <https://doi.org/10.1088/0031-9155/54/12/N01>.
- [2] Mutic S, Dempsey JF. The ViewRay System: Magnetic Resonance-Guided and Controlled Radiotherapy. *Semin Radiat Oncol* 2014;24:196–9. <https://doi.org/10.1016/j.semradonc.2014.02.008>.
- [3] Tofts PS. *Concepts: Measurement and MR. Quant. MRI Brain*, John Wiley & Sons, Ltd; 2003, p. 1–15. <https://doi.org/https://doi.org/10.1002/0470869526.ch1>.
- [4] Stejskal EO, Tanner JE. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *J Chem Phys* 1965;42:288–92. <https://doi.org/10.1063/1.1695690>.
- [5] Le Bihan D, Breton E, Lallemand D, Grenier P, Cabanis E, Laval-Jeantet M. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology* 1986;161:401–7. <https://doi.org/10.1148/radiology.161.2.3763909>.
- [6] Bernstein MA, King KF, Zhou XJ. *Handbook of MRI Pulse Sequences*, Elsevier Academic Press; 2004.
- [7] Le Bihan D, Breton E, Lallemand D, Aubin ML, Vignaud J, Laval-Jeantet M. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* 1988;168:497–505. <https://doi.org/10.1148/radiology.168.2.3393671>.
- [8] Bloch F. Nuclear induction. *Phys Rev* 1946;70:460–74. <https://doi.org/10.1103/PhysRev.70.460>.
- [9] Gilani IA, Sepponen R. Quantitative rotating frame relaxometry methods in MRI. *NMR Biomed* 2016;29:841–61. <https://doi.org/10.1002/nbm.3518>.
- [10] Tofts PS, Brix G, Buckley DL, Evelhoch JL, Henderson E, Knopp M V, et al. Contrast-Enhanced T1-Weighted MRI of a Diffusible Tracer: Standardized Quantities and Symbols 1999;232:223–32.
- [11] Kessler LG, Barnhart HX, Buckler AJ, Choudhury KR, Kondratovich M V., Toledano A, et al. The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. *Stat Methods Med Res* 2015;24:9–26. <https://doi.org/10.1177/0962280214537333>.
- [12] O'Connor JPB, Aboagye EO, Adams JE, Aerts HJWL, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol* 2017;14:169–86. <https://doi.org/10.1038/nrclinonc.2016.162>.
- [13] van Houdt PJ, Yang Y, van der Heide UA. Quantitative Magnetic Resonance Imaging for Biological Image-Guided Adaptive Radiotherapy. *Front Oncol* 2021;10:1–9. <https://doi.org/10.3389/fonc.2020.615643>.
- [14] Keenan KE, Ainslie M, Barker AJ, Boss MA, Cecil KM, Charles C, et al. Quantitative magnetic resonance imaging phantoms: A review and the need for a system phantom. *Magn Reson Med* 2017;00. <https://doi.org/10.1002/mrm.26982>.
- [15] Tijssen RHN, Philippens MEP, Paulson ES, Glitzner M, Chugh B, Wetscherek A, et al. MRI commissioning of 1.5T MR-linac systems – a multi-institutional study. *Radiother Oncol* 2019;132:114–20. <https://doi.org/10.1016/j.radonc.2018.12.011>.

- [16] Hoogcarspel SJ, Zijlema SE, Tijssen RHN, Kerkmeijer LGW, Jürgenliemk-Schulz IM, Lagendijk JJW, et al. Characterization of the first RF coil dedicated to 1.5 T MR guided radiotherapy. *Phys Med Biol* 2018;63:025014. <https://doi.org/10.1088/1361-6560/aaa303>.
- [17] Yang Y, Cao M, Sheng K, Gao Y, Chen A, Kamrava M, et al. Longitudinal diffusion MRI for treatment response assessment: Preliminary experience using an MRI-guided tri-cobalt 60 radiotherapy system. *Med Phys* 2016;43:1369–73. <https://doi.org/10.1118/1.4942381>.
- [18] Gao Y, Han F, Zhou Z, Cao M, Kaprealian T, Kamrava M, et al. Distortion-free diffusion MRI using an MRI-guided Tri-Cobalt 60 radiotherapy system: Sequence verification and preliminary clinical experience. *Med Phys* 2017;44:5357–66. <https://doi.org/10.1002/mp.12465>.
- [19] Shukla-Dave A, Obuchowski NA, Chenevert TL, Jambawalikar S, Schwartz LH, Malyarenko D, et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *J Magn Reson Imaging* 2019;49:e101–21. <https://doi.org/10.1002/jmri.26518>.
- [20] van Houdt PJ, Saeed H, Thorwarth D, Fuller CD, Hall WA, McDonald BA, et al. Integration of quantitative imaging biomarkers in clinical trials for MR-guided radiotherapy: Conceptual guidance for multicentre studies from the MR-Linac Consortium Imaging Biomarker Working Group. *Eur J Cancer* 2021;153:64–71. <https://doi.org/10.1016/j.ejca.2021.04.041>.
- [21] Kerkmeijer LGW, Fuller CD, Verkooijen HM, Verheij M, Choudhury A, Harrington KJ, et al. The MRI-linear accelerator consortium: Evidence-based clinical introduction of an innovation in radiation oncology connecting researchers, methodology, data collection, quality assurance, and technical development. *Front Oncol* 2016;6:1–6. <https://doi.org/10.3389/fonc.2016.00215>.
- [22] de Mol van Otterloo SR, Christodouleas JP, Blezer ELA, Akhlat H, Brown K, Choudhury A, et al. The MOMENTUM Study: An International Registry for the Evidence-Based Introduction of MR-Guided Adaptive Therapy. *Front Oncol* 2020;10. <https://doi.org/10.3389/fonc.2020.01328>.
- [23] Federau C. Intravoxel incoherent motion MRI as a means to measure in vivo perfusion: A review of the evidence. *NMR Biomed* 2017;30:1–15. <https://doi.org/10.1002/nbm.3780>.
- [24] Singh A, Reddy D, Haris M, Cai K, Rajender Reddy K, Hariharan H, et al. T<sub>1ρ</sub> MRI of healthy and fibrotic human livers at 1.5 T. *J Transl Med* 2015;13:1–7. <https://doi.org/10.1186/s12967-015-0648-0>.
- [25] Rauscher I, Eiber M, Ganter C, Martirosian P, Safi W, Umgelter A, et al. Evaluation of T<sub>1ρ</sub> as a potential MR biomarker for liver cirrhosis: Comparison of healthy control subjects and patients with liver cirrhosis. *Eur J Radiol* 2014;83:900–4. <https://doi.org/10.1016/j.ejrad.2014.02.017>.
- [26] Hectors SJCG, Moonen RPM, Strijkers GJ, Nicolay K. T<sub>1ρ</sub> mapping for the evaluation of high intensity focused ultrasound tumor treatment. *Magn Reson Med* 2015;73:1593–601. <https://doi.org/10.1002/mrm.25269>.
- [27] Zahra MA, Hollingsworth KG, Sala E, Lomas DJ, Tan LT. Dynamic contrast-enhanced MRI as a predictor of tumour response to radiotherapy. *The Lancet Oncolgy* 2007;8:63–74. [https://doi.org/10.1016/S1470-2045\(06\)71012-9](https://doi.org/10.1016/S1470-2045(06)71012-9).
- [28] Iima M, Le Bihan D. Clinical Intravoxel Incoherent Motion and Diffusion MR Imaging: Past, Present, and Future. *Radiology* 2016;278:13–32. <https://doi.org/10.1148/radiol.2015150244>.
- [29] Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS Prostate Imaging-Reporting and Data System: 2015, Version 2. *Eur Urol* 2016;69:16–40. <https://doi.org/10.1016/j.eururo.2015.08.052>.

