Multiparametric MRI for focal dose escalation in prostate cancer radiotherapy

Schie, M.A. van

Multiparametric MRI for focal dose escalation in prostate cancer radiotherapy

Marcel Alexander van Schie

Multiparametric MRI for focal dose escalation in prostate cancer radiotherapy
The work described in this thesis was conducted at the Netherlands Cancer Institute in Amsterdam, and was funded by the Dutch Cancer Society, project 10088.
ISBN 978-94-6416-582-1
Cover design Sandra Tukker
Print Ridderprint | www.ridderprint.nl

Multiparametric MRI for focal dose escalation in prostate cancer radiotherapy

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 woensdag 30 juni 2021
klokke 13.45 uur
door

Marcel Alexander van Schie
geboren te Singapore
in 1991

Promotor
Prof.dr. U.A. van der Heide

Co-promotoren

Dr. L.G.W. Kerkmeijer	*Radboudumc Nijmegen*
Dr. F.J. Pos	*NKI-AVL Amsterdam*

Promotiecommissie

Prof.dr. C.R.N. Rasch	
Prof.dr. B.J.M. Heijmen	*Erasmus University Rotterdam*
Prof.dr. C. Ménard	*University of Montréal*

Contents

1

CHAPTER 1

Introduction

Prostate cancer

Prostate cancer is a common disease that presents primarily in elderly men, with an average age at diagnosis of 66 years (Rawla *et al.* 2019). According to the latest global cancer survey, prostate cancer is responsible for 1.28 million new cases and 359 thousand deaths per year (Ferlay *et al.* 2019). While ranked as the second most common type of male cancer worldwide, in the Netherlands prostate cancer is the leading type of male cancer diagnosed, with an incidence of over 13 thousand cases and almost three thousand deaths per year (NKR 2019).

Prostate cancer is a heterogeneous disease that can present in multiple disease stages and levels of aggressiveness. To a large extent the life expectancy of prostate cancer patients depends on the stage of the disease. The majority of prostate cancers are slow-developing indolent tumors that are confined to the prostate gland. Approximately 15% of the patients present with more aggressive, fast-proliferating tumors that may be accompanied with extracapsular extension or metastatic disease (Chang *et al.* 2014).

Due to screening programs, most new prostate cancer cases are diagnosed without apparent complaints of the patient (Donnelly *et al.* 2019). A blood sample measurement showing an increase of the prostate specific antigen (PSA) may raise suspicion, although nonmalignant conditions as benign prostate enlargement (BPE) or prostatitis can cause similar abnormal PSA scores.

Additional imaging and needle biopsies are recommended to differentiate between cancer and benign conditions (Barentsz *et al.* 2012). Upon imaging any visible tumor tissue is localized and staged. Depending on whether the imaging involves a Digital Rectal Examination (DRE) or an MRI examination, either a clinical or radiological tumor (T-) stage is established. The biopsies reveal tissue pathology at multiple locations in the prostate and are assigned to an ISUP (International Society of Urological Pathology) grade group based on the tissue's cell differentiation (van Leenders *et al.* 2020). While previously ultrasound-guided biopsies were taken and MRI examination was performed in case of positive biopsy cores, current clinical practice involves initial MRI acquisition followed by MRI-targeted biopsies to reduce patient burden and improve the detection of clinically significant prostate cancer (Giganti *et al.* 2017).

The combination of PSA, T-stage and grade group represents the clinical condition of the disease and allows to categorize patients by prostate cancer risk groups. These risk groups represent the chance of developing metastatic disease after primary treatment. In Europe the EAU risk classification is adopted that differentiates between low, intermediate and high-risk patients (Mottet *et al.* 2017). Higher risk is generally associated with poorer survival.

Radiotherapy treatment

Alongside prostatectomy, radiotherapy is a suitable primary treatment option for low to high-risk disease (Kishan *et al.* 2017, Moris *et al.* 2020). Radiotherapy can be delivered internally with brachytherapy, using radiation source tubes or seed implants, or externally with radiation beams, referred to as external beam radiotherapy (EBRT). Treatment with a linear accelerator (linac) is the standard of care for EBRT. To deliver EBRT treatment, an anatomical MRI and a simulation CT are acquired to respectively define anatomical regions of interest and retrieve tissue density information to calculate the attenuation of the radiation dose in the patient. With treatment planning software the beam positions, shape and dose rate are calculated to deliver a homogeneous dose distribution to the entire prostate gland and simultaneously limit the dose to surrounding organs at risk, such as the bladder and rectum. The dosimetrically optimized treatment plan is delivered in multiple fractions, which allows for recovery of benign tissue in between fractions. Radiotherapy is often combined with long-term hormonal therapy in the form of androgen deprivation to reduce metastatic spread and recurrence of the disease (Böhmer *et al.* 2016).

Whole gland dose escalation, in which an elevated radiation dose is prescribed to the entire prostate gland, was found to improve biochemical recurrence free survival rates among low to intermediate risk patients (Pollack *et al.* 2002, Peeters *et al.* 2006, Dearnaley *et al.* 2007). Although further dose escalation may lead to improved recurrence free survival among intermediate to high-risk patients as well (Pollack *et al.* 2002, Peeters *et al.* 2006, Morgan *et al.* 2007), it would also induce unacceptable damage to organs at risk. Interestingly, prostate tumors were found to recur predominantly at the location of the primary tumor site (Cellini *et al.* 2002, Pucar *et al.* 2007). Therefore, focal dose escalation based on tumor tissue presence seems a reasonable approach to increase local control while restricting dose to organs at risk.

Multiparametric MRI

To apply a dose escalation within the prostate gland, soft tissue contrast is required to identify the intraprostatic tumor. MRI is a valuable non-invasive imaging technique to reveal excellent soft tissue contrast (Owrangi *et al.* 2018). A T2-weighted sequence can quickly obtain high-resolution anatomical information from the pelvic region and is therefore the most frequently scanned MRI sequence to localize prostate cancer (Cabarrus *et al.* 2017). The T2-weighted image reflects the T2 relaxation time of tissues relative to each other and visualizes malignant tissue in the prostate as hypointense regions. For delineation purpose the display settings are often adjusted to maximize image contrast.

Although an anatomical T2-weighted MRI gives a good indication on the tumor dimensions, additional biological information greatly contributes to determine the tumor extent and, moreover, reveals physical and physiological characteristics of the tumor (Olsson *et al.* 2019). Imaging of such tumor biology is performed with functional imaging. Functional imaging in prostate cancer primarily focusses on diffusion and perfusion measurements but may also involve measurements of tumor metabolism and hypoxia.

Diffusion weighted imaging (DWI) measures the restricted motion of water molecules in biological tissue and allows to reveal details of the microscopic tissue composition. Since prostate tumors have a higher cell density than healthy prostate tissue, diffusion of water is reduced in the tumor, resulting in a hyperintense region on the diffusion-weighted image. Usually, a series of images with different diffusion weighting is acquired (Maurer *et al.* 2017). From this series a per-voxel apparent diffusion coefficient (ADC) map can be derived that eliminates inherent T2-weighting from the diffusion-weighted images and thereby quantifies the apparent local diffusion within the tissue.

Dynamic contrast enhanced (DCE-) MRI involves the recording of a time series of T1-weighted images of a contrast agent distribution in a region of interest. Several tracer kinetic models exist that apply a cell compartment approximation to estimate the true tissue vascularity (Brix *et al.* 1991, Buckley *et al.* 1994, Tofts 1997, Tofts *et al.* 1999). Among multiple parameters that together model the local tissue perfusion, the volume transfer constant K^{trans} is a commonly investigated parameter in prostate cancer. Increased values of K^{trans} are associated with leaky vascularity which indicates the presence of malignant tissue.

The combination of anatomical and functional MRI is referred to as multiparametric MRI (mp-MRI). Mp-MRI improves the sensitivity of tumor detection considerably. While the sensitivity of clinically detectable tumors on T2-weighted images is 0.73, this value increases to 0.85 – 0.89 when combined with DWI and DCE-imaging (Heijmink *et al.* 2007, Zhang *et al.* 2017, Woo *et al.* 2017).

For prostate cancer detection and staging of the disease, the PI-RADS V2 guidelines recommend a combination of T2-weighted and DWI to be scanned, with optional DCE-MRI (Weinreb *et al.* 2016). Although PI-RADS leads to more consistency in localization of the tumor (Rudolph *et al.* 2020), to date no guidelines exist on the use of mp-MRI to delineate the tumor for treatment purpose. This implies that current institutional practice can only be based on local experience and expertise. Several studies have shown that in the absence of guidelines large inter-observer variability exists when delineating the visible tumor on mp-MRI (Bratan *et al.* 2013, Rischke *et al.* 2013, Anwar *et al.* 2014, Steenbergen *et al.* 2015).

Alternatively, uncertainties introduced by human interpretation can be omitted when manual delineations are replaced by machine learning approaches. In such approach computational models are applied to imaging data and optionally combined with clinical parameters, to calculate a probability distribution of tumor presence in the prostate. The performance of the machine learning model depends on the amount and quality of the data the model has learned from in the training phase. Whenever the training data was labelled, the training phase can be considered as supervised learning. For prostate cancer usually a dataset of patients with ground truth information derived from histopathological data forms the labelled training data.

Quantitative MRI

In line with the PI-RADS V2 guidelines, T2-weighted and DW-MRI sequences are scanned to detect and stage prostate tumors. These are popular sequences for their high spatial resolution, high contrast between tumor and benign tissue, and fast imaging protocols. However, for quantification of the tumor tissue these sequences are not suitable. The dimensionless values that are acquired hamper a comparison of image values between patients or institutions.

MR images that contain values with physical meaning are called quantitative MRI. The physical values of quantitative MRI allow to compare image data from different patients at different scanning devices. It also enables the comparison of consecutive images of the same patient over a period of time. This is specifically interesting for assessing the response of tumor and surrounding tissue to the treatment, which could ultimately lead to improved treatment strategies. Quantitative MRI may also contribute to the development of continuous dose prescription maps that are automatically derived according to and at the resolution of the acquired quantitative images (Bentzen 2005).

T2 and ADC maps are commonly investigated quantitative MRI parameters for dose painting purpose and response assessment of prostate cancer. T2 maps are derived from a series of T2-weighted images, analogous to how ADC maps are derived from diffusion-weighted images. Since the T2 values represent the true transverse relaxation within each voxel in the image, T2 times characterize certain tissues. K^{trans} as a physical quantity may also qualify as quantitative parameter. However, K^{trans} values have a high uncertainty due to the variation within and between investigated patient cohorts (Huang *et al.* 2016), and the actual meaning of those K^{trans} values depends on the tracer kinetic model that was used (Khalifa *et al.* 2014). Therefore, K^{trans} maps are less suitable as quantitative image parameters.

Although quantitative MRI parameters by definition should be equivalent between scanners and institutions, differences in imaging protocol, patient setup and scanner settings are sources of variation in obtained quantitative values. Standardization of these aspects is essential to fully utilize the potential of quantitative MRI in radiotherapy applications (Gurney-Champion *et al.* 2020).

Image-based dose escalation strategies

Using MRI for tumor localization, several studies have demonstrated the feasibility of an image-based focal dose escalation in the prostate while preserving dose constraints to surrounding organs at risk. In a planning study with three patients Singh *et al.* considered delivery of 95 Gy to the dominant intraprostatic lesion using EBRT feasible with acceptable levels of toxicity (Singh *et al.* 2007). The HEIGHT trial demonstrated in 35 patients the feasibility of planning a dose escalation up to 89.3 Gy to the intraprostatic tumor while maintaining strict constraints to organs at risk (Bossart *et al.* 2016). Combining EBRT with concurrent brachytherapy with integrated boost was also found feasible with good outcomes for biochemical control, acute and late toxicities (Gomez-Itturiaga *et al.* 2016, Vigneault *et al.* 2016). In the phase II TARGET trial, 80 patients received 76 Gy from EBRT, combined with either an integrated boost of 95 Gy or a single brachytherapy boost of 10 Gy. Initial results have shown acceptable dosimetry and comparable toxicity and quality of life between both arms of the study (Sanmamed *et al.* 2020).

Focal dose escalation: FLAME

The phase III randomized controlled FLAME trial (Focal Lesion Ablative Microboost in ProstatE, NCT01168479) was performed between 2009 and 2015 to investigate the benefit of an integrated focal boost on 5-year biochemical recurrence free survival in a multi-institutional and single blinded setting (Lips *et al.* 2011). On an institutional level, patients were randomly assigned to either a standard treatment with 77 Gy prescribed to the entire prostate gland, or an experimental treatment with an integrated dose escalation up to 95 Gy to the visible tumor. In four participating institutions in total 571 patients were included.

In addition to the CT scan required for dose calculation purpose, an mp-MRI was acquired to identify the tumor in the prostate and delineate the intraprostatic tumor accordingly. This tumor delineation was defined as gross tumor volume (GTV). Dose painting by contours treatment planning was performed using local treatment planning software. Identical to standard prostate radiotherapy, planning target volume (PTV) coverage was prioritized over dose to organs at risk. For patients in the dose-escalated arm of the trial, objectives were

added to increase the dose to the GTV and preserve the dose to organs at risk. The aim was to achieve a tumor dose of 95 Gy, provided that dose constraints to organs at risk were prioritized.