- [30] Morris EA, Comstock CE, Lee CH. ACR BI-RADS® Magnetic Resonance Imaging. ACR BI-RADS® Atlas, Breast Imaging Report. Data Syst., Reston, VA: American College of Radiology; 2013.
- [31] Raunig DL, McShane LM, Pennello G, Gatsonis C, Carson PL, Voyvodic JT, et al. Quantitative imaging biomarkers: A review of statistical methods for technical performance assessment. *Stat Methods Med Res* 2015;24:27–67. <https://doi.org/10.1177/0962280214537344>.
- [32] Decker G, Mürtz P, Gieseke J, Träber F, Block W, Sprinkart AM, et al. Intensity-modulated radiotherapy of the prostate: Dynamic ADC monitoring by DWI at 3.0 T. *Radiother Oncol* 2014;113:115–20. <https://doi.org/10.1016/j.radonc.2014.07.016>.
- [33] Ho JC, Allen PK, Bhosale PR, Rauch GM, Fuller CD, Mohamed ASR, et al. Diffusion-Weighted Magnetic Resonance Imaging as a Predictor of Outcome in Cervical Cancer After Chemoradiation. *Int J Radiat Oncol Biol Phys* 2017;97:546–53. <https://doi.org/10.1016/j.ijrobp.2016.11.015>.
- [34] Das S, Chandramohan A, Rami Reddy JK, Mukhopadhyay S, Kumar RM, Isiah R, et al. Role of conventional and diffusion weighted MRI in predicting treatment response after low dose radiation and chemotherapy in locally advanced carcinoma cervix. *Radiother Oncol* 2015;117:288–93. <https://doi.org/10.1016/j.radonc.2015.10.006>.
- [35] Wang L, Liu L, Han C, Liu S, Tian H, Li Z, et al. The diffusion-weighted magnetic resonance imaging (DWI) predicts the early response of esophageal squamous cell carcinoma to concurrent chemoradiotherapy. *Radiother Oncol* 2016;121:246–51. <https://doi.org/10.1016/j.radonc.2016.10.021>.
- [36] Mahmood F, Johannesen HH, Geertsen P, Hansen RH. Repeated diffusion MRI reveals earliest time point for stratification of radiotherapy response in brain metastases. *Phys Med Biol* 2017;62:2990–3002. <https://doi.org/10.1088/1361-6560/aa5249>.
- [37] Datta A, Aznar MC, Dubec M, Parker GJM, O'Connor JPB. Delivering Functional Imaging on the MRI-Linac: Current Challenges and Potential Solutions. *Clin Oncol* 2018;30:702–10. <https://doi.org/10.1016/j.clon.2018.08.005>.
- [38] Fallone BG. The Rotating Biplanar Linac-Magnetic Resonance Imaging System. *Semin Radiat Oncol* 2014;24:200–2. <https://doi.org/10.1016/j.semradonc.2014.02.011>.
- [39] Keall PJ, Barton M, Crozier S. The Australian Magnetic Resonance Imaging-Linac Program. *Semin Radiat Oncol* 2014;24:203–6. <https://doi.org/10.1016/j.semradonc.2014.02.015>.
- [40] Wojcieszynski AP, Rosenberg SA, Brower J V, Hullett CR, Geurts MW, Labby ZE, et al. Gadoxetate for direct tumor therapy and tracking with real-time MRI-guided stereotactic body radiation therapy of the liver. *Radiother Oncol* 2016;118:416–8. <https://doi.org/10.1016/j.radonc.2015.10.024>.
- [41] Perfusion, Diffusion and Flow-MRI Biomarker Committee. Diffusion-Weighted Magnetic Resonance Imaging (DWI), Quantitative Imaging Biomarkers Alliance. Initial Draft. QIBA 2017.
- [42] DCE MRI Technical Committee. DCE MRI Quantification Profile, Quantitative Imaging Biomarkers Alliance. Version 1.0. Reviewed Draft. QIBA 2012.
- [43] Fram EK, Herfkens RJ, Johnson GA, Glover GH, Karis JP, Shimakawa A, et al. Rapid calculation of T1 using variable flip angle gradient refocused imaging. *Magn Reson Imaging* 1987;5:201–8. [https://doi.org/10.1016/0730-725x\(87\)90021-x](https://doi.org/10.1016/0730-725x(87)90021-x).
- [44] Deoni SCL, Rutt BK, Peters TM. Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state. *Magn Reson Med* 2003;49:515–26. <https://doi.org/10.1002/mrm.10407>.

- [45] Milford D, Rosbach N, Bendszus M, Heiland S. Mono-exponential fitting in T2-relaxometry: Relevance of offset and first echo. *PLoS One* 2015;10:1–13. <https://doi.org/10.1371/journal.pone.0145255>.
- [46] Schabel MC, Parker DL. Uncertainty and bias in contrast concentration measurements using spoiled gradient echo pulse sequences. *Phys Med Biol* 2008;53:2345–73. <https://doi.org/10.1088/0031-9155/53/9/010>.
- [47] Bane O, Hectors SJ, Wagner M, Arlinghaus LL, Aryal MP, Cao Y, et al. Accuracy, repeatability, and interplatform reproducibility of T1 quantification methods used for DCE-MRI: Results from a multicenter phantom study. *Magn Reson Med* 2017;00:1–12. <https://doi.org/10.1002/mrm.26903>.
- [48] Stikov N, Boudreau M, Levesque IR, Tardif CL, Barral JK, Pike GB. On the accuracy of T1 mapping: Searching for common ground. *Magn Reson Med* 2015;73:514–22. <https://doi.org/10.1002/mrm.25135>.
- [49] Murase K. Efficient Method for Calculating Kinetic Parameters Using T1-Weighted Dynamic Contrast-Enhanced Magnetic Resonance Imaging. *Magn Reson Med* 2004;51:858–62. <https://doi.org/10.1002/mrm.20022>.
- [50] Parker GJM, Roberts C, Macdonald A, Buonaccorsi GA, Cheung S, Buckley DL, et al. Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI. *Magn Reson Med* 2006;56:993–1000. <https://doi.org/10.1002/mrm.21066>.
- [51] Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8:135–60. <https://doi.org/10.1177/096228029900800204>.
- [52] Bland JM, Altman DG. Statistics Notes: Measurement error proportional to the mean. *BMJ* 1996;313:106–106.
- [53] Andersen C, Tågehøj Jensen F. Precision, accuracy, and image plane uniformity in NMR relaxation time imaging on a 1.5 T whole-body MR imaging system. *Magn Reson Imaging* 1994;12:775–84. [https://doi.org/10.1016/0730-725X\(94\)92202-0](https://doi.org/10.1016/0730-725X(94)92202-0).
- [54] Graves MJ, Emmens D, Lejay H, Hariharan H, Polzin J, Lomas DJ. T2 and T2\* Quantification Using Optimal B1 Image Reconstruction for Multicoil Arrays. *J Magn Reson Imaging* 2008;28:278–81. <https://doi.org/10.1002/jmri.21420>.
- [55] Liney GP, Knowles AJ, Manton DJ, Turnbull LW, Blackband SJ, Horsman A. Comparison of conventional single echo and multi-echo sequences with a Fast Spin-Echo sequence for quantitative T2 mapping: Application to the prostate. *J Magn Reson Imaging* 1996;6:603–7. <https://doi.org/10.1002/jmri.1880060408>.
- [56] de Bazelaire CMJ, Duhamel GD, Rofsky NM, Alsop DC. MR Imaging Relaxation Times of Abdominal and Pelvic Tissues Measured in Vivo at 3.0 T: Preliminary Results. *Radiology* 2004;230:652–9. <https://doi.org/10.1148/radiol.2303021331>.
- [57] van Houdt PJ, Agarwal HK, van Buuren LD, Heijmink SWTPJ, Haack S, van der Poel HG, et al. Performance of a fast and high-resolution multi-echo spin-echo sequence for prostate T2 mapping across multiple systems. *Magn Reson Med* 2017;79:1586–94. <https://doi.org/10.1002/mrm.26816>.
- [58] Malyarenko D, Galbán CJ, Londy FJ, Meyer CR, Johnson TD, Rehemtulla A, et al. Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom. *J Magn Reson Imaging* 2013;37:1238–46. <https://doi.org/10.1002/jmri.23825>.



- [59] Grech-Sollars M, Hales PW, Miyazaki K, Raschke F, Rodriguez D, Wilson M, et al. Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain. *NMR Biomed* 2015;28:468–85. <https://doi.org/10.1002/nbm.3269>.
- [60] Moreau B, Iannessi A, Hoog C, Beaumont H. How reliable are ADC measurements? A phantom and clinical study of cervical lymph nodes. *Eur Radiol* 2018;28:3362–71. <https://doi.org/10.1007/s00330-017-5265-2>.
- [61] Schmidt MA, Payne GS. Radiotherapy planning using MRI. *Phys Med Biol* 2015;60:R323–61. <https://doi.org/10.1088/0031-9155/60/22/R323>.
- [62] Gurney-Champion OJ, Mahmood F, van Schie M, Julian R, George B, Philippens MEP, et al. Quantitative imaging for radiotherapy purposes. *Radiother Oncol* 2020;146:66–75. <https://doi.org/10.1016/j.radonc.2020.01.026>.
- [63] Keenan KE, Biller JR, Delfino JG, Boss MA, Does MD, Evelhoch JL, et al. Recommendations towards standards for quantitative MRI (qMRI) and outstanding needs. *J Magn Reson Imaging* 2019;49:e26–39. <https://doi.org/10.1002/jmri.26598>.
- [64] Obuchowski NA, Buckler A, Kinahan P, Chen-Mayer H, Petrick N, Barboriak DP, et al. Statistical Issues in Testing Conformance with the Quantitative Imaging Biomarker Alliance (QIBA) Profile Claims. *Acad Radiol* 2016;23:496–506. <https://doi.org/10.1016/j.acra.2015.12.020>.
- [65] Lambrecht M, Vandecaveye V, De Keyzer F, Roels S, Penninckx F, Van Cutsem E, et al. Value of diffusion-weighted magnetic resonance imaging for prediction and early assessment of response to neoadjuvant radiochemotherapy in rectal cancer: Preliminary results. *Int J Radiat Oncol Biol Phys* 2012;82:863–70. <https://doi.org/10.1016/j.ijrobp.2010.12.063>.
- [66] Lambrecht M, Van Herck H, De Keyzer F, Vandecaveye V, Slagmolen P, Suetens P, et al. Redefining the target early during treatment. Can we visualize regional differences within the target volume using sequential diffusion weighted MRI? *Radiother Oncol* 2014;110:329–34. <https://doi.org/10.1016/j.radonc.2013.09.023>.
- [67] Park SY, Kim CK, Park BK, Park W, Park HC, Han DH, et al. Early changes in apparent diffusion coefficient from diffusion-weighted MR imaging during radiotherapy for prostate cancer. *Int J Radiat Oncol Biol Phys* 2012;83:749–55. <https://doi.org/10.1016/j.ijrobp.2011.06.2009>.
- [68] Liu Y, Bai R, Sun H, Liu H, Zhao X, Li Y. Diffusion-weighted imaging in predicting and monitoring the response of uterine cervical cancer to combined chemoradiation. *Clin Radiol* 2009;64:1067–74. <https://doi.org/10.1016/j.crad.2009.07.010>.
- [69] Padhani AR, Liu G, Mu-Koh D, Chenevert TL, Thoeny HC, Takahara T, et al. Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations. *Neoplasia* 2009;11:102–25. <https://doi.org/10.1593/neo.81328>.
- [70] Perfusion, Diffusion and Flow-MRI Biomarker Committee. QIBA Profile: Diffusion-Weighted Magnetic Resonance Imaging (DWI). Public comment draft. QIBA 2019.
- [71] Baltzer AP, Mann RM, Iima M, Sigmund EE, Clauser P, Gilbert F. Diffusion-Weighted Imaging of the breast – A consensus and mission statement from the EUSOBI International Breast Diffusion-Weighted Imaging working group 2019:1–25.
- [72] Kooreman ES, van Houdt PJ, Nowee ME, van Pelt VWJ, Tijssen RHN, Paulson ES, et al. Feasibility and accuracy of quantitative imaging on a 1.5 T MR-linear accelerator. *Radiother Oncol* 2019;133:156–62. <https://doi.org/10.1016/j.radonc.2019.01.011>.

- [73] Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data (vol 34, pg 910, 1995). *Magn Reson Med* 1996;36:332.
- [74] Tijssen RHN, Philippens MEP, Paulson ES, Glitzner M, Chugh B, Wetscherek A, et al. MRI commissioning of 1.5T MR-linac systems – a multi-institutional study. *Radiother Oncol* 2019;132:114–20. <https://doi.org/10.1016/j.radonc.2018.12.011>.
- [75] Baig TN, Eagan TP, Petropoulos LS, Kidane TK, Edelstein WA, Brown RW. Gradient coil with active endcap shielding. *Concepts Magn Reson Part B Magn Reson Eng* 2007;31B:12–23. <https://doi.org/10.1002/cmr.b.20079>.
- [76] Chan RW, Von Deuster C, Giese D, Stoeck CT, Harmer J, Aitken AP, et al. Characterization and correction of eddy-current artifacts in unipolar and bipolar diffusion sequences using magnetic field monitoring. *J Magn Reson* 2014;244:74–84. <https://doi.org/10.1016/j.jmr.2014.04.018>.
- [77] Meier C, Zwanger M, Feiweier T, Porter D. Concomitant field terms for asymmetric gradient coils: Consequences for diffusion, flow, and echo-planar imaging. *Magn Reson Med* 2008;60:128–34. <https://doi.org/10.1002/mrm.21615>.
- [78] Bammer R, Markl M, Barnett A, Acar B, Alley MT, Pelc NJ, et al. Analysis and generalized correction of the effect of spatial gradient field distortions in diffusion-weighted imaging. *Magn Reson Med* 2003;50:560–9. <https://doi.org/10.1002/mrm.10545>.
- [79] Jackson SJ, Glitzner M, Tijssen RHN, Raaymakers BW. MRI  $B_0$  homogeneity and geometric distortion with continuous linac gantry rotation on an Elekta Unity MR-linac. *Phys Med Biol* 2019;0–7. <https://doi.org/10.1088/1361-6560/ab231a>.
- [80] Le Bihan D. What can we see with IVIM MRI? *Neuroimage* 2019;187:56–67. <https://doi.org/10.1016/j.neuroimage.2017.12.062>.
- [81] Lemke A, Laun FB, Simon D, Stieltjes B, Schad LR. An in vivo verification of the intravoxel incoherent motion effect in diffusion-weighted imaging of the abdomen. *Magn Reson Med* 2010;64:1580–5. <https://doi.org/10.1002/mrm.22565>.
- [82] Riches SF, Hawtin K, Charles-Edwards EM, de Souza NM. Diffusion-weighted imaging of the prostate and rectal wall: Comparison of biexponential and monoexponential modelled diffusion and associated perfusion coefficients. *NMR Biomed* 2009;22:318–25. <https://doi.org/10.1002/nbm.1328>.
- [83] Barnhart HX, Barboriak DP. Applications of the repeatability of quantitative imaging biomarkers: a review of statistical analysis of repeat data sets. *Transl Oncol* 2009;2:231–5. <https://doi.org/10.1593/tlo.09268>.
- [84] Bito Y, Hirata S, Yamamoto E. Optimal gradient factors for ADC measurements. B. Abstr. Third Annu. Meet. Int. Soc. Magn. Resonance Med. ISMRM, Berkeley, CA: ISMRM; 1995, p. 913.
- [85] Xing D, Papadakis NG, Huang CLH, Lee VM, Carpenter TA, Hall LD. Optimised diffusion-weighting for measurement of apparent diffusion coefficient (ADC) in human brain. *Magn Reson Imaging* 1997;15:771–84. [https://doi.org/10.1016/S0730-725X\(97\)00037-4](https://doi.org/10.1016/S0730-725X(97)00037-4).
- [86] Saritas EU, Lee JH, Nishimura DG. SNR dependence of optimal parameters for apparent diffusion coefficient measurements. *IEEE Trans Med Imaging* 2011;30:424–37. <https://doi.org/10.1109/TMI.2010.2084583>.
- [87] Dietrich O, Heiland S, Sartor K. Noise correction for the exact determination of apparent diffusion coefficients at low SNR. *Magn Reson Med* 2001;45:448–53. [https://doi.org/10.1002/1522-2594\(200103\)45:3<448::AID-MRM1059>3.0.CO;2-W](https://doi.org/10.1002/1522-2594(200103)45:3<448::AID-MRM1059>3.0.CO;2-W).

- [88] Otazo R, Lambin P, Pignol JP, Ladd ME, Schlemmer HP, Baumann M, et al. MRI-guided Radiation Therapy: An Emerging Paradigm in Adaptive Radiation Oncology. *Radiology* 2021;298:248–60. <https://doi.org/10.1148/RADIOL.2020202747>.
- [89] van Houdt PJ, Yang Y, van der Heide UA. Quantitative Magnetic Resonance Imaging for Biological Image-Guided Adaptive Radiotherapy. *Front Oncol* 2021;10:1–9. <https://doi.org/10.3389/fonc.2020.615643>.
- [90] Hall WA, Paulson ES, van der Heide UA, Fuller CD, Raaymakers BW, Lagendijk JJW, et al. The transformation of radiation oncology using real-time magnetic resonance guidance: A review. *Eur J Cancer* 2019;122:42–52. <https://doi.org/10.1016/j.ejca.2019.07.021>.
- [91] Vaupel P, Mayer A. Hypoxia in cancer: Significance and impact on clinical outcome. *Cancer Metastasis Rev* 2007;26:225–39. <https://doi.org/10.1007/s10555-007-9055-1>.
- [92] Cao Y. The Promise of Dynamic Contrast-Enhanced Imaging in Radiation Therapy. *Semin Radiat Oncol* 2011;21:147–56. <https://doi.org/10.1016/j.semradonc.2010.11.001>.
- [93] Hompland T, Hakon Hole K, Ragnum HB, Aarnes E-K, Vlatkovic L, Lie AK, et al. Combined MR Imaging of Oxygen Consumption and Supply Reveals Tumor Hypoxia and Aggressiveness in Prostate Cancer Patients. *Cancer Res* 2018;78:4774–86. <https://doi.org/10.1158/0008-5472.CAN-17-3806>.
- [94] Zelfhof B, Pickles M, Liney G, Gibbs P, Rodrigues G, Kraus S, et al. Correlation of diffusion-weighted magnetic resonance data with cellularity in prostate cancer. *BJU Int* 2009;103:883–8. <https://doi.org/10.1111/j.1464-410X.2008.08130.x>.
- [95] Kato H, Esaki K, Yamaguchi T, Tanaka H, Kajita K, Furui T, et al. Predicting early response to chemoradiotherapy for uterine cervical cancer using intravoxel incoherent motion mr imaging. *Magn Reson Med Sci* 2019;18:293–8. <https://doi.org/10.2463/mrms.tn.2018-0138>.
- [96] Bian H, Liu F, Chen S, Li G, Song Y, Sun M, et al. Intravoxel incoherent motion diffusion-weighted imaging evaluated the response to concurrent chemoradiotherapy in patients with cervical cancer. *Medicine (Baltimore)* 2019;98:e17943. <https://doi.org/10.1097/MD.0000000000017943>.
- [97] Marzi S, Piludu F, Sanguineti G, Marucci L, Farneti A, Terrenato I, et al. The prediction of the treatment response of cervical nodes using intravoxel incoherent motion diffusion-weighted imaging. *Eur J Radiol* 2017;92:93–102. <https://doi.org/10.1016/j.ejrad.2017.05.002>.
- [98] Paudyal R, Oh JH, Riaz N, Venigalla P, Li J, Hatzoglou V, et al. Intravoxel incoherent motion diffusion-weighted MRI during chemoradiation therapy to characterize and monitor treatment response in human papillomavirus head and neck squamous cell carcinoma. *J Magn Reson Imaging* 2017;45:1013–23. <https://doi.org/10.1002/jmri.25523>.
- [99] Foltz WD, Wu A, Chung P, Catton C, Bayley A, Milosevic M, et al. Changes in apparent diffusion coefficient and T2 relaxation during radiotherapy for prostate cancer. *J Magn Reson Imaging* 2013;37:909–16. <https://doi.org/10.1002/jmri.23885>.
- [100] van Schie MA, van Houdt PJ, Ghobadi G, Pos FJ, Walraven I, de Boer HCJ, et al. Quantitative MRI Changes During Weekly Ultra-Hypofractionated Prostate Cancer Radiotherapy With Integrated Boost. *Front Oncol* 2019;9. <https://doi.org/10.3389/fonc.2019.01264>.
- [101] Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol* 2017;71:618–29. <https://doi.org/10.1016/j.eururo.2016.08.003>.