Due to the integrated boost to an extreme dose in combination with strict constraints to organs at risk, the constitution of a focal dose escalation plan was a complex procedure. Since the optimization function of the treatment planning system was expanded with additional dose objectives, and the iterative manual optimization steps needed to be performed within a reasonable time frame, it is not guaranteed that the GTVs received the highest possible dose that the patient anatomy would allow for.

Recently, the FLAME trial was found to demonstrate a significant increase in 5-year biochemical disease-free survival, from 85% in the standard to 92% in the dose-escalated treatment arm, and a significant GTV dose response relation was observed (Kerkmeijer *et al.* 2021). The nonzero biochemical recurrence rate in the dose-escalated arm of the trial may partially be explained by potential undertreatment of GTVs, suggesting that more consistent treatment planning methods may increase the biochemical recurrence free survival rate even further. A common approach to improve consistency between treatment plans regarding optimal target coverage and organ at risk sparing is knowledge-based planning (KBP) (Wu *et al.* 2009). KBP methods utilize a large dataset of optimized dose distributions for different types of patient anatomy. The achievable doses to target volumes and organs at risk of a new patient are predicted from a subset of similar patient anatomies and corresponding dose distributions from the database (Wu *et al.* 2009, Appenzoller *et al.* 2012, Yuan *et al.* 2012, Wang *et al.* 2013, Wang *et al.* 2017, Wall *et al.* 2018). In standard prostate radiotherapy KBP has led to improved dose distributions and enabled automated plan quality assessment (Janssen *et al.* 2019). In future clinical settings of focal dose escalated treatment, KBP may find similar applications.

Hypofractionated focal dose escalation

Fractionation of the prescribed radiation dose allows for tissue recovery in between consecutive treatment fractions. The sensitivity of tissue to radiation can be described with the linear-quadratic (LQ) model (McMahon *et al.* 2018). The surviving fraction of cells is dependent on the radiation dose D, the intrinsic radiosensitivity parameter α, and the repair capability parameter β. The ratio between α and β describes the fractionation sensitivity of cells and relates to the response time of the cells to radiation.

Evidence was found that the α/β ratio of prostate tumors is lower than the surrounding normal tissue, which is opposite to most other tumor sites (Brenner *et al.* 1999, Vogelius *et al.* 2013). This implies that a sufficiently large dose per fraction will do relatively more damage to tumor

cells than to normal tissue. As a result, patient outcome would benefit from a treatment delivered in fewer treatment fractions with larger fraction doses (Ritter *et al.* 2008, Benjamin *et al.* 2017).

In a recent meta-analysis of 13 randomized trials that studied conventional fractionation with 1.8 – 2.0 per fraction and moderate hypofractionation with 2.4 – 3.4 Gy per fraction, the low α/β ratio was confirmed with a highly significant dose response (Vogelius *et al.* 2018). Indeed, in low to intermediate risk prostate cancer, moderate hypofractionation was shown to be non-inferior and with comparable complication rates as conventional fractionation (Dearnaley *et al.* 2016, Brand *et al.* 2019, Widmark *et al.* 2019). Besides the radiobiological advantage, also practical aspects are in favor of a hypofractionated treatment approach: overall treatment time is reduced, the patient comfort is improved, and resources can be utilized more efficiently.

Based on the aforementioned evidence, recommendation guidelines on the delivery of moderate and even extreme hypofractionation up to 5 Gy per fraction have been published (Morgan *et al.* 2018). As a result, moderate hypofractionation has become the new standard of care for all risk groups. Extreme hypofractionation is considered a save treatment option for low to intermediate risk disease, while intermediate to high-risk patients should only receive such treatment in trial setting. A recent meta-analysis of phase III randomized trials observed similar levels of safety and efficacy in conventional fractionation, moderate hypofractionation and extreme hypofractionation schemes among low to high-risk patients (Lehrer *et al.* 2020).

During enrollment of the FLAME trial, conventional fractionation was delivered with 2.0 Gy per fraction to the prostate, with an optional boost to 2.2 Gy to the delineated tumor. The increasing attention for moderate and extreme hypofractionation over the last years has led to the hypothesis that intermediate to high-risk patients could also benefit from improved treatment outcome and reduced treatment time if such hypofractionated radiation scheme is combined with focal dose escalation. Several studies currently investigate if extreme hypofractionation combined with a focal boost can be safely and effectively delivered to these patients. In the phase II DELINEATE trial, both conventional and moderate hypofractionation were combined with a focal dose escalation (Murray *et al.* 2020). Patients received either 74 Gy in 37 fractions or 60 Gy in 20 fractions, with integrated boost doses of 82 and 67 Gy respectively. In the phase II Hypo-FLAME trial, extreme hypofractionation of 35 Gy in five weekly fractions was delivered with an integrated boost up to 50 Gy (Draulans *et al.* 2020). Similarly, the phase II SPARC trial treated patients in five fractions with 36.25 Gy and up to 47.5 Gy to the intraprostatic tumor (Nicholls *et al.* 2020). While these trials have reported efficacy of the treatment with acceptable toxicity levels, the primary endpoint will eventually

confirm the overall benefit of combined hypofractionation and focal dose escalation for intermediate to high-risk patients.

Dose Painting by Numbers

Dose painting was introduced as a novel planning technique to incorporate both tumor location and characteristics as derived from imaging (Ling *et al.* 2000). Focal dose escalation as performed in the FLAME and hypo-FLAME trials involved a discrete elevated dose prescription to the identified intraprostatic lesion, and is referred to as Dose Painting by Contours. As observed in pathology, the actual prostate tumor boundary is more irregular than delineations suggest (Steenbergen *et al.* 2015). It has been shown that certain types of tumor tissue may be completely missed on mp-MRI (van Houdt *et al.* 2020). Moreover, multiple levels of cell differentiation and aggressiveness may present in the tumor, with different levels of radioresistance. Due to the finite resolution of mp-MRI and the inability to visualize all tumor tissue, delineated tumor boundaries may be inaccurate, and the uniform dose escalation may not always match with the local tumor tissue characteristics.

Dose Painting by Numbers (DPbN) is a treatment strategy that reflects uncertainties of the target definition in terms of boundary irregularities and tumor tissue heterogeneity (Bentzen 2005). DPbN allows to omit manual contouring and instead prescribe dose at the resolution of the mp-MRI. In addition, besides the radiological images DPbN could also be performed on higher order image features that may contain tumor characteristics invisible to the human eye. Often biological heterogeneity is modelled to base the dose prescription on (Vanderstraeten *et al.* 2006, Thorwarth *et al.* 2007, Differding *et al.* 2017, Grönlund *et al.* 2019, Yan *et al.* 2019). Specifically in prostate cancer the modelling of tumor presence is valuable for the purpose of DPbN. Voxel-wise conversion from tumor presence probability to prescription dose is enabled by a calibration function (Bowen *et al.* 2009). In such calibration function dose levels range between a minimum value to guarantee sufficient tumor control and a maximum value that could be delivered safely in performed trials.

Irrespective of the chosen calibration function, the omission of manual contouring and discrete dose levels may lead to propagation of image value uncertainties to the planned dose distribution. Therefore, investigation of the repeatability of image-based dose prescription is essential for the development of DPbN as a robust candidate for dose escalation treatment.

Treatment response monitoring

Not only for treatment planning purpose but also during treatment delivery mp-MRI may potentially further improve current practice. In one of the side studies of the Hypo-FLAME trial (NCT02853110) five weekly mp-MRI were performed to prepare for future MRI-guided treatment. The mp-MRI acquired at each treatment fraction visit allowed to investigate the potential role of mp-MRI for treatment response monitoring and early-adaptive treatment strategies. To achieve such treatment strategies with imaging, an assessment of anatomical, functional or molecular image characteristics, called imaging biomarkers, is required (ESR 2015, Dregely *et al.* 2018). Tracking changes of the tumor appearance in the Hypo-FLAME dataset is a suitable candidate for assessment. In radiotherapy an imaging biomarker, such as tumor intensity change, has a prognostic value if it relates to outcome regardless of the radiation dose. The prognostic value of the imaging biomarker alone, however, will not be suitable to base treatment adaptation on. Imaging biomarkers that do qualify for adaptive treatment need to have predictive value as well (Oldenhuis *et al.* 2008, Gurney-Champion *et al.* 2020). Predictive imaging biomarkers predict patient outcome depending on the radiation dose that was delivered. They allow to relate radiation dose to patient outcome using a calibration curve and act upon early changes during treatment.

Thesis outline

Dose escalation to the intraprostatic tumor has been shown to improve outcome of patients with intermediate to high-risk disease (Kerkmeijer *et al.* 2021). The effectiveness of such novel treatment depends to a large extent on the accuracy of the delivered treatment and the optimal use of available information about the tumor physiology. Over the past decades mp-MRI has found application at all stages of the radiotherapy workflow (Kerkmeijer *et al.* 2018, Olsson *et al.* 2019). This thesis describes mp-MRI applications and models that were specifically developed for, and may play an important role in the evolution of dose escalated treatment of prostate cancer.

Chapters 2 and 3 evaluate to what extent the treatment was realized as intended, and to what extent clinical practice could be further improved. Retrospective analysis of the FLAME dataset was performed to evaluate the quality of tumor delineations and escalated dose levels. Prediction models were applied to demonstrate potential improvements to MRI-based tumor delineations and realized dose escalations on an individual basis. In chapter 2 it was investigated if upon delineation of mp-MRI, besides variation between observers also institutional differences apply. The soundness of individual delineations was verified with an

automatic tumor detection model that may ultimately serve as delineation QA tool in clinical practice. In chapter 3 the realized dose to the tumor was evaluated and compared with the highest achievable dose escalation as predicted by a KBP model to estimate how well the dose escalation to the visible tumor was realized with respect to the prescribed dose and with respect to the patient anatomy.

In chapter 4 and 5, the role for mp-MRI to future dose escalated treatment was investigated. This included DPbN treatment planning and response monitoring. In DPbN, the direct conversion from image parameters to prescription dose, may allow image value uncertainties to propagate into the planned dose distribution. In chapter 4 a test-retest planning study is described to estimate the robustness of DPbN to these uncertainties. In chapter 5 repeated quantitative MRI were analyzed to evaluate if early-responding tissue can be identified during treatment. Such tissue changes would mark the start of establishing predictive imaging biomarkers that could be used in adaptive treatment strategies.

In chapter 6 the findings of this thesis are discussed to answer the main questions: how well can mp-MRI-based focal dose escalation with an extreme dose to the prostate tumor be delivered, how can predictive models assist to improve such dose escalated treatment, and how can potential mp-MRI applications contribute to the evolution of dose painting strategies for prostate cancer.

2

CHAPTER 2

Contouring of prostate tumors on multiparametric MRI: evaluation of clinical delineations in a multicenter radiotherapy trial

Marcel A. van Schie
Cuong V. Dinh
Petra J. van Houdt
Floris J. Pos
Stijn W.T.J.P. Heijmink
Linda G.W. Kerkmeijer
Alexis N.T.J. Kotte
Raymond Oyen
Karin Haustermans
Uulke A. van der Heide

Radiotherapy & Oncology, Volume 128, pp321–326 (2018)

Abstract

Purpose

To date no guidelines are available for contouring prostate cancer inside the gland, as visible on multiparametric (mp-) MRI. We assessed inter-institutional differences in interpretation of mp-MRI in the multicenter phase III FLAME trial.

Materials and methods

We analyzed clinical delineations on mp-MRI and clinical characteristics from 260 patients across three institutes. We performed a logistic regression analysis to examine each institute's weighting of T2w, ADC and K^{trans} intensity maps in the delineation of the cancer. As reviewing of all delineations by an expert panel is not feasible, we made a selection based on discrepancies between a published tumor probability (TP) model and each institute's clinical delineations using Areas Under the ROC Curve (AUC) analysis.

Results

Regression coefficients for the three institutes were -0.07, -0.27 and -0.11 for T2w, -1.96, -0.53 and -0.65 for ADC and 0.15, 0.20 and 0.62 for K^{trans}, with significant differences between institutes for ADC and K^{trans}. AUC analysis showed median AUC values of 0.92, 0.80 and 0.79. Five patients with lowest AUC values were reviewed by a uroradiologist.

Conclusions

Regression coefficients revealed considerably different interpretations of mp-MRI in tumor contouring between institutes and demonstrated the need for contouring guidelines. Based on AUC values outlying delineations could efficiently be identified for review.

Introduction

Radiotherapy is one of the standard treatment options for prostate cancer. Although it has been shown that tumor foci are non-uniformly distributed over the prostate (Chen *et al.* 2000), the prostate is usually irradiated with a more or less homogeneous dose distribution. Local recurrence of the disease has been observed at the original location of the tumor, suggesting an insufficient radiation dose at that location (Cellini *et al.* 2002, Pucar *et al.* 2007). As dose escalation to the entire gland would likely increase treatment-related toxicity, a focal dose escalation was proposed (Lips *et al.* 2011). Recently, accrual of patients in the FLAME trial (clinicaltrials.gov identifier NTC01168479) was ended. This large multi-center single-blinded randomized controlled phase III trial aimed to investigate the clinical benefit of focal escalation of the radiation dose to the visible cancer to 95 Gy. This required delineation of the tumor as visible on multiparametric (mp-) MRI, consisting of a T2-weighted (T2w) scan, a diffusion-weighted MRI (DWI) and a dynamic contrast-enhanced (DCE) MRI.