- [102] Kooreman ES, van Houdt PJ, Keesman R, Pos FJ, van Pelt VWJ, Nowee ME, et al. ADC measurements on the Unity MR-linac – A recommendation on behalf of the Elekta Unity MR-linac consortium. *Radiother Oncol* 2020. <https://doi.org/10.1016/j.radonc.2020.09.046>.
- [103] Turkbey B, Rosenkrantz AB, Haider MA, Padhani AR, Villeirs G, Macura KJ, et al. Prostate Imaging Reporting and Data System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2. *Eur Urol* 2019;76:340–51. <https://doi.org/10.1016/j.eururo.2019.02.033>.
- [104] Taouli B, editor. *Extra-Cranial Applications of Diffusion-Weighted MRI*. Cambridge: Cambridge University Press; 2010. <https://doi.org/10.1017/CBO9780511778070>.
- [105] Obuchowski NA. Interpreting Change in Quantitative Imaging Biomarkers. *Acad Radiol* 2018;25:372–9. <https://doi.org/10.1016/j.acra.2017.09.023>.
- [106] Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Soft* 2015;67:1–48. <https://doi.org/10.18637/jss.v067.i01>.
- [107] Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. *J Stat Software, Artic* 2017;82:1–26. <https://doi.org/10.18637/jss.v082.i13>.
- [108] He N, Li Z, Li X, Dai W, Peng C, Wu Y, et al. Intravoxel Incoherent Motion Diffusion-Weighted Imaging Used to Detect Prostate Cancer and Stratify Tumor Grade: A Meta-Analysis. *Front Oncol* 2020;10. <https://doi.org/10.3389/fonc.2020.01623>.
- [109] Jafari-Khouzani K, Paynabar K, Hajighasemi F, Rosen B. Effect of Region of Interest Size on the Repeatability of Quantitative Brain Imaging Biomarkers. *IEEE Trans Biomed Eng* 2019;66:864–72. <https://doi.org/10.1109/TBME.2018.2860928>.
- [110] Alonzi R, Padhani AR, Taylor NJ, Collins DJ, D'Arcy JA, Stirling JJ, et al. Antivascular effects of neoadjuvant androgen deprivation for prostate cancer: An in vivo human study using susceptibility and relaxivity dynamic MRI. *Int J Radiat Oncol Biol Phys* 2011;80:721–7. <https://doi.org/10.1016/j.ijrobp.2010.02.060>.
- [111] Hötter AM, Mazaheri Y, Zheng J, Moskowitz CS, Berkowitz J, Lantos JE, et al. Prostate Cancer: assessing the effects of androgen-deprivation therapy using quantitative diffusion-weighted and dynamic contrast-enhanced MRI. *Eur Radiol* 2015;25:2665–72. <https://doi.org/10.1007/s00330-015-3688-1>.
- [112] Barker HE, Paget JTE, Khan AA, Harrington KJ. The tumour microenvironment after radiotherapy: Mechanisms of resistance and recurrence. *Nat Rev Cancer* 2015;15:409–25. <https://doi.org/10.1038/nrc3958>.
- [113] Dinis Fernandes C, van Houdt PJ, Heijmink SWTPJ, Walraven I, Keesman R, Smolic M, et al. Quantitative 3T multiparametric MRI of benign and malignant prostatic tissue in patients with and without local recurrent prostate cancer after external-beam radiation therapy. *J Magn Reson Imaging* 2019;50:269–78. <https://doi.org/10.1002/jmri.26581>.
- [114] García-Figueiras R, Baleato-González S, Padhani AR, Luna-Alcalá A, Vallejo-Casas JA, Sala E, et al. How clinical imaging can assess cancer biology. *Insights Imaging* 2019;10:28. <https://doi.org/10.1186/s13244-019-0703-0>.
- [115] O'Connor JPB, Robinson SP, Waterton JC. Imaging tumour hypoxia with oxygen-enhanced MRI and BOLD MRI. *Br J Radiol* 2019;92:20180642. <https://doi.org/10.1259/bjr.20180642>.
- [116] Cao Y. The Promise of Dynamic Contrast-Enhanced Imaging in Radiation Therapy. *Semin Radiat Oncol* 2011;21:147–56. <https://doi.org/10.1016/j.semradonc.2010.11.001>.

- [117] Park JJ, Kim CK, Park SY, Simonetti AW, Kim EJ, Park BK, et al. Assessment of early response to concurrent chemoradiotherapy in cervical cancer: Value of diffusion-weighted and dynamic contrast-enhanced MR imaging. *Magn Reson Imaging* 2014;32:993–1000. <https://doi.org/10.1016/j.mri.2014.05.009>.
- [118] Xie T, Ye Z, Pang P, Shao G. Quantitative Multiparametric MRI May Augment the Response to Radiotherapy in Mid-Treatment Assessment of Patients with Esophageal Carcinoma. *Oncol Res Treat* 2019;42:326–33. <https://doi.org/10.1159/000499322>.
- [119] Wang P, Popovtzer A, Eisbruch A, Cao Y. An approach to identify, from DCE MRI, significant subvolumes of tumors related to outcomes in advanced head-and-neck cancer. *Med Phys* 2012;39:5277–85. <https://doi.org/10.1118/1.4737022>.
- [120] Baer AH, Hoff BA, Srinivasan A, Galbán CJ, Mukherji SK. Feasibility Analysis of the Parametric Response Map as an Early Predictor of Treatment Efficacy in Head and Neck Cancer. *Am J Neuroradiol* 2015;36:757–62. <https://doi.org/10.3174/ajnr.A4296>.
- [121] Wong KH, Panek R, Dunlop A, Mcquaid D, Riddell A, Welsh LC, et al. Changes in multimodality functional imaging parameters early during chemoradiation predict treatment response in patients with locally advanced head and neck cancer. *Eur J Nucl Med Mol Imaging* 2018;45:759–67. <https://doi.org/10.1007/s00259-017-3890-2>.
- [122] Kooreman ES, van Houdt PJ, Keesman R, van Pelt VWJ, Nowee ME, Pos F, et al. Daily Intravoxel Incoherent Motion (IVIM) In Prostate Cancer Patients During MR-Guided Radiotherapy—A Multicenter Study. *Front Oncol* 2021;11:1–9. <https://doi.org/10.3389/fonc.2021.705964>.
- [123] Lawrence LSP, Chan RW, Chen H, Keller B, Stewart J, Ruschin M, et al. Accuracy and precision of apparent diffusion coefficient measurements on a 1.5 T MR-Linac in central nervous system tumour patients. *Radiother Oncol* 2021;164:155–62. <https://doi.org/10.1016/j.radonc.2021.09.020>.
- [124] Joo I, Lee JM, Grimm R, Han JK, Choi BI. Monitoring vascular disrupting therapy in a rabbit liver tumor model: Relationship between tumor perfusion parameters at IVIM Diffusion-weighted MR imaging and those at dynamic contrast-enhanced MR Imaging. *Radiology* 2016;278:104–13. <https://doi.org/10.1148/radiol.2015141974>.
- [125] Wang J, Salzillo T, Jiang Y, Mackeyev Y, David Fuller C, Chung C, et al. Stability of MRI contrast agents in high-energy radiation of a 1.5T MR-Linac. *Radiother Oncol* 2021;161:55–64. <https://doi.org/10.1016/j.radonc.2021.05.023>.
- [126] Zhang X, Petersen ET, Ghariq E, De Vis JB, Webb AG, Teeuwisse WM, et al. In vivo blood T1 measurements at 1.5 T, 3 T, and 7 T. *Magn Reson Med* 2013;70:1082–6. <https://doi.org/10.1002/mrm.24550>.
- [127] Rohrer M, Bauer H, Mintonovitch J, Requardt M, Weinmann H-J. Comparison of Magnetic Properties of MRI Contrast Media Solutions at Different Magnetic Field Strengths. *Invest Radiol* 2005;40:715–24.
- [128] Georgiou L, Wilson DJ, Sharma N, Perren TJ, Buckley DL. A functional form for a representative individual arterial input function measured from a population using high temporal resolution DCE MRI. *Magn Reson Med* 2019;81:1955–63. <https://doi.org/10.1002/mrm.27524>.
- [129] Chang LC, Cheng GK, Basser PJ, Pierpaoli C. Linear least-squares method for unbiased estimation of T1 from SPGR signals. *Magn Reson Med* 2008;60:496–501. <https://doi.org/10.1002/mrm.21669>.
- [130] Cheong LH, Koh TS, Hou Z. An automatic approach for estimating bolus arrival time in dynamic contrast MRI using piecewise continuous regression models. *Phys Med Biol* 2003;48:N83–8. <https://doi.org/10.1088/0031-9155/48/5/403>.