Guidelines on detection, localization, characterization and risk stratification of suspected prostate cancer using recommended mp-MRI were published in the Prostate Imaging – Reporting And Data System (PI-RADS) in 2012, and updated to PI-RADS v2 in 2015 (Barentsz *et al.* 2012, Dickinson *et al.* 2013, Weinreb *et al.* 2016). These guidelines were however not available when the majority of patients were included in the FLAME trial. Moreover, guidelines on contouring of prostate tumors based on mp-MRI are not available to date. Steenbergen *et al.* showed the large inter-observer variability that exists in a prostate tumor delineation study using mp-MRI (Steenbergen *et al.* 2015). Such variability can also be expected in the FLAME trial. In the absence of guidelines, institutional differences in contouring practice caused by differences in interpretation and weighting of the various sequences in mp-MRI scans may have occurred as well.

In this work we investigated the contours of prostate tumors in the FLAME trial, focusing on the weighting of the individual mp-MRI sequences within three institutes. We combined mp-MRI data with the actual clinical delineations to assess the relative contribution of each MRI sequence to the tumor contouring decision. As revision of all contours by a panel of experts is not feasible, we applied a pathology validated model for prostate tumor localization in order to identify cases that showed discrepancies between clinical delineations and MRI data. We selected the patients with the highest inconsistency between predicted tumor location and delineation and reviewed these retrospectively.

Materials and methods

Patient characteristics

We analyzed 260 prostate cancer patients who were included in the FLAME trial and randomized in the escalated dose arm. These patients had biopsy-proven prostate cancer, clinically localized intermediate or high-risk disease and no evidence of metastatic disease, according to Ash *et al.* (2000). Institutional review board approval was obtained and all patients provided written informed consent. The patients were treated in three institutes: 160 patients in the University Medical Center Utrecht (UMCU), 54 patients in the Netherlands Cancer Institute (NKI), and 46 patients in the University Hospitals in Leuven (UZL). Thirty-five patients were excluded because they had missing MRI data (15), missing biopsy reports (3), missing delineations (10), registration artifacts (2) or they did not receive the escalated dose (5), which led to analysis of 140, 33 and 52 patients from UMCU, UZL and NKI, respectively.

MRI data

All patients received an mp-MRI exam consisting of a T2w, DWI and DCE sequence. Specifications of the scanner type and sequences for each of the institutes are listed in Table 1. An apparent diffusion coefficient (ADC) map was derived from the b-values of the DWI using a mono-exponential fit. We determined the volume transfer constant (K^{trans}) values with the Tofts model using a population-based arterial input function (Tofts *et al.* 1999, Murase *et al.* 2004). Within each institute a radiation oncologist in consultation with a radiologist had prospectively delineated the prostate and all tumors visible on mp-MRI.

Image processing

We processed the MRI data according to the method of Dinh *et al.* (2017). To minimize the impact of differences between acquisition protocols among the participating institutes, this method applies a normalization of T2w and K^{trans} to the median signal intensity in the peripheral zone (PZ). Since no PZ delineations were made in our cohort, we assumed that 75% of the prostate volume was PZ tissue and 25% central gland (Weinreb *et al.* 2016). The T2w signal in the central gland tends to be lower than in the PZ, while the K^{trans} is higher (Kayhan *et*

Table 1. *mp-MRI parameters per institute.*

MRI parameters	I-1	I-2	I-3
Scanner type	Siemens 1.5 T	Philips 3.0 T	Philips 3.0 T
Scanner sequence			
T2w			
Pixel size / slice thickness (mm)	0.78 / 4.0	0.40 / 3.0	0.49–1.0 / 2.5–4.0
TE / TR (ms)	124 / 11250	120 / 3126–3828	110–130 / 2698–6717
ADC			
Pixel size / slice thickness (mm)	2.73 / 4.0	1.07–1.11 / 3.0–3.7	1.17–2.38 / 2.5–4.0
TE / TR (ms)	67 / 7110–9900	58–73 / 2712–3500	59–94 / 3119–10036
b-values (s/mm^2)	0, 50, 100, 500, 750, 1000	0, 188, 375, 563, 750 or 200, 400, 600, 800 *	0, 300, 500, 1000 or 0, 300, 1000 *
Ktrans			
Pixel size / slice thickness (mm)	1.37–1.68 / 4.0	1.02–1.36 / 3.0	0.94–2.5 / 2.5–7.0
TE / TR (ms)	1.5 / 4–5	1.9 / 4–5	1.0–1.7 / 4
Dynamic scan time (s)	4.4	2.5	2.5

* b = 0 s/mm^2 was acquired but not used for ADC map calculation.

al. 2010). Considering the upper and lower 75% of the T2w and K^{trans} signal intensity histograms as belonging to the PZ, we normalized the signal to the upper and lower 37.5% respectively.

The data set per patient consisted of normalized T2w and K^{trans} images, ADC, biopsy map and tumor prevalence map, plus the clinically delineated tumor and prostate. From the delineations a labeling mask was derived that contained labels for healthy and tumor tissue within the prostate. The data sets were resampled to an in-plane resolution of 0.49 mm, equal to the resolution of the image data used by Dinh *et al.* (2017), and a slice thickness of 1.0 mm.

Institutional interpretation

We evaluated the institutional differences on interpretation of the mp-MRI with a logistic regression analysis on voxel level of three intensity features, i.e. the T2w, ADC and K^{trans} intensity images. A transformation of each feature *i* to zero mean and unit variance was applied to allow comparison between features. The logistic regression function is:

$$F(x) = \frac{1}{1 + e^{-\left(\beta_0 + \sum \beta_i x_i\right)}}, \quad (1)$$

where $F(x)$ is the probability that voxel x is included in the tumor delineation, x_i is the intensity value of feature i, β_i is the regression coefficient of feature i and represents weight factor, and β_0 is the offset.

Probability model

For the automatic evaluation of the manual delineations we used a published tumor probability (TP) model (Dinh *et al.* 2017), which is a logistic regression model trained on mp-MRI and biopsy data and validated on histology data from 40 patients in two institutes. The coefficients of the TP model are found in Table 2. We combined 29 features from the normalized mp-MRI with biopsy and prevalence information, and applied the TP model to calculate a TP per voxel within the prostate.

For each calculated TP map and labeling mask we derived the Area Under the receiver operating characteristic Curve (AUC). We selected the patients that had a large disagreement between calculated TP map and labeling mask with AUC values below 0.50 and reviewed the

Table 2. *Regression coefficients β_i and offset β_0 of the TP model (Dinh et al. 2017). $G_i(s)$ stands for the Gaussian derivative in direction i, with standard deviation s (in mm).*

Feature description	β	Feature description	β
Offset - β_0	0.1	T2w – $G_y(7.6)$	-18.7
T2w – Intensity	-5.3 e-4	T2w – $G_{xx}(7.6)$	-3.0
T2w – G(3.0)	-10.9	T2w – $G_{yy}(7.6)$	2.6
T2w – $G_x(3.0)$	2.0	T2w – $G_{xy}(7.6)$	-16.6
T2w – $G_y(3.0)$	-3.3	T2w – G(12.0)	39.5
T2w – $G_{xx}(3.0)$	-2.9	T2w – $G_x(12.0)$	-1.4
T2w – $G_{yy}(3.0)$	-0.9	T2w – $G_y(12.0)$	-6.4
T2w – $G_{xy}(3.0)$	-4.0	T2w – $G_{xx}(12.0)$	3.1
T2w – G(4.8)	-16.2	T2w – $G_{yy}(12.0)$	-3.5
T2w – $G_x(4.8)$	-3.8	T2w – $G_{xy}(12.0)$	-6.9
T2w – $G_y(4.8)$	-7.8	ADC – Intensity	-2753
T2w – $G_{xx}(4.8)$	2.2	ADC – Blobness	-376.8
T2w – $G_{yy}(4.8)$	-0.1	K^{trans} – Intensity	0.3
T2w – $G_{xy}(4.8)$	-9.1	K^{trans} – Blobness	0.2
T2w – G (7.6)	-13.2	Biopsy map	1.3
T2w – $G_x(7.6)$	4.1	Prevalence map	0.1

clinical delineations. For each of the review cases we related the TP map with the clinical delineations and described the likely cause of disagreement.

Statistics

We tested patient characteristics and properties of the delineations as well as logistic regression coefficients for statistically significant difference on an institutional level. Median age, iPSA, prostate volume, tumor volume and regression coefficients were pairwise tested with a Mann-Whitney U test at a significance level of $\alpha = 0.05$. Clinical T stage, Gleason Score, number of tumors per patient and tumor location were pairwise tested with a Fisher exact test. Statistical tests were performed with the Statistics and Machine Learning Toolbox in MATLAB (version R2017a, MathWorks, Natick, MA). A Bonferroni correction was applied to correct for multiple testing.

Results

Table 3 shows the patient characteristics and delineation properties for the three institutes. We found significant differences in age (between institute (I-) 1 and I-2/I-3), iPSA (between institute I-1 and I-2/I-3), number of delineated tumors per patient (between I-1 and I-3) and tumor volume (between I-1 and I-2/I-3). No significant differences were observed for clinical T stage, Gleason Score, tumor location and prostate volume.

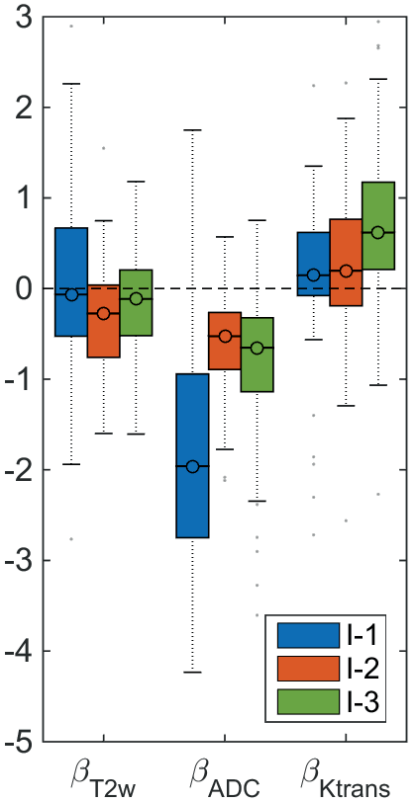
Figure 1 displays the obtained coefficients β_i from equation (1). Mean coefficients for I-1, I-2 and I-3 were -0.07, -0.27 and -0.11 for T2w, -1.96, -0.53 and -0.65 for ADC and 0.15, 0.20 and 0.62 for K^{trans}. The strongest negative coefficient was observed for ADC in I-1, the strongest positive coefficient for K^{trans} in I-3. All institutes weighted the ADC map the most. Comparing within each MRI sequence, T2w, ADC and K^{trans} were weighted the most by I-2, I-1 and I-3, respectively. Statistically significant differences were found between I-1 and I-2/I-3 for ADC and between I-1/I-2 and I-3 for K^{trans} ($p < 0.001$). This demonstrates a different interpretation of MRI sequences among the three institutes.

Median AUC values per institute were 0.92, 0.80 and 0.79, and ranged between 0.50 – 0.99, 0.47 – 0.98 and 0.19 – 0.98, for I-1, I-2 and I-3 respectively. A histogram of the distribution of calculated AUC values is shown in Figure 2. We found a majority of 92% with AUC values above

0.60. Five delineations scored an AUC below 0.50. These were reviewed as cases R-1 to R-5 by a uroradiologist with 14 years of experience. Figure 3 shows the T2w, ADC and K^{trans}, biopsy, prevalence and the obtained TP map of R-1 to R-5 at a representative slice in the prostate. An example patient P-1 with a high AUC is added for comparison.

Table 3. *Patient characteristics and delineation properties per institute. Pairwise p-values between institutes are reported, significant values are printed bold (Bonferroni correction for 24 tests). Median values are reported for age, iPSA, prostate volume and tumor volume; mean values for number of tumors per patient. Numbers in parentheses represent range.*

	I-1	I-2	I-3	p-value		
				p_{1-2}	p_{2-3}	p_{1-3}
Patient characteristics						
Age (y)	75 (64–82)	68 (51–78)	71 (56–80)	**< 0.001**	0.008	**0.001**
iPSA (ng/mL)	8.0 (1.8–29.0)	12.7 (3.6–44.7)	12.4 (2.6–100)	**0.002**	0.94	**< 0.001**
% Clinical T stage				0.017	0.22	0.033
T1c	3.0	3.9	12.1			
T2a	12.1	9.6	7.9			
T2b	0.0	11.5	4.3			
T2c	18.2	15.4	15.0			
T3a	48.5	30.8	38.6			
T3b	9.1	28.9	21.4			
T4	9.1	0.0	0.7			
% Gleason score				0.028	0.78	0.004
≤ 6	3.0	21.2	21.4			
7	69.7	46.2	40.7			
≥ 8	27.3	32.7	37.9			
% Tumor location				0.17	0.022	0.80
PZ	61.5	45.6	55.9			
PZ and TZ	23.1	41.2	23.9			
TZ	15.4	13.2	20.3			
Delineation characteristics						
Tumors (n)	1.2 (1–3)	1.3 (1–3)	1.6 (1–5)	0.72	0.057	**0.002**
Prostate volume (cm^3)	40.4 (16.9–101)	34.8 (17.9–105)	43.8 (13.8–157)	0.23	0.009	0.27
Tumor volume (cm^3)	0.69 (0.05–14.2)	3.3 (0.36–27.2)	3.6 (0.11–50.3)	**< 0.001**	0.97	**< 0.001**



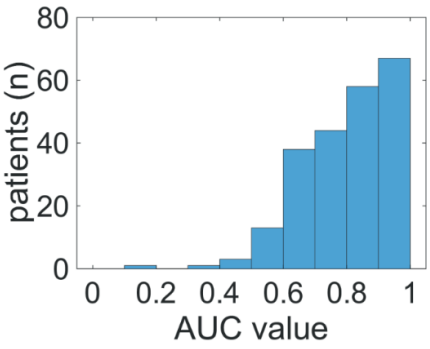


Figure 2. *Histogram of AUC values calculated with the TP model for 225 patients.*

Figure 1. *Box plot of regression coefficients β_i from equation (1) for all patients per institute. Positive regression coefficients represent a direct relation between intensity value and tumor presence, negative coefficients represent an inverse relation.*

R-1 and R-2 showed a transurethral resection of the prostate (TURP) cavity. In the TURP cavity T2w and ADC signal intensities are high and K^{trans} intensities are low, which causes low TP values. In R-1, with an AUC of 0.19, the tumor was delineated around the TURP. Upon review however, suspected tissue was localized in the left PZ, based on low ADC and high K^{trans} signal intensity. The area around the TURP cavity was considered as post-operative tissue response. In R-2 (AUC = 0.31) the clinical tumor delineation enclosed almost the whole prostate, with exception of the dorsolateral left PZ. The reviewer suspected tumor tissue in the latter region as well.

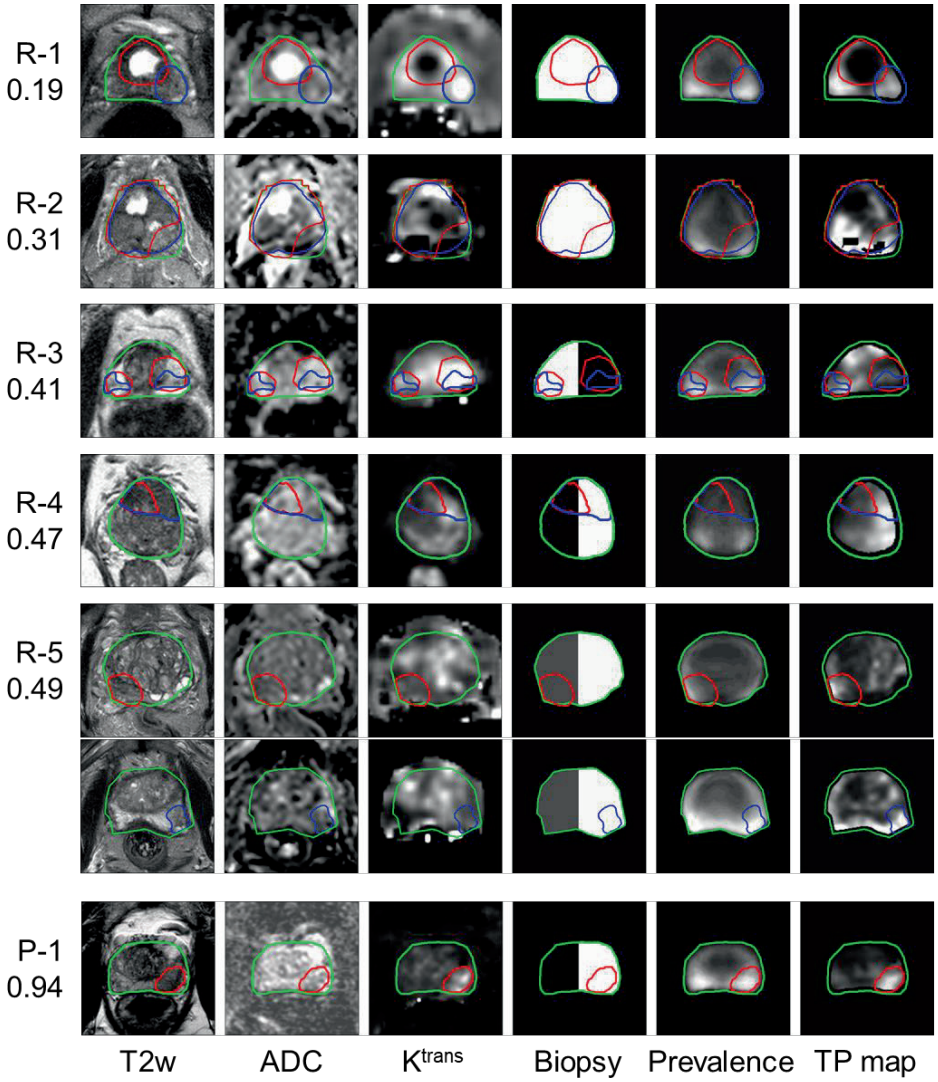


Figure 3. *Overview of T2w, ADC and K^{trans} signal intensity images, biopsy and prevalence map, and the calculated TP map of the reviewed patients R-1 to R-5 and example patient P-1 with corresponding AUC values. Prostates are delineated in green, clinical tumor delineations in red. Tumor delineations after review are blue. The reviewed delineation in R-5 was in a different axial plane and is therefore displayed on two rows.*

The location of the delineations in R-3 and R-4 were consistent with the mp-MRI. In R-3 (AUC = 0.41) the reviewer agreed with the locations of both contours based on T2w and K^{trans}, but the contoured volumes were too large in some parts, while missing parts of the suspected tissue elsewhere. R-4 had an AUC of 0.47. Upon review, the TZ delineation was considered too small and should have included the left ventral part of the TZ as well, based on low ADC and high K^{trans} signal.

In R-5, with an AUC of 0.49, the reviewer regarded the clinical tumor delineation as nonmalignant central gland tissue. Instead, a PI-RADS 2 region on the left apex side on the T2w image was mildly suspect for tumor and supported with positive biopsy map on the left side of the prostate (not indicated in Figure 3). ADC and K^{trans} values did not confirm this finding. However, TP values were moderately high in the clinical tumor delineation, while higher TP values were found in the left apex region, indicated by the reviewer.

Discussion

Delineation of the visible cancer on mp-MRI was required for focal dose escalation in the FLAME trial. Because of the absence of contouring guidelines for prostate cancer inside the gland at the time of treatment, the accuracy of these delineations is uncertain. For this reason, we evaluated the delineations of the FLAME trial in the three institutes and observed statistically significant differences between institutes in the weighting of each of the MRI sequences during contouring of the visible cancer.

Guidelines for the detection and localization of prostate cancer were established in 2012 in PI-RADS and have been updated in version 2 in 2015 with a decreased relevance of DCE-MRI (Barentsz *et al.* 2012, Weinreb *et al.* 2016). Several studies have already demonstrated the need for contouring guidelines based on inter-observer studies (Bratan *et al.* 2013, Jung *et al.* 2013, Rischke *et al.* 2013, Anwar *et al.* 2014, Steenbergen *et al.* 2015). With established contouring guidelines, training of radiation oncologists could improve the accuracy of clinical tumor delineations on mp-MRI.

Using logistic regression, we now showed differences in interpretation of mp-MRI on an institutional level. We limited the logistic regression analysis to T2w, ADC and K^{trans} intensity images only. Nonetheless, other MRI features may be considered by the radiation oncologists and radiologists, but these are not made explicit. Structural appearance for example plays an important role as well. However, for the purpose of investigating differences between

institutes, the intensity values allow for a more objective analysis. Therefore, we considered the intensity images as most informative for contouring of the visible cancer in the clinic. Logistic regression performed on an increased number of features would furthermore increase the risk of correlations and compromise the separation between obtained coefficients.

In Table 1 significant differences between institutes are observed for age, iPSA, number and volume of delineated tumors, after Bonferroni correction of the significance level for 24 tests. The significant differences we found for the delineated number and volume of tumors are considered to be of clinical relevance, since they have direct impact on the escalated dose volume and thereby may propagate into tumor control. Age and iPSA were not considered during manual contouring of the tumor, so we regarded the clinical relevance of the significant differences observed for these patient characteristics to be limited.

Retrospective evaluation of 225 clinical delineations by a panel of experts is not feasible in practice. Instead, we applied a published tumor probability model to assess the delineations automatically (Dinh *et al.* 2017). Most models used in other studies are focused on the classification or risk stratification of prostate cancer (Vos *et al.* 2008, Langer *et al.* 2009, Ozer *et al.* 2010, Rouvière *et al.* 2012). Some that do predict tumor probability in the prostate use information from imaging only (Viswanath *et al.* 2012, Groenendaal *et al.* 2012). The TP model that we applied is comparable to the latter, but benefits from the inclusion of clinical information and knowledge of the prevalence of prostate cancer in the gland, that is also available to the radiation oncologist.

As a proof of principle, we reviewed the five patients that showed the largest discrepancies between model predictions and delineations based on their AUC value. We found low AUC values caused by TURP regions for two patients and inaccurate tumor delineations for another two. Also, biopsy information that was contradictive to MRI lowered the AUC for two patients. We found that all the selected patients indeed needed revision of the clinical delineations.

In patients treated with radiotherapy, no whole mount section histology information can be available to validate the delineations. Therefore, discrepancies between the TP model and clinical delineations can be reviewed by an expert, but confirmation with ground truth is not feasible. Discrepancies may be attributed to inaccurate delineations as well as to inaccuracy of the model. Since the model was trained on a multi-institutional dataset with comparable MRI scanners and scanning protocols as used in this study, we attempted to minimize the latter. Limitations in the MRI normalization method could translate into inaccurate model predictions, although the risk is small when the tumor accounts for more than 50% of the PZ area. Because of the absence of ground truth information and possible model inaccuracies we

treated the model as a screening tool to select patients that need further review by experts, rather than a method to improve or even substitute clinical delineations.

The absence of histology information also complicates the assessment of the impact of delineation variation on the delivered focal dose escalation. The high median AUC values suggest however that this impact may be small overall. Still insufficient dose escalation to the tumor tissue and increased dose levels to normal tissue in the prostate may have occurred in some instances.

Conclusions

In conclusion, we demonstrated inter-institutional differences in the interpretation of mp-MRI for delineation of the visible cancer. This supports the need for contouring guidelines for prostate cancer inside the gland. AUC values were high in general, suggesting good agreement between TP model predictions and clinical delineations. Observed discrepancies based on low AUC values were a clear indication for further review by experts.

3

CHAPTER 3

Knowledge-based assessment of focal dose escalation treatment plans in prostate cancer

Marcel A. van Schie
Tomas M. Janssen
Dave Eekhout
Iris Walraven
Floris J. Pos
Peter de Ruiter
Alexis N.T.J. Kotte
Evelyn M. Monninkhof
Linda G.W. Kerkmeijer
Cédric Draulans
Robin de Roover
Karin Haustermans
Martina Kunze-Busch
Robert Jan Smeenk
Uulke A. van der Heide

International Journal of Radiation Oncology • Biology • Physics, Volume 108, pp1055–1062 (2020)

Abstract

Purpose

In a randomized focal dose escalation radiation therapy trial for prostate cancer (FLAME), up to 95 Gy was prescribed to the tumor in the dose-escalated arm, with 77 Gy to the entire prostate in both arms. As dose constraints to organs at risk had priority over dose escalation and suboptimal planning could occur, we investigated how well the dose to the tumor was boosted. We developed an anatomy-based prediction model to identify plans with suboptimal tumor dose and performed replanning to validate our model.

Materials and methods

We derived dose-volume parameters from planned dose distributions of 539 FLAME trial patients in four institutions and compared them between both arms. In the dose-escalated arm, we determined overlap volume histograms and derived features representing patient anatomy. We predicted tumor $D_{98\%}$ with a linear regression on anatomic features and performed replanning on 21 plans.

Results

In the dose-escalated arm, the median tumor $D_{50\%}$ and $D_{98\%}$ were 93.0 and 84.7 Gy, and 99% of the tumors had a dose escalation greater than 82.4 Gy (107% of 77 Gy). In both arms organs at risk constraints were met. Five out of 73 anatomic features were found to be predictive for tumor $D_{98\%}$. Median predicted tumor $D_{98\%}$ was 4.4 Gy higher than planned $D_{98\%}$. Upon replanning, median tumor $D_{98\%}$ increased by 3.0 Gy. A strong correlation between predicted increase in $D_{98\%}$ and realized increase upon replanning was found ($\rho = 0.86$).

Conclusions

Focal dose escalation in prostate cancer was feasible with a dose escalation to 99% of the tumors. Replanning resulted in an increased tumor dose that correlated well with the prediction model. The model was able to identify tumors on which a higher boost dose could be planned. The model has potential as a quality assessment tool in focal dose escalated treatment plans.