- [131] Tofts PS, Brix G, Buckley DL, Evelhoch JL, Henderson E, Knopp M V, et al. Estimating Kinetic Parameters From Dynamic Contrast-Enhanced T<sub>1</sub>-Weighted MRI of a Diffusible Tracer: Standardized Quantities and Symbols. *J Magn Reson Imag* 1999;10:223–32. [https://doi.org/10.1002/\(SICI\)1522-2586\(199909\)10](https://doi.org/10.1002/(SICI)1522-2586(199909)10).
- [132] Bakdash JZ, Marusich LR. Repeated Measures Correlation. *Front Psychol* 2017;8:1–13. <https://doi.org/10.3389/fpsyg.2017.00456>.
- [133] Gao P, Shi C, Zhao L, Zhou Q, Luo L. Differential diagnosis of prostate cancer and noncancerous tissue in the peripheral zone and central gland using the quantitative parameters of DCE-MRI. *Medicine (Baltimore)* 2016;95:e5715. <https://doi.org/10.1097/MD.0000000000005715>.
- [134] Le Bihan D, Turner R. The Capillary Network: A Link between IVIM and Classical Perfusion. *Magn Reson Med* 1992;27:171–8.
- [135] Pang Y, Turkbey B, Bernardo M, Kruecker J, Kadoury S, Merino MJ, et al. Intravoxel incoherent motion MR imaging for prostate cancer: An evaluation of perfusion fraction and diffusion coefficient derived from different b -value combinations. *Magn Reson Med* 2013;69:553–62. <https://doi.org/10.1002/mrm.24277>.
- [136] Blyth CR. On Simpson's paradox and the sure-thing principle. *J Am Stat Assoc* 1972;67:364–6. <https://doi.org/10.1080/01621459.1972.10482387>.
- [137] Klawer EME, van Houdt PJ, Simonis FFJ, van den Berg CAT, Pos FJ, Heijmink SWTPJ, et al. Improved repeatability of dynamic contrast-enhanced MRI using the complex MRI signal to derive arterial input functions: a test-retest study in prostate cancer patients. *Magn Reson Med* 2019;81:3358–69. <https://doi.org/10.1002/mrm.27646>.
- [138] Klawer EME, Houdt PJ, Simonis FFJ, den Berg CAT, Pos FJ, Heijmink SWTPJ, et al. Erratum to: Improved repeatability of dynamic contrast-enhanced MRI using the complex MRI signal to derive arterial input functions: a test-retest study in prostate cancer patients (*Magn Reson Med*. 2019; 81: 3358–3369). *Magn Reson Med* 2021;85:2334–6. <https://doi.org/10.1002/mrm.28606>.
- [139] Alonzi R, Taylor NJ, Stirling JJ, D'Arcy JA, Collins DJ, Saunders MI, et al. Reproducibility and correlation between quantitative and semiquantitative dynamic and intrinsic susceptibility-weighted MRI parameters in the benign and malignant human prostate. *J Magn Reson Imaging* 2010;32:155–64. <https://doi.org/10.1002/jmri.22215>.
- [140] Cohen J, Cohen P, West SG, Aiken LS. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Third. Lawrence Erlbaum Associates, Inc.; 2003.
- [141] Joiner M, van der Kogel A. *Basic Clinical Radiobiology*. Fourth Edi. Hodder Education; 2009. <https://doi.org/10.1201/b13224>.
- [142] Cobb JG, Xie J, Gore JC. Contributions of Chemical and Diffusive Exchange to T<sub>1</sub> r Dispersion 2013;1366:1357–66. <https://doi.org/10.1002/mrm.24379>.
- [143] Wang L, Regatte RR. T<sub>1ρ</sub> MRI of Human Musculoskeletal System. *J Magn Reson Imaging* 2015;41:586–600. <https://doi.org/10.1002/jmri.24677>.
- [144] Jambor I, Pesola M, Taimen P, Merisaari H, Boström PJ, Minn H, et al. Rotating frame relaxation imaging of prostate cancer: Repeatability, cancer detection, and Gleason score prediction. *Magn Reson Med* 2016;75:337–44. <https://doi.org/10.1002/mrm.25647>.
- [145] Santyr GE, Henkelman RM, Bronskill MJ. Spin locking for magnetic resonance imaging with application to human breast. *Magn Reson Med* 1989;12:25–37. <https://doi.org/10.1002/mrm.1910120104>.

- [146] Zhou N, Chu C, Dou X, Li M, Liu S, Guo T, et al. Early Changes of Irradiated Parotid Glands Evaluated by T1rho-Weighted Imaging. *J Comput Assist Tomogr* 2017;41:472–6. <https://doi.org/10.1097/RCT.0000000000000547>.
- [147] Witschey WRT, Borthakur A, Elliott MA, Mellon E, Niyogi S, Wallman DJ, et al. Artifacts in T1ρ-weighted imaging: Compensation for B1 and B0 field imperfections. *J Magn Reson* 2007;186:75–85. <https://doi.org/10.1016/j.jmr.2007.01.015>.
- [148] Witschey WRT, Borthakur A, Elliott MA, Mellon E, Niyogi S, Wallman DJ, et al. Artifacts in T1ρ-Weighted Imaging: Compensation for B1 and B0 Field Imperfections. *J Magn Reson* 2007;186:75–85. <https://doi.org/10.1016/j.molcel.2009.10.020.The>.
- [149] Buck FM, Bae WC, Diaz E, Du J, Statum S, Han ET, et al. Comparison of T1rho measurements in agarose phantoms and human patellar cartilage using 2D multislice spiral and 3D magnetization prepared partitioned k-space spoiled gradient-echo snapshot techniques at 3 T. *Am J Roentgenol* 2011;196:174–9. <https://doi.org/10.2214/AJR.10.4570>.
- [150] Fokas E, Appelt A, Glynn-Jones R, Beets G, Perez R, Garcia-Aguilar J, et al. International consensus recommendations on key outcome measures for organ preservation after (chemo)radiotherapy in patients with rectal cancer. *Nat Rev Clin Oncol* 2021;18:805–16. <https://doi.org/10.1038/s41571-021-00538-5>.
- [151] Buck FM, Bae WC, Diaz E, Du J, Statum S, Han ET, et al. Comparison of T1rho measurements in agarose phantoms and human patellar cartilage using 2D multislice spiral and 3D magnetization prepared partitioned k-space spoiled gradient-echo snapshot techniques at 3 T. *AJR Am J Roentgenol* 2011;196:W174–9. <https://doi.org/10.2214/AJR.10.4570>.
- [152] Xie S, Li Q, Cheng Y, Zhang Y, Zhuo Z, Zhao G, et al. Impact of liver fibrosis and fatty liver on T1rho measurements: A prospective study. *Korean J Radiol* 2017;18:898–905. <https://doi.org/10.3348/kjr.2017.18.6.898>.
- [153] Lambregts DMJ, Boellaard TN, Beets-tan RGH. Response evaluation after neoadjuvant treatment for rectal cancer using modern MR imaging : a pictorial review 2019.
- [154] Ai QYH, Zhang H, Jiang B, So TY, Mo FKE, Qamar S, et al. Test-retest repeatability of T1rho (T1ρ) MR imaging in the head and neck. *Eur J Radiol* 2021;135:109489. <https://doi.org/10.1016/j.ejrad.2020.109489>.
- [155] Keenan KE, Gimbutas Z, Dienstfrey A, Stupic KF. Assessing effects of scanner upgrades for clinical studies. *J Magn Reson Imaging* 2019;50:1948–54. <https://doi.org/10.1002/jmri.26785>.
- [156] Subashi E, Dresner A, Tyagi N. Longitudinal assessment of quality assurance measurements in a 1.5 T MR-linac: Part II—Magnetic resonance imaging. *J Appl Clin Med Phys* 2022. <https://doi.org/10.1002/acm2.13586>.
- [157] Keesman R, van de Lindt TN, Juan-Cruz C, van den Wollenberg W, van der Bijl E, Nowee ME, et al. Correcting geometric image distortions in slice-based 4D-MRI on the MR-linac. *Med Phys* 2019;3044–54. <https://doi.org/10.1002/mp.13602>.
- [158] McDonald BA, Salzillo T, Mulder S, Ahmed S, Dresner A, Preston K, et al. In Vivo and Phantom Repeatability of Diffusion-Weighted MRI Sequences on 1.5 T MRI-Linear Accelerator (MR-Linac) and MR Simulator Devices for Head and Neck Cancers: Results from a Prospective R-IDEAL Stage 2a Evaluation of Tumor and Normal Tissue Apparent. *MedRxiv* 2022.

- [159] Meyer HJ, Wienke A, Surov A. Pre-treatment apparent diffusion coefficient does not predict therapy response to radiochemotherapy in cervical cancer: A systematic review and meta-analysis. *Anticancer Res* 2021;41:1163–70. <https://doi.org/10.21873/anticancer.14873>.
- [160] Harry VN, Persad S, Bassaw B, Parkin D. Diffusion-weighted MRI to detect early response to chemoradiation in cervical cancer: A systematic review and meta-analysis. *Gynecol Oncol Reports* 2021;38:100883. <https://doi.org/10.1016/j.gore.2021.100883>.
- [161] Zijlema SE, Tijssen RHN, Van Dijk L, Hackett SL, Wolthaus JWH, Breimer W, et al. Improving the imaging performance of the 1.5 T MR-linac using a flexible, 32-channel, on-body receive array. *Phys Med Biol* 2020;65. <https://doi.org/10.1088/1361-6560/aba87a>.
- [162] Sun H, Xu Y, Xu Q, Shi K, Wang W. Short-term reproducibility of intravoxel incoherent motion parameters in 3.0T magnetic resonance imaging. *Medicine (Baltimore)* 2017;96:e6866. <https://doi.org/10.1097/MD.0000000000006866>.
- [163] Gurney-Champion OJ, Klaassen R, Froeling M, Barbieri S, Stoker J, Engelbrecht MRW, et al. Comparison of six fit algorithms for the intra-voxel incoherent motion model of diffusion-weighted magnetic resonance imaging data of pancreatic cancer patients. *PLoS One* 2018;13:e0194590. <https://doi.org/10.1371/journal.pone.0194590>.
- [164] Lemke A, Stieltjes B, Schad LR, Laun FB. Toward an optimal distribution of b values for intravoxel incoherent motion imaging. *Magn Reson Imaging* 2011;29:766–76. <https://doi.org/10.1016/j.mri.2011.03.004>.
- [165] Merisaari H, Jambor I. Optimization of b-value distribution for four mathematical models of prostate cancer diffusion-weighted imaging using b values up to 2000 s/mm<sup>2</sup>: Simulation and repeatability study. *Magn Reson Med* 2015;73:1954–69. <https://doi.org/10.1002/mrm.25310>.
- [166] Wetscherek A, McDonald BA, Kooreman ES, Lau AZ, Ramesh P, Shukla-Dave A, et al. Data-driven optimization of intravoxel incoherent motion imaging for clinical endpoints in radiotherapy on a 1.5 T MR-linac [abstract]. *Jt. Annu. Meet. ISMRM-ESMRMB ISMRT 31st Annu. Meet.*, London: 2022.
- [167] Kaandorp MPT, Barbieri S, Klaassen R, van Laarhoven HWM, Crezee H, While PT, et al. Improved unsupervised physics-informed deep learning for intravoxel incoherent motion modeling and evaluation in pancreatic cancer patients. *Magn Reson Med* 2021;86:2250–65. <https://doi.org/10.1002/mrm.28852>.
- [168] Vannesjo SJ, Haerberlin M, Kasper L, Pavan M, Wilm BJ, Barmet C, et al. Gradient system characterization by impulse response measurements with a dynamic field camera. *Magn Reson Med* 2013;69:583–93. <https://doi.org/10.1002/mrm.24263>.
- [169] Jezzard P, Barnett AS, Pierpaoli C. Characterization of and correction for eddy current artifacts in echo planar diffusion imaging. *Magn Reson Med* 1998;39:801–12. <https://doi.org/10.1002/mrm.1910390518>.
- [170] Le Bihan D, Poupon C, Amadon A, Lethimonnier F. Artifacts and pitfalls in diffusion MRI. *J Magn Reson Imaging* 2006;24:478–88. <https://doi.org/10.1002/jmri.20683>.
- [171] Corkum MT, Achard V, Morton G, Zilli T. Ultrahypofractionated Radiotherapy for Localised Prostate Cancer: How Far Can We Go? *Clin Oncol* 2022;34:340–9. <https://doi.org/10.1016/j.clon.2021.12.006>.
- [172] Johnson LM, Turkbey B, Figg WD, Choyke PL. Multiparametric MRI in prostate cancer management. *Nat Rev Clin Oncol* 2014;11:346–53. <https://doi.org/10.1038/nrclinonc.2014.69>.
- [173] American College of Radiology. PI-RADS Version 2.1. Prostate Imaging-Reporting Data Syst 2019:1–69.



- [174] Thorwarth D, Ege M, Nachbar M, Mönnich D, Gani C, Zips D, et al. Quantitative magnetic resonance imaging on hybrid magnetic resonance linear accelerators: Perspective on technical and clinical validation. *Phys Imaging Radiat Oncol* 2020;16:69–73. <https://doi.org/10.1016/j.phro.2020.09.007>.
- [175] Chan RW, Lawrence LSP, Oglesby RT, Chen H, Stewart J, Theriault A, et al. Chemical exchange saturation transfer MRI in central nervous system tumours on a 1.5 T MR-Linac. *Radiother Oncol* 2021;162:140–9. <https://doi.org/10.1016/j.radonc.2021.07.010>.
- [176] Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duerk JL, et al. Magnetic resonance fingerprinting. *Nature* 2013;495:187–92. <https://doi.org/10.1038/nature11971>.
- [177] Mickevicius NJ, Kim JP, Zhao J, Morris ZS, Hurst NJ, Glide-Hurst CK. Toward magnetic resonance fingerprinting for low-field MR-guided radiation therapy. *Med Phys* 2021;48:6930–40. <https://doi.org/10.1002/mp.15202>.
- [178] Mickevicius NJ, Glide-Hurst CK. Low-rank inversion reconstruction for through-plane accelerated radial MR fingerprinting applied to relaxometry at 0.35 T. *Magn Reson Med* 2022;88:840–8. <https://doi.org/10.1002/mrm.29244>.
- [179] Bruijnen T, van der Heide O, Intven MPW, Mook S, Lagendijk JJW, Berg CAT van den, et al. Technical feasibility of Magnetic Resonance Fingerprinting on a 1.5T MRI-Linac 2020:1–17.
- [180] Press RH, Shu HKG, Shim H, Mountz JM, Kurland BF, Wahl RL, et al. The Use of Quantitative Imaging in Radiation Oncology: A Quantitative Imaging Network (QIN) Perspective. *Int J Radiat Oncol Biol Phys* 2018;102:1219–35. <https://doi.org/10.1016/j.ijrobp.2018.06.023>.
- [181] Yankeelov TE, Mankoff DA, Schwartz LH, Lieberman FS, Buatti JM, Mountz JM, et al. Quantitative imaging in cancer clinical trials. *Clin Cancer Res* 2016;22:284–90. <https://doi.org/10.1158/1078-0432.CCR-14-3336>.
- [182] Manfrini E, Smits M, Thust S, Geiger S, Bendella Z, Petr J, et al. From research to clinical practice: a European neuroradiological survey on quantitative advanced MRI implementation. *Eur Radiol* 2021;31:6334–41. <https://doi.org/10.1007/s00330-020-07582-2>.
- [183] McGee KP, Hwang KP, Sullivan DC, Kurhanewicz J, Hu Y, Wang J, et al. Magnetic resonance biomarkers in radiation oncology: The report of AAPM Task Group 294. *Med Phys* 2021;48:e697–732. <https://doi.org/10.1002/mp.14884>.



## SUMMARY

Hybrid systems that combine an MRI scanner with a linear accelerator for radiation treatment, called MR-linacs, were first introduced in the clinic in 2014. The MRI images provide better soft tissue contrast when compared to CT, enabling excellent depictions of the tumor and surrounding tissue. By integrating this with a linac, scans of the patient can be made during treatment. Besides the potential for reducing treatment margins, the MRI can also be used to acquire images that provide quantitative information about the patient biology (qMRI), such as diffusion weighted imaging (DWI) from which apparent diffusion coefficient (ADC) maps can be derived. Values from these images have the potential to be used as quantitative imaging biomarkers (QIBs) for treatment response monitoring or prediction. The MR-linac provides the benefit of acquiring qMRI with a high frequency, where daily acquisition is possible in most cases without increasing patient burden. The work in this thesis is focused on the acquisition of daily qMRI parameters on the 1.5 T Unity MR-linac system.

The Unity MR-linac was first used clinically in 2018. Because it is a newly introduced machine, it is important to assess its qMRI capabilities in phantoms. In the study described in Chapter 2, we showed that the Unity MR-linac is capable of acquiring  $T_1$ ,  $T_2$ , and ADC maps, with comparable accuracy and repeatability to diagnostic systems, although for the ADC maps some adjustments related to SNR had to be made to produce comparable metrics. A series of phantom tubes with an increasing amount of DCE contrast agent showed that we could accurately determine the concentration up to around 0.5 mM, but underestimated larger concentrations. Finally, we demonstrated feasibility of acquiring these qMRIs on the Unity MR-linac in a prostate cancer patient.

One of the benefits of the Unity MR-linac is that the system is identical in all centers. This reduces challenges with reproducibility which are common when acquiring the same QIBs on different systems. However, it is then also important to unify the way that specific QIBs are acquired on the Unity MR-linac. In Chapter 3, we provide recommendations for the acquisition of the ADC. These recommendations specifically focus on the Unity MR-linac, and how the adjusted design of the system might influence the ADC value. Specific recommendations include limiting the highest b-value to 500 s/mm<sup>2</sup>, scanning in close proximity of the iso-center to reduce the influence of spatial variations. We also showed that scanning while the gantry holding the linac is moving does not significantly increase the accuracy or repeatability, concluding that acquiring the ADC during treatment is possible.

Perfusion imaging is of interest as it has been shown to relate to hypoxia which is a known biomarker for treatment response. One clinically established way to measure quantitative parameters that are related to perfusion on MRI is dynamic contrast-

enhanced (DCE) MRI, which necessitates the use of a contrast agent. Using a contrast agent on a daily basis, is undesirable and therefore investigating alternatives that may provide perfusion information without the use of contrast agent is of interest. In Chapters 4 and 5, IVIM, which is such a technique is investigated in terms of treatment response monitoring (Chapter 4) and correlation to DCE (Chapter 5). In the multi-center study described in Chapter 4, 43 prostate cancer patients who received 20 fractions of 3 Gy were included, and daily IVIM scans were acquired. The change over time of the IVIM parameters  $D$ ,  $f$ , and  $D^*$  were analyzed using a mixed effects model. The model indicated no change in  $D$  in the healthy prostate tissue, while the  $D$  increased in the tumor. The perfusion parameters  $f$  and  $D^*$  increased in both the prostate and healthy tissue, showing that IVIM parameters could potentially be used for treatment response monitoring.

Correlations between IVIM- and DCE parameters have been investigated before, but mainly on values acquired at a single time point. For treatment response purposes, it is also valuable to determine if they correlate longitudinally, as that could indicate that IVIM can be used instead of DCE for treatment response monitoring. In Chapter 5, we investigated the longitudinal correlations between IVIM and DCE measurements. In a cohort of 20 prostate cancer patients, weekly DCE scans were acquired in the same session as an IVIM scan. The repeated measures correlation was calculated for different ROIs. On the group level, an increase in IVIM and DCE parameters was found over the course of treatment, except for  $D$  in the peripheral- and transition zones. Statistically significant longitudinal correlations, although all below 0.5, were found between IVIM- and DCE parameters in the transition zone and peripheral zone of the prostate, while no statistically significant correlations were found in the tumor. The results suggest that IVIM could potentially be used as an alternative for DCE when measuring longitudinally.

$T_{1\rho}$  is a type of qMRI sequence that is most commonly used in cartilage imaging. Recently, some studies suggested that it might also be useful as a QIB for tumor localization and treatment response monitoring in oncology. In Chapter 6, a study is described where a  $T_{1\rho}$  sequence was implemented and tested on the Unity MR-linac. Accuracy and repeatability were tested with a custom-made Agar phantom, and showed values comparable to previously reported values from diagnostic systems. Ten rectal cancer patients were scanned during their treatment of 5 fractions of 5 Gy. Results showed an increase in the GTV  $T_{1\rho}$  for most patients, with a statistically significant difference between the first and last fraction for the incomplete responder group ( $n = 5$ ). Evaluation in a larger cohort is needed to determine if there is a clinical benefit for the use of  $T_{1\rho}$  as a treatment response QIB.

In conclusion, this thesis describes studies assessing the acquisition of daily qMRI on the Unity MR-linac. Initial feasibility was shown, including accuracy, repeatability and reproducibility in phantoms and patients, and changes in different qMRI parameters over the course of radiation treatment were measured in several patient groups. Daily qMRI has potential in terms of treatment response monitoring, and MR-linacs seem to be an excellent platform to help advance this field.