Introduction

Focal dose escalation to the tumor in prostate cancer radiotherapy has been hypothesized to improve patient outcome without increasing acute and late toxicities (Lips *et al.* 2011). In the multicenter randomized Focal Lesion Ablative Microboost in prostatE cancer (FLAME) trial, patients in the dose-escalated arm received an escalated dose up to 95 Gy to the visible tumor. The aim of the trial was to increase the 5-year biochemical recurrence-free survival rate by 10%. To prevent increased toxicity compared to the standard arm, strict dose-volume constraints were imposed on the organs at risk (OARs). During treatment planning, these OAR constraints had priority over dose escalation. Indeed, no significant increase in toxicity was found up to two years after treatment (Monninkhof *et al.* 2018). Because of the OAR constraints being prioritized however, the planned dose escalation to the tumor was limited by the spatial separation between the tumor and the OARs. This raises the question as to how much dose escalation to the GTV was really achieved in the dose-escalated arm of the trial. For this reason, we investigated in the first part of this study to what extent a dose escalation to the visible tumor was realized via comparison of dose-volume parameters between both arms of the trial.

Besides patient anatomy, the degree of dose escalation to the tumor can also be affected by decisions made during optimization of the treatment plan. In clinical practice, it is difficult for a planner or radiation oncologist to assess if a treatment plan can be considered optimal. Over the past years Knowledge Based Planning (KBP) techniques have been introduced to enable automated plan quality assessment (QA) in radiotherapy (Wu *et al.* 2009, Moore *et al.* 2011, Appenzoller *et al.* 2012, Yuan *et al.* 2012, Good *et al.* 2013, Wang *et al.* 2013, Schreibmann *et al.* 2014, Nwankwo *et al.* 2015, Song *et al.* 2016, Shiraishi *et al.* 2016, Wall *et al.* 2018). These studies utilize a database of previously treated patients to guide treatment planning of a new patient, based on similarities of the patient's anatomy with existing ones in the database. For standard prostate treatment planning several KBP techniques were utilized to predict OAR dose from the patient anatomy (Yuan *et al.* 2012, Wang *et al.* 2017, Janssen *et al.* 2019). However, to date no KBP methods have been published to predict achievable focal dose escalation in prostate cancer. In the second part of the study we therefore developed an anatomy-based prediction model using our own database to predict the achievable dose in the tumor. We compared the predicted achievable tumor dose with the tumor dose realized in the clinical plans. We tested the validity of our model and the potential for a QA tool with a replanning of a subset of treatment plans based on our model's predictions.

In this work we present how much of the intended dose was actually planned, how much dose escalation could have been achieved, and how much of the predicted dose escalation could

be realized upon replanning. Focal dose escalation is a promising strategy in prostate cancer. By combining dosimetric evaluation with knowledge-based planning predictions, this study gives a comprehensive overview of the current feasibility and limitations of this dose escalation strategy for prostate cancer, as well as an indication of potential improvements that could be realized upon future clinical implementation.

Materials and methods

Patient characteristics

Data from 571 prostate cancer patients who participated in the randomized FLAME trial (clinicaltrials.gov NCT01168479) were used. All patients had biopsy-proven, clinically localized, intermediate or high-risk prostate cancer (Ash *et al.* 2000). Patients were excluded from the trial if they received previous pelvic radiation or underwent prostatectomy, if they had a World Health Organization (WHO) score > 2, an International Prostate Symptom Score (IPSS) ≥ 20, a transurethral resection of the prostate (TURP) less than three months prior to treatment, contraindications for MRI, or if they could not discontinue anti-coagulate usage which was required for implanting gold fiducial markers. We obtained approval from the institutional review board and written informed consent from all participating patients.

Patients were treated at four institutions: 320 patients at the University Medical Center Utrecht (UMCU), 93 at the University Hospitals in Leuven (UZL), 109 at the Netherlands Cancer Institute in Amsterdam (NKI), and 49 at the Radboud University Medical Center in Nijmegen (Radboudumc). At each institution patients were randomly and in a 1:1 ratio assigned to the standard and dose-escalated treatment arm. Treating physicians were not blinded for the randomization in order to evaluate and approve the treatment plans.

The primary endpoint of the trial was to achieve an increase in 5-year biochemical recurrence free survival rate of 10% among patients in the dose-escalated arm of the trial as compared to the standard arm patients. To identify biochemical recurrence, the Prostate Specific Antigen (PSA) level in the blood was measured twice per year, and biochemical recurrence was defined as a PSA rise of 2 n/mL above nadir PSA level, according to the Phoenix definition (Roach *et al.* 2006). Treatment-related acute and late toxicity, measures by the Common Toxicity Criteria for adverse events version 3.0 (CTCAE) (Trotti *et al.* 2003), as well as Quality of Life and disease-specific survival were secondary endpoints of the trial.

For this study we considered patients who were included in the per-protocol analyses of the trial (Monninkhof *et al.* 2018). Patients that did not receive the assigned treatment or decided to discontinue the treatment due to anxiety for increased toxicity in the dose-escalated arm were not included in the per-protocol analyses. From the patients eligible for the per-protocol analyses, we excluded three patients who were assigned to the standard treatment arm for which no bladder was delineated. In total 274 patients in the standard arm and 265 patients in the dose-escalated arm were available for analysis.

Treatment planning and delivery

All patients received a planning CT scan and a pretreatment multiparametric (mp-) MRI exam, including a T2-weighted, diffusion weighted imaging and dynamic contrast-enhanced sequence. The prostate gland was delineated on the T2-weighted MRI by a radiation oncologist. The clinical target volume (CTV) consisted of the prostate gland and, depending on the risk of tumor involvement, the seminal vesicles (SV). For patients who were randomized to the dose-escalated treatment arm, any tumor tissue in the CTV that was visible on the mp-MRI was contoured and defined as gross tumor volume (GTV). After registration of MRI to CT, target volumes and organs at risk (OARs) were defined and delineated. The planning target volume (PTV) was defined as the CTV with a margin of 5 – 8 mm, according to institutional practice. Based on negligible dosimetric impact of PTV margins around intraprostatic GTVs, in this trial no margins were applied to the GTV (van Haaren *et al.* 2009, Lips *et al.* 2009).

The study protocol prescribed a radiation dose of 77 Gy to the PTV in 35 fractions, with an integrated boost up to 95 Gy to the identified tumors of patients in the dose-escalated arm. Depending on institutional practice, 55 to 77 Gy was prescribed to the SV whenever it was included in the CTV. Dose constraints to the OARs followed institutional practice and applied to both arms of the trial. In addition, dose constraints of 77 Gy to 1 cc of the rectum and 80 Gy to 1 cc of the bladder were included. One institution applied an endorectal balloon to further reduce dose to the rectal wall (Radboudumc).

Among the participating institutions different Treatment Planning Systems (TPSs) and delivery techniques were used. The UMCU used PLATO (Nucletron, Veenendaal, The Netherlands) and Monaco (Elekta, Stockholm, Sweden) to generate 7-beam intensity modulated radiation therapy treatment plans for 126 and 183 patients, respectively. The UZL generated 2-arc volumetric-modulated arc therapy treatment plans with Eclipse (Varian, Palo Alto, CA). The NKI and Radboudumc used Pinnacle TPS (Philips Radiation Oncology Systems, Fitchburg, WI) to generate 1- or 2- arc volumetric-modulated arc therapy plans.

Dose evaluation

A prescription dose map was constructed using the CTV and GTV masks with corresponding prescription dose levels. All dose distributions were resampled to a 1 mm isotropic voxel grid. From the dose distributions we derived dose-volume parameters within the PTV, all GTVs in the prostate, the CTV minus GTV, the bladder and the rectum.

In both study arms we determined the near-maximum dose $D_{2\%}$ and high-dose volume $V_{107\%}$ in the CTV minus GTV. The $V_{107\%}$ was calculated as the volume percentage with a dose escalation above 107% of the prescribed 77 Gy (82.4 Gy). We chose to evaluate the GTV coverage in terms of CTV prescription dose, since the trial prioritized organ at risk sparing over achieving GTV coverage and therefore GTV coverage was not explicitly required. Furthermore, we derived the $V_{95\%}$ in the PTV and the near-maximum doses D_{1cc} and D_{2cc} in the bladder and the rectum. For plans in the dose-escalated arm, we evaluated to what extent we reached the prescribed dose escalation of 95 Gy. We determined the number of plans with a GTV $D_{50\%}$ and $D_{98\%}$ above 82.4 Gy. Statistically significant differences between both arms were examined with one-way ANOVA tests. Since we applied several tests, Bonferroni correction for multiple testing was used to correct the significance level.

At one institution (Radboudumc), an endorectal balloon was applied to reduce rectal wall dose and decrease inter- and intrafraction motion. For this relatively small patient group we merged rectal wall and balloon contours to represent the rectum, on which we report dose volume parameters to be in accordance with literature. We compared GTV and rectum dose-volume parameters between patients with and without endorectal balloon in situ to decide if both patient cohorts could be combined for development of a prediction model.

Another 25 patients from UMCU and NKI received adaptive treatment. For these patients a rigid registration of planning and adaptive CT scan was performed, and a weighted sum was applied to the co-registered planned dose distributions. The weights corresponded to the number of treatment fractions that each dose distribution was delivered. Three patients had a replanning CT. In addition to the rigid registration of CT scans, we extracted binary masks of prostate, bladder and rectum and pairwise deformably registered the masks between first and second planning session. The deformable registration involved an implementation of the b-spline deformation algorithm described by Rueckert *et al.* (1999). The normalized cross-correlation similarity measure was used for optimization, and registrations were visually assessed. Replanned dose distributions were mapped accordingly, resulting in locally deformed dose distributions for prostate, bladder and rectum. Planed dose distributions were weighted separately for prostate, bladder and rectum to allow for dose-volume parameter

derivation. In the second part of the study we only considered the initial treatment plans to develop a prediction model on.

Prediction model

We developed a prediction model that calculated the highest achievable $D_{98\%}$ in the GTV based on the anatomy of all patients. We chose to predict the near-minimum dose, as this was regarded to be most sensitive to trade-offs between OAR dose and tumor coverage.

We derived Overlap Volume Histograms (OVHs) of delineated PTV, GTV, bladder and rectum to encode the patient's anatomy (Wu *et al.* 2009, Yuan *et al.* 2012). We defined ten structure pairs (PTV→Bladder, PTV→Rectum, GTV→Bladder, GTV→Rectum, PTV→GTV, and vice versa) and derived the OVH of each structure pair on a 1 mm resolution. In case of multiple GTVs per patient, we performed our analysis per GTV to allow for a dose prediction per GTV. Per structure pair we combined the OVHs of all patients and performed Principal Component Analysis (PCA) to reduce dimensionality (Yuan *et al.* 2012). We determined the set of principal components (PCs) that described 90% of the variance in OVHs. We reconstructed the OVHs using the derived PCs and defined the obtained patient-specific coefficients as PC scores. In addition, we added radii $r_{5\%}$, $r_{50\%}$ and $r_{95\%}$ corresponding with 5%, 50% and 95% fractional overlap between two structures. In contrast to the PC scores these radii were only dependent on the patient's individual OVHs.

The model we developed combined automatic feature selection with a modified linear regression algorithm to predict the $D_{98\%}$ in the GTV. Given the complexity of a dose escalated treatment plan, it is difficult to manually assess if a treatment plan was made optimal in terms of highest GTV $D_{98\%}$ for a given set of anatomical constraints. Due to the large size of the study, we expected the plans in our dataset to range between not optimal and close to optimal planned dose distributions. Therefore, we modified the regression algorithm such that for treatment plans with similar anatomy, a larger weighting was applied to tumors with higher planned $D_{98\%}$.

Depending on the anatomy, values for planned $D_{98\%}$ are expected to lie in the range of 77 – 95 Gy. In some cases a GTV with a low $D_{98\%}$ may be optimal given the anatomy. To account for a non-uniform distribution of $D_{98\%}$ values over the dose range, we also applied a weighting of the planned $D_{98\%}$ that compensated for the sparsity of data points at lower dose. Details on the model's training and validation scheme can be found in the supplementary material.

To verify if inclusion of patients with endorectal balloon in situ did not bias the performance of the model, we retrained the model after exclusion of patients with balloon and compared the pairwise difference between the predicted GTV $D_{98\%}$ by the two models.

Evaluation of the model

We determined the dose difference between predicted and planned $D_{98\%}$ for all GTVs in the dose-escalated arm. We ranked the GTVs according to predicted dose difference in order to make a selection of treatment plans for replanning. We selected five treatment plans with the largest predicted dose difference and another 16 random plans: eight with a GTV with at least 10 Gy predicted dose difference and eight without. Among the largest predicted dose differences no bias towards any of the institutions was observed. Replanning was performed by planning specialists (DE, PR and RR) with ten, nine and three years of experience in treatment planning. The planning specialists were blinded for the predicted dose difference by the model, and instructed to plan the highest achievable dose to the GTVs while adhering to the existing target objectives and OAR constraints. New treatment plans were generated in the original treatment planning system, based on original delineations and according to the FLAME study treatment protocol. Because of decommissioning, seven treatment plans originally planned with the PLATO treatment planning system were replanned using Pinnacle. We compared our predicted tumor $D_{98\%}$ with the $D_{98\%}$ obtained upon replanning to evaluate our model. All analyses were performed in MATLAB (MathWorks, Natick, MA, USA).

Results

Dose evaluation

The observed dose-volume parameters for the PTV, GTV, CTV minus GTV, bladder and rectum are described in Table 1. The median $D_{50\%}$ to the GTV was 93.0 Gy, the median $D_{98\%}$ was 84.7 Gy. The percentage of GTVs that received a $D_{50\%}$ above 82.4 Gy (107% of 77 Gy) was 98.7%, and 70.4% received a $D_{98\%}$ above that level. Histograms of the distribution of GTV $D_{50\%}$ and $D_{98\%}$ in the dose-escalated arm are shown in Figure 1.

Table 1. *Comparison of dose-volume parameters in both arms of the FLAME trial. Median and IQR values are reported.*

Structure	Dose-volume parameters	Standard arm (n =274)	Dose-escalated arm (n =265)	P-value*
PTV	$V_{95\%}$ (%)	98.3 (95.5–98.8)	98.1 (95.3–98.7)	0.127
CTV – GTV	$D_{2\%}$ (Gy)	79.3 (78.8–79.8)	91.2 (88.6–92.7)	**< 0.001**
	$V_{107\%}$ (%)	0.7 (0.0–2.7)	25.9 (17.3–39.2)	**< 0.001**
GTV	$D_{50\%}$ (Gy)		93.0 (90.3–94.5)	
	$D_{98\%}$ (Gy)		84.7 (81.3–88.4)	
	$V_{95\%}$ (%)		77.6 (50.6–92.0)	
Bladder	D_{1cc} (Gy)	75.5 (74.4–76.7)	76.2 (75.0–77.6)	**< 0.001**
	D_{2cc} (Gy)	74.6 (73.7–76.0)	75.2 (74.0–76.6)	**0.009**
Rectum	D_{1cc} (Gy)	74.1 (73.5–74.8)	74.9 (73.7–75.9)	**< 0.001**
	D_{2cc} (Gy)	73.3 (72.5–74.0)	73.5 (72.4–74.4)	0.037

***** Differences were tested with a one-way ANOVA test. A post hoc Bonferroni method was applied to correct the significance level for multiple testing.

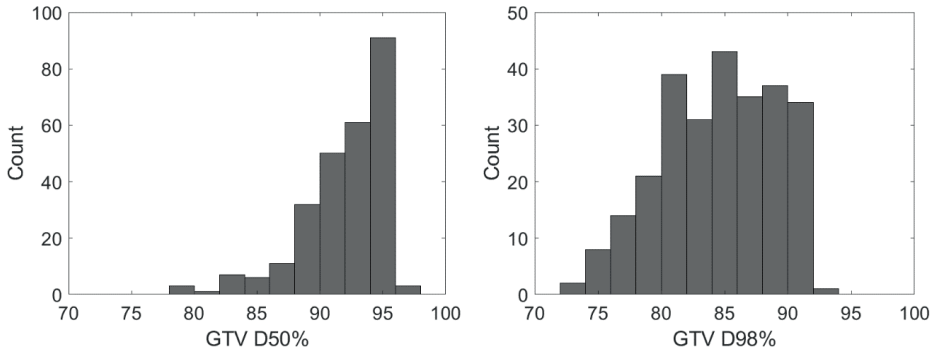


Figure 1. *Histograms of planned $D_{50\%}$ (left) and $D_{98\%}$ (right) of 265 patients in the dose-escalated arm.*

The median $V_{95\%}$ in the PTV was 98% in both study arms. The median near-maximum dose $D_{2\%}$ and high-dose volume $V_{107\%}$ in the CTV minus GTV were respectively 79.3 Gy and 0.7% in the standard arm, and 91.2 Gy and 25.9 % in the dose-escalated arm, and differed significantly between both arms. The difference is explained by dose gradients surrounding the GTVs in the dose-escalated arm. The median D_{1cc} in the bladder and rectum were 75.5 and 74.1 Gy in the standard arm, and 76.2 and 74.9 Gy in the dose-escalated arm, respectively. The median

bladder and rectum D_{2cc} were 74.6 and 73.3 Gy in the standard arm and 75.2 and 73.5 Gy in the dose-escalated arm, respectively.

Between patients treated with and without an endorectal balloon in situ we observed minor differences in GTV $D_{50\%}$ and $D_{98\%}$ that were non-significant, and small differences in rectum D_{1cc} and D_{2cc} that were significant but not exceeding clinical dose constraints. The results of this comparison are presented in the supplementary material, Table A2.

Prediction model

The model was trained on 382 GTVs. After PCA, four to five PCs were extracted from each OVH. The trained model consisted of five features, listed with corresponding coefficients in the supplementary material, Table A1. The predicted GTV $D_{98\%}$ is plotted against the planned $D_{98\%}$ in Figure 2. The influence of the larger contribution of plans with a higher planned GTV $D_{98\%}$ is reflected by the small fraction of data points above the identity line. Planned GTV $D_{98\%}$ values are observed up to 95 Gy, which reflects the aim of the trial. Predicted GTV $D_{98\%}$ in some cases however exceeded the 95 Gy, suggesting that according to the model the anatomy of these patients would allow for further dose escalation. In one extreme case a GTV $D_{98\%}$ of 104.6 Gy was predicted, which appeared to be a small tumor at a relative large distance from the rectum.

The median dose difference between predicted and planned $D_{98\%}$ was 4.4 Gy, and dose differences ranged between -2.8 Gy and 16.7 Gy. In 135 of 265 patients who received a focal dose escalation, at least for one GTV an achievable increase of 5 Gy was predicted.

Between the prediction models trained with and without inclusion of patients treated with endorectal balloon in situ, we observed a median pairwise difference of 0.0 Gy (95% confidence interval -0.4 – 0.1 Gy), which justified the inclusion of patients with balloon in our presented prediction model. The difference between predicted and planned $D_{98\%}$ of both models can be found in the supplementary material, as well as a scatter plot of the pairwise difference in predicted $D_{98\%}$ between the two models (Figure A2).

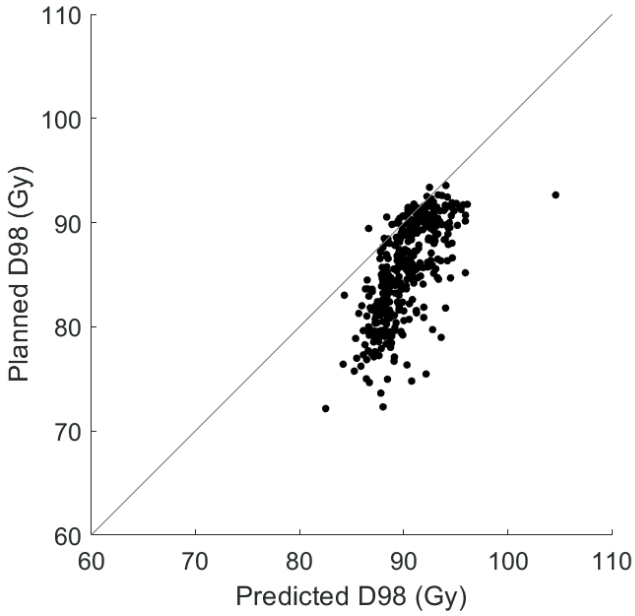


Figure 2. *Scatterplot of predicted $D_{98\%}$ versus planned $D_{98\%}$ in the GTV after training of the modified linear regression model. The majority of data points can be observed below the identity line as a result of the asymmetric cost function.*

Evaluation of the model

The 21 treatment plans selected for replanning involved 43 GTVs. We plotted the predicted GTV $D_{98\%}$ from our model versus the planned $D_{98\%}$ and the replanned $D_{98\%}$ in Figure 3. Before replanning we observed a median dose difference between predicted and planned GTV $D_{98\%}$ of 7.5 Gy (0.5 – 16.7 Gy). After replanning the median dose difference between predicted and replanned GTV $D_{98\%}$ was 3.8 Gy (-5.2 – 7.7 Gy). A strong correlation ($\rho = 0.86$) was observed between the predicted increase in $D_{98\%}$ and realized increase in $D_{98\%}$ after replanning.

A median increase from planned to replanned GTV $D_{98\%}$ of 3.0 Gy (-4.0 – 16.9 Gy) was found. For GTVs with a planned $D_{98\%}$ below 80 Gy we observed a median increase of 10.4 Gy. GTVs with a planned $D_{98\%}$ between 80 and 85 Gy had a median increase of 4.9 Gy, between 85 and 90 Gy a decrease of 0.7 Gy, and above 90 Gy a decrease of 1.7 Gy. Below 85.9 Gy, all GTV's

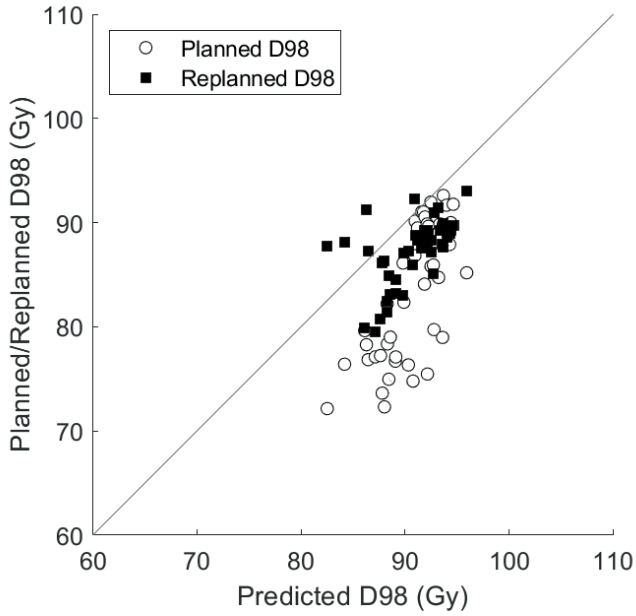


Figure 3. *Scatterplot of the predicted $D_{98\%}$ in the GTV versus the planned and replanned $D_{98\%}$. Upon replanning an increase in GTV $D_{98\%}$ can be observed which correlates with the predicted increase in $D_{98\%}$.*

$D_{98\%}$ increased upon replanning. In 16 out of 43 GTVs we observed a decreased $D_{98\%}$ after replanning, with a median decrease of 1.4 Gy (range 0.2 – 4.1 Gy). These GTVs had a relative high median $D_{98\%}$ of 90.1 Gy, which reduced the likelihood of improved tumor dose after replanning. For the five treatment plans that were selected based on largest predicted dose difference, the median difference between planned and replanned GTV $D_{98\%}$ was 9.0 Gy. For the 16 randomly selected treatment plans this was 1.7 Gy.

We observed comparable median PTV $V_{95\%}$ of 97.1% before and 97.9% after replanning. Median bladder D_{1cc} and D_{2cc} were respectively 77.0 and 75.8 Gy before, and 76.8 and 75.4 Gy after replanning, while median rectum D_{1cc} and D_{2cc} were respectively 74.3 and 73.3 Gy before, and 75.8 and 74.4 Gy after replanning. The small increase in dose to the rectum was expected to correlate with an increased dose to the GTV. Maximum bladder and rectum dose were still in accordance with clinical constraints.

Discussion

We showed that integrated focal dose escalation in the prostate is feasible with a median dose above 107% of the standard dose of 77 Gy achieved in 99% of the patients. Observed dose-volume parameters show a median GTV $D_{50\%}$ of 93.0 Gy, which was close to the intended 95 Gy, and a median $D_{98\%}$ of 84.7 Gy.

We also developed a prediction model based on overlap volume histograms and planned $D_{98\%}$ to identify GTVs for which a higher escalated dose was regarded feasible. After replanning of a subset of treatment plans, we observed a considerable increase in planned $D_{98\%}$ which strongly correlated with the predicted increase by the model.

A recent trial on dose painting in prostate cancer (HEIGHT trial) with up to 89.3 Gy in 38 fractions found a GTV $V_{95\%}$ (above 84.8 Gy) between 95.2% and 99.8% (Bossart *et al.* 2016). Because of the different fractionation scheme and level of dose escalation, a comparison with our results could not be made.

$D_{2\%}$ and $V_{107\%}$ in the CTV minus GTV as well as D_{1cc} in the bladder and the rectum showed a significant increase of dose in the dose-escalated arm, while D_{2cc} in the bladder and the rectum did not. These findings can partially be explained by the study protocol that allowed for dose escalation in the healthy prostate, provided that dose-volume constraints to OARs were not violated.