## SAMENVATTING

MRI-versnellers zijn nieuwe, hybride systemen, waarin een MRI-scanner wordt gecombineerd met een lineaire deeltjesversneller. Met deze systemen kunnen patiënten die behandeld worden met radiotherapie, tijdens de bestralingssessie in beeld gebracht worden met de MRI. De MRI-beelden bieden een beter contrast van zacht weefsel in vergelijking met CT, waardoor uitstekende afbeeldingen van de tumor en het omliggende weefsel gemaakt kunnen worden. Naast de mogelijkheid om de behandelmarges te verkleinen, kan de MRI ook worden gebruikt om beelden te verkrijgen die kwantitatieve informatie geven over de tumor en het omliggende weefsel. Een voorbeeld hiervan is diffusie-gewogen imaging (DWI) waaruit zogenoemde ADC-maps kunnen worden afgeleid, die de plaatselijke diffusiecoëfficiënt weergeven. Waarden uit kwantitatieve beelden hebben het potentieel om te worden gebruikt als biomarkers voor het in de gaten houden van de behandeling of het voorspellen van de behandelrespons. De MR-linac biedt hier het voordeel dat het dagelijks verkrijgen van deze kwantitatieve beelden in de meeste gevallen mogelijk is zonder de patiënt extra te belasten. In dit proefschrift is een vijftal onderzoeken gebundeld waar gebruik is gemaakt van de 1.5 T Unity MR-linac, ontworpen door Elekta en Philips, en die zijn gericht op de acquisitie van dagelijkse kwantitatieve MRI-beelden op dit systeem.

De Unity MR-linac werd voor het eerst klinisch gebruikt in 2018. Omdat het een nieuw systeem is, wat op MRI-gebied aangepast is om ruimte te maken voor de lineaire versneller, is de eerste stap om de mogelijkheden met betrekking tot kwantitatieve MRI te karakteriseren. In de studie beschreven in Hoofdstuk 2 hebben we met behulp van fantomen, buisjes gevuld met vloeistoffen waarvan de MRI-karakteristieken bekend zijn, aangetoond dat de Unity MR-linac in staat is om  $T_1$ -,  $T_2$ - en ADC-maps te meten, met vergelijkbare nauwkeurigheid en herhaalbaarheid als diagnostische systemen. Voor de ADC-maps waren wel enkele aanpassingen met betrekking tot signal-to-noise ratio nodig om resultaten te verkrijgen die vergelijkbaar waren met diagnostische systemen. Een reeks fantoombuisjes met een toenemende hoeveelheid contrastmiddel toonde aan dat we de concentratie tot ongeveer 0.5 mM nauwkeurig konden bepalen, maar grotere concentraties onderschatten. Ten slotte hebben we de haalbaarheid aangetoond van het verkrijgen van deze kwantitatieve beelden op de Unity MR-linac bij een prostaatankerpatiënt.

Een van de voordelen van de Unity MR-linac is dat behandelcentra die het systeem hebben aangeschaft, beschikken over precies hetzelfde systeem. Dit vermindert problemen met reproduceerbaarheid die vaak voorkomen bij het verkrijgen van dezelfde kwantitatieve beelden op verschillende systemen. Het is echter ook belangrijk om de specifieke manier waarop kwantitatieve beelden worden verkregen op de Unity

MR-linac te standaardiseren. In Hoofdstuk 3 doen we aanbevelingen voor het meten van de ADC. Deze aanbevelingen richten zich specifiek op de Unity MR-linac en hoe het aangepaste ontwerp van het systeem de ADC-waarde kan beïnvloeden. Deze aanbevelingen bevatten onder andere het beperken van de hoogste b-waarde tot 500 s/mm<sup>2</sup> en het scannen in de nabijheid van het iso-centrum om de invloed van ruimtelijke variaties te verminderen. We hebben ook aangetoond dat het scannen van de ADC-map terwijl de lineaire versneller in beweging is de nauwkeurigheid en herhaalbaarheid niet significant beïnvloedt, en het dus mogelijk is om te meten terwijl de patiënt bestraald wordt.

Het meten van perfusie is van belang omdat is aangetoond dat het verband houdt met hypoxie, een biomarker voor respons. De gevestigde manier om kwantitatieve parameters te meten die verband houden met perfusie op MRI is met DCE-MRI, waarbij de patiënt contrastmiddel toegediend krijgt. Dagelijks gebruik van contrastmiddel is ongewenst en daarom is het interessant om alternatieven te onderzoeken die perfusie-informatie kunnen geven zonder het gebruik van contrastmiddel. In Hoofdstukken 4 en 5 wordt IVIM, een dergelijke techniek, onderzocht met betrekking tot het monitoren van de behandelrespons (hoofdstuk 4) en de correlatie met DCE (hoofdstuk 5). In de multicenter studie beschreven in Hoofdstuk 4 zijn 43 prostaatkankerpatiënten geïncludeerd die behandeld werden met 20 fracties van 3 Gy. Bij deze patiënten werden dagelijks IVIM-scans gemaakt. De verandering in de tijd van de IVIM-parameters  $D$ ,  $f$  en  $D^*$  werden geanalyseerd met behulp van een mixed effects-model. Het model gaf geen significante verandering in de  $D$  van het gezonde prostaatweefsel aan, terwijl de  $D$  in de tumor significant toenam. De perfusieparameters  $f$  en  $D^*$  namen toe in zowel de prostaat als het gezonde weefsel, wat aantoont dat IVIM-parameters mogelijk kunnen worden gebruikt voor het meten of monitoren van de behandelrespons.

Correlaties tussen IVIM- en DCE-parameters zijn eerder onderzocht, maar voornamelijk op metingen die op een enkel tijdstip zijn verkregen. Voor behandelresponsdoeleinden is het ook waardevol om te bepalen of ze longitudinaal correleren, aangezien dat erop zou kunnen wijzen dat IVIM kan worden gebruikt in plaats van DCE voor het monitoren van de behandelingsrespons. In Hoofdstuk 5 hebben we de longitudinale correlaties tussen IVIM- en DCE-metingen onderzocht. In een cohort van 20 prostaatkankerpatiënten werden wekelijkse DCE-scans gemeten in dezelfde sessie als een IVIM-scan. De correlatie met herhaalde metingen werd berekend voor verschillende regio's in de prostaat. Op groepsniveau werd een toename van IVIM- en DCE-parameters gevonden in de loop van de behandeling, behalve voor  $D$  in de perifere- en overgangszones. Statistisch significante longitudinale correlaties, hoewel allemaal onder de 0,5, werden gevonden tussen IVIM- en DCE-parameters in de overgangszone en perifere zone van de prostaat, terwijl er geen statistisch significante correlaties werden gevonden in de

tumor. De resultaten suggereren dat IVIM mogelijk kan worden gebruikt als alternatief voor DCE bij longitudinale metingen.

$T_{1\rho}$  is een kwantitatieve MRI-techniek die het meest wordt gebruikt bij de beeldvorming van kraakbeen. Enkele recente studies hebben gesuggereerd dat het ook nuttig zou kunnen zijn als een biomarker voor het lokaliseren van de tumor en het monitoren van de behandelrespons in de oncologie. In Hoofdstuk 6 wordt een studie beschreven waarbij een  $T_{1\rho}$ -sequentie werd geïmplementeerd en getest op de Unity MR-linac. Nauwkeurigheid en herhaalbaarheid werden getest met een op maat gemaakt Agar-fantom. Metingen toonden waarden die vergelijkbaar waren met eerder gerapporteerde waarden van diagnostische systemen. Tien rectumkankerpatiënten werden gescand tijdens hun behandeling van 5 fracties van 5 Gy. De resultaten toonden een toename van de  $T_{1\rho}$  in het GTV van de meeste patiënten, met een statistisch significant verschil tussen de eerste en de laatste fractie voor de patiënten met incomplete respons ( $n = 5$ ). Evaluatie in een groter cohort is nodig om te bepalen of er een klinisch voordeel is voor het gebruik van  $T_{1\rho}$  als een biomarker voor het meten van de behandelrespons.

Dit proefschrift beschrijft onderzoek naar het meten van dagelijkse kwantitatieve MRI-beelden op de Unity MR-linac. De mogelijkheid om dit soort metingen te doen werd aangetoond, waarbij de nauwkeurigheid, herhaalbaarheid en reproduceerbaarheid is gekarakteriseerd in fantomen en patiënten. Veranderingen in verschillende kwantitatieve MRI-parameters tijdens de bestralingsbehandeling werden gemeten in verschillende patiëntengroepen. Dagelijkse kwantitatieve MRI is veelbelovend als toepassing voor het monitoren van de behandelrespons, en MR-linacs lijken een uitstekend platform om dit onderzoeksveld vooruit te helpen.



## LIST OF PUBLICATIONS

### This thesis

Feasibility and accuracy of quantitative imaging on a 1.5 T MR-linear accelerator

**Kooreman, E. S.**, van Houdt, P. J., Nowee, M. E., van Pelt, V. W. J., Tijssen, R. H. N., Paulson, E. S., Gurney-Champion, O. J., Wang, J., Koetsveld, E., van Buuren, L. D., Ter Beek, L. C., & van der Heide, U. A. (2019)

*Radiotherapy and Oncology*, 133, 156–162 (2019)

DOI: 10.1016/j.radonc.2019.01.011

ADC measurements on the Unity MR-linac – A recommendation on behalf of the Elekta Unity MR-linac consortium

**Kooreman, E. S.**, van Houdt, P. J., Keesman, R., Pos, F. J., van Pelt, V. W. J., Nowee, M. E., Wetscherek, A., Tijssen, R. H. N., Philippens, M. E. P., Thorwarth, D., Wang, J., Shukla-Dave, A., Hall, W. A., Paulson, E. S., & van der Heide, U. A.