In the dose-escalated arm there were 382 GTVs in 265 plans, which on average was 1.4 GTVs per plan. These findings are in agreement with Van Schie *et al.* (2018). A higher average of 2.0 GTVs per plan was observed in the replanning selection of 21 plans. The higher average in the replanning selection can partly be explained with statistics since a plan with multiple GTVs had an a priori higher chance of inclusion in the replanning selection. We also observed an overestimation of the achievable $D_{98\%}$ as compared to the planned dose upon replanning. One explanation is the design of the trial, in which we aimed for an escalated dose up to 95 Gy. Our model however was not restricted by this dose constraint and, based on patient anatomy, could in principle predict a higher achievable escalated dose than 95 Gy. While both observations can partly be assigned to statistics and trial design, we do believe they are to some extent also explained by the limitation of our prediction model that did not consider the effect of multiple GTVs per prostate. The model determined the achievable GTV $D_{98\%}$ for each GTV individually, which could lead to violation of OAR dose constraints in case of multiple GTVs within the prostate. During replanning, this likely has resulted in a reduced GTV $D_{98\%}$, since OAR dose constraints were prioritized.

We demonstrated that focal dose escalation was achieved in almost all patients in the dose-escalated arm of the trial. Although several trials have hypothesized clinical benefits of focal dose escalation, no KBP methods exist to predict the highest achievable integrated boost dose to the tumor. Here we demonstrated a novel methodology that, using anatomical features and based on a heterogeneous dataset, could predict the highest achievable dose in the GTV and allowed to identify GTVs for which the escalated dose could be improved. A limitation of existing KBP methods is that the predicted dose range reflects the range of clinical plans. In our model we introduced an upward bias to predict the highest achievable dose, by putting extra weight on the better optimized plans in the database. Since it was trained on data from multiple institutions, the model is robust to different treatment planning systems. We recognize that our model does not allow for a precise estimation of the achievable tumor dose. We do however believe that our model can assist as a QA tool to identify GTVs that could be planned with a higher escalated dose.

Focal dose escalation is a promising dose escalation strategy in prostate cancer. By combining dosimetric evaluation with knowledge-based planning predictions we were able to demonstrate the feasibility of focal dose escalation up to 95 Gy in the prostate, as well as presenting a methodology to potentially improve on focal dose escalation treatment plans in a clinical setting. Although developed for a novel dose escalation strategy in prostate cancer, we believe our methodology can be of general applicability to other treatment sites and radiation strategies as well.

Conclusions

Focal dose escalation in prostate cancer was feasible in almost all GTVs, with an escalated dose well above the standard prescribed dose. We developed a prediction model to identify GTVs for which a higher escalated dose was considered achievable. Using this model to select plans for replanning, a considerable increase in $D_{98\%}$ was found achievable, specifically for lower planned $D_{98\%}$. Our prediction model has potential as a QA tool and identify suboptimal GTV doses to be optimized via replanning.

Supplementary Material

Prediction model

Materials and methods

To predict the highest achievable $D_{98\%}$ in the GTV, we developed a model based on patient anatomy, expressed by principal component (PC) scores and fractional overlap scores from patient's OVHs. We applied a stepwise multiple regression model (SMRM) to find the optimal combination of anatomical features to predict the $D_{98\%}$ (Yuan *et al.* 2012). We modified MATLAB's *stepwiselm* function available from the *Statistics and Machine Learning Toolbox*. The cost function was expanded with a penalty term such that for GTVs with comparable anatomy, the GTV with highest $D_{98\%}$ would have a larger influence on the regression. A weighting of the planned $D_{98\%}$ was added to compensate for sparsity of data at lower dose. The resulting asymmetric least squares optimization algorithm used the following cost function:

$$\text{Cost} = \sum_{i=1}^{n} \left(D_{\max} - D_{\text{plan,i}}\right)^2 \cdot \left(D_{\text{pred,i}} - D_{\text{plan,i}}\right)^2 \cdot \left(\text{sgn}\left(D_{\text{pred,i}} - D_{\text{plan,i}}\right) - \alpha\right)^2. \quad \text{(A1)}$$

The first term in this equation describes the weighting function, the second term corresponds to ordinary least squares optimization, and the third term penalizes predictions below the planned $D_{98\%}$. $D_{\text{plan,i}}$ and $D_{\text{pred,i}}$ are the planned and predicted $D_{98\%}$ of data point *i*, *sgn* is the sign function and α the asymmetry parameter. $D_{\max}$ was set to 102 Gy (107% of 95 Gy), α was set to 0.9.

We had 265 treatment plans with 382 GTVs available and trained the prediction model on the entire dataset. During training, feature set selection was based on 10-fold cross validation, evaluating the Cost in the validation set. Since the SMRM performs feature set selection based on a preset p-value, we tested a range of p-values (0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1). The range of p-values tested was determined during development of our model. For lower p-values the model would not include any feature, and for higher p-values calculation time increased drastically and resulted in decreased model's accuracy. For each p-value we identified the most frequently selected feature set and corresponding Cost. If feature sets were identical to the most frequent selected feature set, we selected the minimum of corresponding Costs. To be able to select the optimal p-value and optimal feature set, and increase robustness of the selection strategy, we repeated this procedure ten times. After ten repetitions we calculated the median Cost per p-value, and selected the p-value corresponding to the lowest median Cost. The final selected feature set was the feature set

that corresponded to the selected p-value and lowest Cost. Final weights were determined via fitting of the selected features to the complete data set using the asymmetric cost function.

To investigate the robustness of our training strategy, we also randomly designated 25% of the GTVs as test set, stratified by treatment plan. A schematic representation of our training strategy is depicted in Figure A1.

Results

During OVH calculation, from most OVHs four PCs were extracted that explained 90% of the variance in the OVHs. For the structure pairs PTV→Bladder, PTV→Rectum, and Bladder→GTV these were five PCs. Together with the 5^{th}, 50^{th} and 95^{th} OVH curve fractional overlaps scores, this yielded 73 anatomical features that were used to train the prediction model. A p-value of 0.01 resulted in the lowest Cost of 0.627 during training of the model.

In the robustness test we found an average Cost of 0.609 in the training data, and an average Cost of 0.748 in the test data, demonstrating our model was robust when applied to different datasets.

Endorectal Balloon

DVH parameters

In Table A2 dose-volume parameters in the GTV and rectum are compared between patients treated with and without endorectal balloon in situ. Significant dose differences are observed for the rectum, although D_{1cc} and D_{2cc} values were within accepted dose limits.

Prediction model

In Table A3 we reported the selected features of the prediction model after exclusion of patients treated with an endorectal balloon in situ. The observed median difference between predicted and planned GTV $D_{98\%}$ was 4.4 Gy (95% CI -0.2 – 10.7 Gy). For these patients, the median difference in the model trained on all patients was 4.3 Gy (95% CI of -0.3 – 10.4 Gy). The difference between predicted GTV $D_{98\%}$ by both models was minimal: 0.0 Gy (95% CI -0.4 – 0.1 Gy). The difference between the scatter plot distributions of both models is shown in Figure A2. Considering the minimal differences in GTV $D_{98\%}$ and scatter plot distributions, the model including all patients was favored for presentation in the article and selection of treatment plans for replanning.

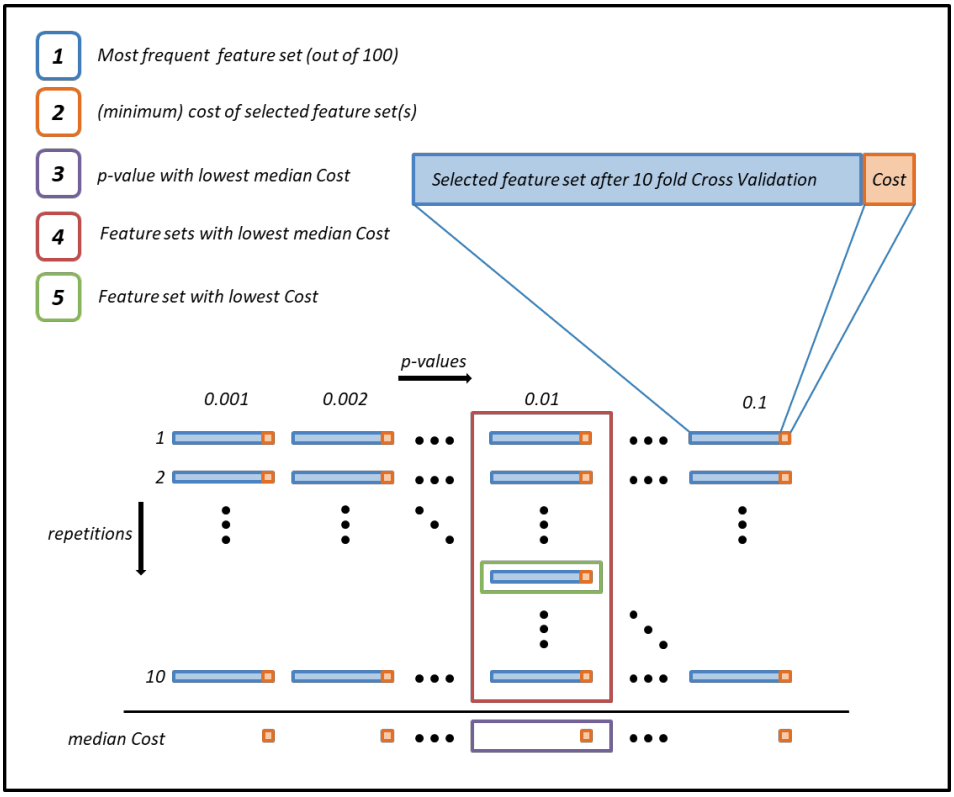


Figure A1. *Schematic representation of the training strategy of our prediction model. Steps in the training strategy are color coded in the top left corner.*

Table A1. *List of features in the trained prediction model, together with coefficients. $D_{98\%}$ was determined using $D_{98\%} = D_0 + \Sigma(\beta_i * x_i)$*

Description of selected features x_i	Coefficients β_i
Offset	66.96 Gy
PC2 of OVH Bladder → GTV	0.048
PC4 of OVH Bladder → GTV	-0.092
$r_{5\%}$ of OVH GTV → Rectum	0.350
PC2 of OVH Rectum → GTV	0.076
PC3 of OVH Rectum → GTV	-0.405

Table A2. *Comparison of dose-volume parameters between patients treated with and without endorectal balloon in situ. Median and IQR values are reported.*

Structure	Dose-volume parameter	Treatment arm	Without balloon (n=493)	With balloon (n=46)	P-value*
GTV	$D_{50\%}$ (Gy)		84.8 (81.4 – 88.4)	82.5 (79.2 – 88.0)	0.05
	$D_{98\%}$ (Gy)		92.8 (90.1 – 94.4)	93.9 (91.6 – 95.7)	0.29
Rectum	D_{1cc} (Gy)	Standard	74.1 (73.5 – 74.7)	75.1 (74.1 – 76.3)	**<0.001**
		Dose-escalated	74.8 (73.6 – 75.9)	75.3 (74.8 – 76.0)	0.14
	D_{2cc} (Gy)	Standard	73.2 (72.5 – 73.8)	74.7 (73.7 – 76.0)	**<0.001**
		Dose-escalated	73.3 (72.3 – 74.3)	74.6 (73.9 – 75.1)	**0.001**

***** Differences were tested with a one-way ANOVA test. A post hoc Bonferroni method was applied to correct the significance level for multiple testing.

Table A3. *List of features and corresponding coefficients in the trained prediction model using patients treated without endorectal balloon in situ.*

Description of selected features x_i	Coefficients β_i
Offset	66.57 Gy
PC2 of OVH Bladder → GTV	0.050
PC4 of OVH Bladder → GTV	-0.111
$r_{5\%}$ of OVH GTV → Rectum	0.358
PC3 of OVH Rectum → GTV	-0.377

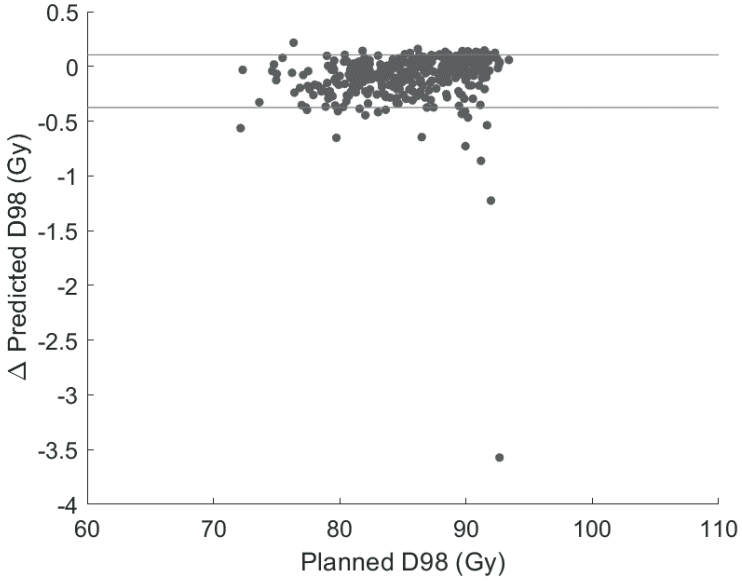


Figure A2. *Scatter plot of the planned GTV $D_{98\%}$ versus the difference in predicted GTV $D_{98\%}$ between the trained prediction models using all patients and only patients treated without endorectal balloon in situ. A minimum difference of -3.6 Gy was observed, corresponding to the extreme case in Figure 2. Due to the relative small size and large distance from the rectum, this GTV was sensitive to the composition of the training dataset.*

4

CHAPTER 4

Repeatability of dose painting by numbers treatment planning in prostate cancer radiotherapy based on multiparametric magnetic resonance imaging

Marcel A. van Schie
Peter Steenbergen
Cuong V. Dinh
Ghazaleh Ghobadi
Petra J. van Houdt
Floris J. Pos
Stijn W.T.J.P. Heijmink
Henk G. van der Poel
Steffen Renisch
Torbjørn Vik
Uulke A. van der Heide

Physics in Medicine & Biology, Volume 62, pp5575–5588 (2017)

Abstract

Purpose

Dose Painting by Numbers (DPbN) refers to a voxel-wise prescription of radiation dose modelled from functional image characteristics, in contrast to dose painting by contours which requires delineations to define the target for dose escalation. The direct relation between functional imaging characteristics and DPbN implies that random variations in images may propagate into the dose distribution. The stability of MR-only prostate cancer treatment planning based on DPbN with respect to these variations is yet unknown. We conducted a test-retest study to investigate the stability of DPbN for prostate cancer in a semi-automated MR-only treatment planning workflow.