*Radiotherapy and Oncology*, 153, 106–113 (2020)

DOI: 10.1016/j.radonc.2020.09.046

Daily Intravoxel Incoherent Motion (IVIM) In Prostate Cancer Patients During MR-Guided Radiotherapy–A Multicenter Study

**Kooreman, E. S.**, van Houdt, P. J., Keesman, R., van Pelt, V. W. J., Nowee, M. E., Pos, F., Sikorska, K., Wetscherek, A., Müller, A. C., Thorwarth, D., Tree, A. C., & van der Heide, U. A. *Frontiers in oncology*, 11, 705964 (2021)

DOI: 10.3389/fonc.2021.705964

Longitudinal Correlations Between Intravoxel Incoherent Motion (IVIM) and Dynamic Contrast-Enhanced (DCE) MRI During Radiotherapy in Prostate Cancer Patients

**Kooreman, E. S.**, van Pelt, V., Nowee, M. E., Pos, F., van der Heide, U. A., & van Houdt, P. J. *Frontiers in oncology*, 12, 897130 (2022)

DOI: 10.3389/fonc.2022.897130

T<sub>10</sub> for Radiotherapy Treatment Response Monitoring in Rectal Cancer Patients: A Pilot Study

**Kooreman, E. S.**, Tanaka, M., Ter Beek, L. C., Peters, F. P., Marijnen, C. A. M., van der Heide, U. A., & van Houdt, P. J.

*Journal of clinical medicine*, 11(7), 1998 (2022)

DOI: 10.3390/jcm11071998

**Other publications**

Controlled tip wear on high roughness surfaces yields gradual broadening and rounding of cantilever tips

Vorselen, D., **Kooreman, E. S.**, Wuite, G. J., & Roos, W. H.

*Scientific reports*, 6, 36972 (2016)

DOI: 10.1038/srep36972

Evaluation of bi-objective treatment planning for high-dose-rate prostate brachytherapy-A retrospective observer study

Maree, S. C., Luong, N. H., **Kooreman, E. S.**, van Wieringen, N., Bel, A., Hinnen, K. A., Westerveld, H., Pieters, B. R., Bosman, P. A. N., & Alderliesten, T.

*Brachytherapy*, 18(3), 396–403 (2019)

DOI: 10.1016/j.brachy.2018.12.010

Integration of quantitative imaging biomarkers in clinical trials for MR-guided radiotherapy: Conceptual guidance for multicentre studies from the MR-Linac Consortium Imaging Biomarker Working Group

van Houdt, P. J., Saeed, H., Thorwarth, D., Fuller, C. D., Hall, W. A., McDonald, B. A., Shukla-Dave, A., **Kooreman, E. S.**, Philippens, M. E. P., van Lier, A. L. H. M. W., Keesman, R., Mahmood, F., Coolens, C., Stanescu, T., Wang, J., Tyagi, N., Wetscherek, A., & van der Heide, U. A.

*European journal of cancer (Oxford, England : 1990)*, 153, 64–71 (2021)

DOI: 10.1016/j.ejca.2021.04.041

Double delay alternating with nutation for tailored excitation facilitates banding-free isotropic high-resolution intracranial vessel wall imaging

Coolen, B. F., Schoormans, J., Gilbert, G., **Kooreman, E. S.**, de Winter, N., Viessmann, O., Zwanenburg, J. J. M., Majoie, C. B. L. M., Strijkers, G. J., Nederveen, A. J., & Siero, J. C. W.

*NMR in biomedicine*, 34(9), e4567 (2021)

DOI: 10.1002/nbm.4567

## Conferences

Repeatability of Quantitative Imaging on the MR-linac for Treatment Response Monitoring

**Kooreman, E.S.\***, van Houdt, P.J.†, Nowee, M.E., van Pelt, V.W.J., Koetsveld, F., ter Beek, L.C., Peeters, J.M., van der Heide, U.A.

*10<sup>th</sup> Annual Meeting of the Benelux ISMRM Chapter, Antwerp, Belgium (2018)*

\*Oral presentation

*Joint Annual Meeting ISMRM-ESMRMB, Paris, France (2018)*

\*Electronic poster presentation

*37<sup>th</sup> ESTRO conference, Barcelona, Spain (2018)*

†Poster presentation

Repeatability of T<sub>2</sub> and ADC mapping on the MR-linac in prostate cancer patients

**Kooreman, E.S.\***, van Houdt, P.J., Nowee, M.E., van Pelt, V.W.J., Koetsveld, F., ter Beek, L.C., Peeters, J.M., van der Heide, U.A.

*6<sup>th</sup> MR in RT Symposium, Utrecht, The Netherlands (2018)*

\*Poster presentation

Daily IVIM and weekly DCE time trends in prostate cancer patients during radiation treatment

**Kooreman, E.S.\***, van Houdt, P.J., Nowee, M.E., Tree, A.C., Hall, W.A., van der Heide, U.A.

*36<sup>th</sup> Annual Scientific Meeting of the ESMRMB, Rotterdam, The Netherlands (2019)*

\*Oral presentation

Influence of the linear accelerator of a 1.5 T MR-linac on diffusion imaging during radiation treatment

**Kooreman, E.S.**, van Houdt, P.J., van Pelt, V.W.J., Nowee, M.E., van der Heide, U.A.\*

*7<sup>th</sup> MR in RT Symposium, Toronto, Canada (2019)*

\*Oral presentation

Intravoxel incoherent motion MRI, a potential alternative for DCE MRI in prostate cancer patients

**Kooreman, E.S.**, van Houdt, P.J.\*, van Pelt, V.W.J., Nowee, M.E., Pos, F., van der Heide, U.A.

*ESTRO 2021, Madrid, Spain & Online (2021)*

\*Oral presentation

First T<sub>1ρ</sub> mapping results on a 1.5 T MR-linac

**Kooreman, E.S.\***, van Houdt, P.J., ter Beek, L.C., Peters, F.P., Nowee, M.E., van der Heide, U.A.

*Virtual 8<sup>th</sup> MR in RT Symposium, Online (2021)*

\*Poster presentation, 2<sup>nd</sup> poster prize

## ACKNOWLEDGEMENTS

Beste Uulke, bedankt voor je begeleiding in mijn vorming tot onderzoeker en je onvermoeibare steun en aanmoediging door de afgelopen jaren. Met name ben ik dankbaar voor het begrip en de vrijheid die je me gaf toen ik er voor mijn gezin moest zijn.

Petra, door jou voelde ik me vanaf dag één thuis op het NKI. Heel fijn dat je altijd tijd had voor wat voor vraag dan ook en bedankt voor al je steun tijdens mijn promotieonderzoek.

Roque, good morning my workplace proximity acquaintance. Thank you for all the fun moments in our room in the office, teaching me Spanish, and supporting me throughout the PhD. I am honored that you are my paronymph.

All colleagues from the Uulke group, Ahmed, Anke, Catarina, Chavelli, Edzo, Ghazaleh, Karolina, Mar, Marcel, Rita, Robin, Roelant, Stefan, and Thijs, I had an amazing time with you all. Thank you for the great discussions, input in my work, and support over the last years.

Rick, ik heb altijd erg genoten van onze discussies, met name over de MRI. Vaak ging het mijn pet te boven. Je hebt in die twee jaar veel betekend voor mijn wetenschappelijke vorming en het huidige resultaat was er zonder jou niet geweest.

Maartje en Maarten, heel erg leuk dat ik jullie heb begeleid bij jullie masterstages. Beide projecten, hoewel bemoeilijkt door corona, hebben tot mooie eindresultaten geleid.

Beste collega's van de RT-afdeling en de Jan-Jakob groep, in het bijzonder Anja, Celia, Diedie, Leon, Marlies, Patricia, Tessa, Vineet, en Vivian, heel erg bedankt voor al jullie ondersteuning en hulp bij mijn onderzoeken.

Iris, ik vind het erg leuk dat wij de afgelopen 20 jaar zoveel samen hebben opgetreden in orkesten en op tournee zijn geweest. Bedankt dat je mijn paranimf wilt zijn.

Familie, ik wil jullie heel hartelijk bedanken voor al jullie interesse en ondersteuning tijdens mijn onderzoek. Jullie continue aanmoedigingen en oprechte interesse hebben me enorm geholpen om door te zetten.

David, Dewi, Florian, Josephine, en Nick, ik wil jullie erg bedanken voor alle fijne en gezellige afleiding die er vaak voor zorgde dat ik er weer tegenaan kon.

Lieve Frederike, heel erg bedankt voor al je ondersteuning en begrip de afgelopen jaren. Je hebt me altijd zo veel mogelijk geholpen en eindeloos geduld getoond. Toen een analyse opnieuw moest heb je hele avonden scriptjes voor me gedraaid. Zonder jou was dit proefschrift er zeker niet geweest. Ik kijk erg uit naar onze toekomst samen met Jasmijn.

## **CURRICULUM VITAE**

Ernst Kooreman werd op 16 november 1990 geboren in Zeist. In 2010 behaalde hij zijn vwo-diploma aan de Stichtse Vrije School, waarna hij aan de Vrije Universiteit in Amsterdam Medische Natuurwetenschappen ging studeren. Hierop volgde hij de bijbehorende master met specialisatie in “medical physics”. In zijn afstudeerjaar liep hij stage op twee afdelingen in het AMC in Amsterdam, waar hij eerst onderzoek heeft gedaan naar een MRI-sequentie voor de beeldvorming van de intracraniale bloedvaatjes, en daarna op de radiotherapie-afdeling heeft bijgedragen aan een onderzoek naar de verbetering van automatisch gegenereerde brachytherapieplannen voor prostaatkankerpatienten. In zijn promotieonderzoek, waar hij in 2017 aan begon op de radiotherapie-afdeling van het Antoni van Leeuwenhoek, kwamen beide onderwerpen, MRI en radiotherapie, samen. Hier deed hij onderzoek naar de inzet van kwantitatieve MRI voor het voorspellen van de behandeluikomst van patiënten die behandeld werden met een MRI-versneller. Dit onderzoek resulteerde in het proefschrift met de titel “Daily Quantitative MRI for Radiotherapy Response Monitoring”. Sinds 2023 is hij actief als data scientist bij Valcon.