Materials and methods

Twelve patients received a multiparametric MRI on two separate days prior to prostatectomy. The tumor probability (TP) within the prostate was derived from image features with a logistic regression model. Dose mapping functions were applied to acquire a DPbN prescription map that served to generate an intensity modulated radiation therapy (IMRT) treatment plan. Dose calculations were done on a pseudo-CT derived from the MRI. The TP and DPbN map and the IMRT dose distribution were compared between both MRI sessions, using the intraclass correlation coefficient (ICC) to quantify repeatability of the planning pipeline. The quality of each treatment plan was measured with a quality factor (QF).

Results

Median ICC values for the TP and DPbN map and the IMRT dose distribution were 0.82, 0.82 and 0.88 for linear dose mapping, and 0.82, 0.84 and 0.94 for square root dose mapping. A median QF of 3.4% was found among all treatment plans.

Conclusions

We demonstrated the stability of DPbN radiotherapy treatment planning in prostate cancer, with excellent overall repeatability and acceptable treatment plan quality. Using validated tumor probability modelling and simple dose mapping techniques it was shown that despite day-to-day variations in imaging data still consistent treatment plans were obtained.

Introduction

In prostate cancer radiotherapy dose escalation to the tumor based on image characteristics is referred to as dose painting. While dose painting is not a standard procedure in prostate cancer treatment (Bauman *et al.* 2013), it has been hypothesized that a clinical benefit without increased toxicity can be achieved with this approach (Pickett *et al.* 1999, van Lin *et al.* 2006, Singh *et al.* 2007, Fonteyne *et al.* 2008). The clinical benefit of focal dose escalation is currently investigated in the FLAME trial, a multicenter phase III randomized clinical trial (Lips *et al.* 2011). In dose painting by contours tumors are manually delineated based on multiparametric (mp-) MRI, consisting of a T2-weighted (T2w), diffusion-weighted (DWI) and a Dynamic Contrast-Enhanced (DCE) sequence (Barentsz *et al.* 2012, Dickinson *et al.* 2013). However, it has been demonstrated by Steenbergen *et al.* (2015) that a large inter-observer variation in manual tumor delineations on mp-MRI exists. Since delineation of the tumor is an inherently binary procedure, variability in tumor delineations may have an impact on the dose coverage of the actual tumor.

As an alternative to manual contouring automated methods to derive a tumor probability (TP) map of the prostate have been developed (Groenendaal *et al.* 2012, Viswanath *et al.* 2012, Vos *et al.* 2012, Dinh *et al.* 2016, Dinh *et al.* 2017). These maps represent the likelihood of tumor presence in each voxel. We showed earlier that this tumor probability correlates with the level of consensus among observers (Dinh *et al.* 2016). Dose Painting by Numbers (DPbN) refers to the concept of deriving a prescription dose from image characteristics directly. For this, dose mapping functions converting characteristics into a dose prescription need to be applied (Bowen *et al.* 2009).

In the current study we incorporated DPbN and tumor probability modelling in the radiotherapy treatment planning pipeline of prostate cancer. Based on dose prescription maps treatment plans were realized with conventional planning objectives and dose-volume constraints. Additional planning objectives allowed us to modulate the heterogeneous dose distribution within the prostate.

While DPbN does not suffer from the inter-observer variation of manual contouring, the MRI data are affected by image noise and day-to-day patient variation. The direct relationship between functional imaging characteristics and dose painting by numbers implies that such variation in images may propagate into the dose distribution. The impact of the image data variations and further processing on the treatment plan quality in DPbN is yet unknown. Therefore, we performed a test-retest mp-MRI study to investigate whether variability due to

imaging noise propagates through the planning pipeline and influences the repeatability of semi-automated DPbN treatment planning in an MR-only workflow.

Materials and methods

The repeatability of the proposed semi-automated MR-only workflow was investigated with a test-retest study, scanning each patient two times on a different day with a mp-MRI. For each mp-MRI exam a radiation treatment plan was simulated via four planning stages. A repeatability analysis between both treatment planning sessions was performed for each of the stages in this workflow. The images of the first and second MRI exams were registered with a B-spline deformable image registration.

Patient inclusion

We included twelve patients with biopsy-proven stage T1-T2 prostate cancer (median age 67, range 54 – 71 years) between October 2014 and March 2016. Each patient underwent two mp-MRI examinations with a median interval of 21 days (range 7 – 37 days), in which no change in tumor appearance was assumed. The second MRI examination was followed by a radical prostatectomy. Institutional review board approval was obtained, and all patients provided written informed consent.

Image acquisition

Images were acquired on a 3.0 T Philips Achieva MR scanner (Philips Healthcare, Best, the Netherlands). We scanned nine patients with a 6-channel phased-array coil in combination with an endorectal coil. After a dStream upgrade of the scanner, we scanned the last three patients with a 16-channel dS anterior coil and a 12-channel dS posterior coil.

In accordance with recent recommendations (Barentsz *et al.* 2012, Dickinson *et al.* 2013), the mp-MRI exam included an axial, sagittal and coronal T2-weighted turbo spin echo sequence, a diffusion weighted single-shot echo-planar imaging sequence, and a DCE scan using a 3D spoiled gradient echo sequence. A pre-contrast T1-weighted (T1w) 3D gradient echo sequence

was used to detect hemorrhage areas as a result of preceding diagnostic biopsies. For dose calculations we scanned an mDIXON sequence prior to contrast, from which a pseudo-CT scan was derived using the MRCAT algorithm (Philips Medical Systems MR Finland, Vantaa, Finland).

T2w scans were acquired with a TE/TR of 12/2000–5000 ms, and a reconstructed voxel size of 0.27 x 0.27 x 3.0 mm^3 before and 0.4 x 0.4 x 3.0 mm^3 after dStream upgrade. T1w scans were obtained with a TE/TR of 1.8/3.7 ms, while the mDIXON sequence had a TE1/TE2/TR of 1.16/2.4/3.7 ms. The DWI sequence had a TE/TR of 59/3500 ms, using b-values of 200, 600 and 1000 s/mm^2 to derive an apparent diffusion coefficient (ADC) map using a mono-exponential model (Bammer 2002). The DCE-MRI had a TE/TR of 1.9/4–5 ms and was acquired at a flip angle of 20°, after injection of 15 ml 0.5 M gadolinium-based contrast agent (DOTAREM®, Guerbet, Paris, France). In total the DCE sequence consisted of 110 acquisitions with an interval of 2.7 – 2.9 s. Signal intensities were converted to concentration values using a baseline T1-map obtained from variable flip angle scans at 3°, 6°, 10°, 20° and 30°. (Schabel *et al*. 2008). Volume transfer constant (K^{trans}) values were estimated with the Tofts model using a population-based arterial input function (Tofts *et al.* 1999, Murase 2004).

Any organ motion that occurred between the different mp-MRI sequences was corrected with a rigid registration to the prostate on the axial T2w scan.

Planning pipeline

Data from each examination was processed separately through the planning pipeline. The prostate gland and organs at risk (OARs) were delineated manually using the T2w and pseudo-CT images. In total 30 image features from T2w, ADC and K^{trans} were combined to derive a per-voxel TP map for the prostate using a logistic regression model (Dinh *et al.* 2016, Dinh *et al.* 2017). Prior knowledge about tumor location was included in the TP model in the form of a tumor prevalence map, derived from radical prostatectomy patients (Ou *et al.* 2009). The TP model was trained previously on data of 17 patients from the same institute, and was validated with pathology data using a leave-one-out approach. After resampling of the TP map to a 2-mm cubic voxel grid compatible with treatment planning software, polynomial dose mapping functions were applied to translate the TP map to a dose prescription map (DPbN map).

The pseudo-CT, OAR structure set and DPbN map were imported in the treatment planning software to generate an intensity modulated radiation therapy (IMRT) treatment plan, with heterogeneous dose prescription to the prostate. A schematic overview of the data flow from mp-MRI examination to treatment plan realization is depicted in Figure 1.

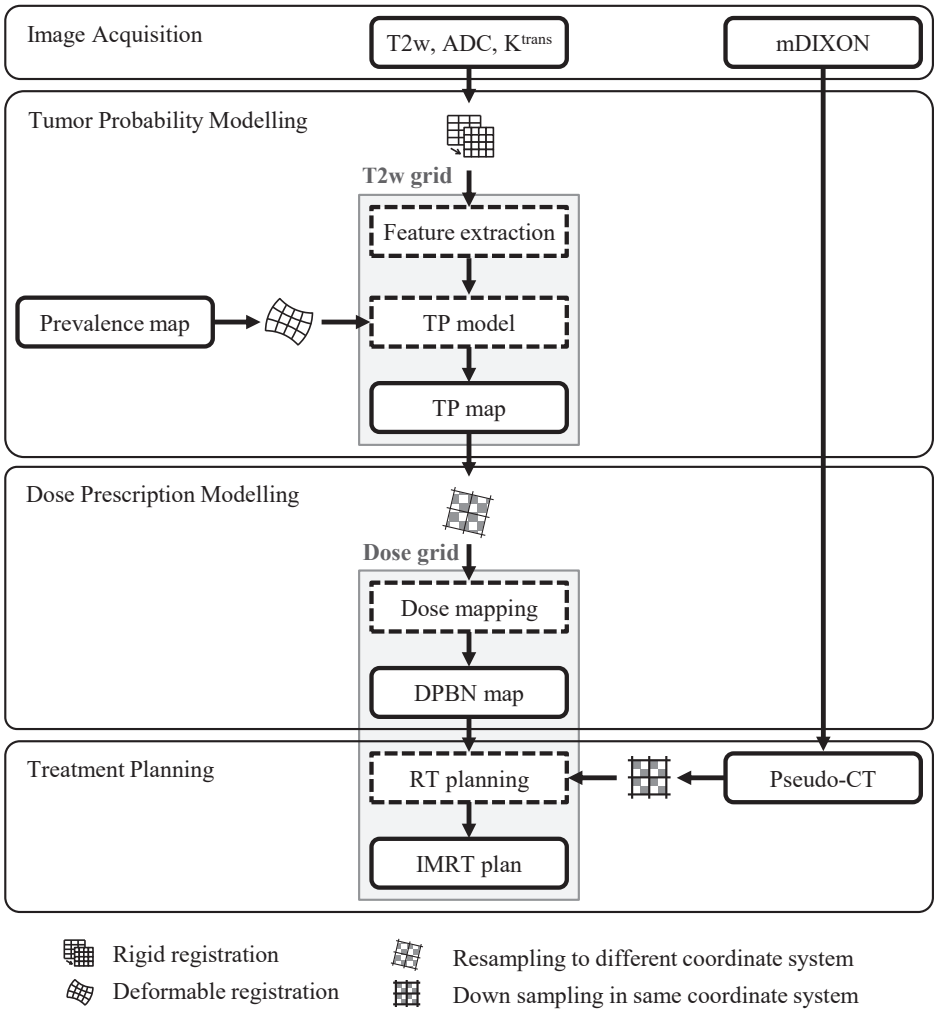


Figure 1. *Schematic overview of the planning pipeline in DPbN prostate cancer treatment. The treatment planning pipeline involves four stages: image acquisition, tumor probability modelling, dose prescription modelling and treatment planning. Processing steps are indicated with dashed contours. Registration and resampling steps to different coordinate systems are indicated with symbols.*

Dose prescription modelling

For each voxel in the prostate the TP was translated into a prescription dose D_{presc} using a polynomial dose mapping function (Bowen *et al.* 2009):

$$D_{presc}^{n} = D_{\min} + \left(D_{\max} - D_{\min}\right) \cdot TP^{n}, \quad (1)$$

with D_{min} and D_{max} the minimum and maximum prescribed dose respectively, and n the polynomial order of the mapping function. DPbN prescription maps were created with values for n of 1 and 0.5, corresponding to linear and square root mappings of TP. D_{min} was set to a safe lower bound of 68 Gy, which reflects the standard treatment in the Dutch dose escalation trial (Peeters *et al.* 2006). A D_{max} of 102 Gy was allowed, which is 107% of 95 Gy and corresponds to the escalated dose to the visible tumor in the FLAME trial (Lips *et al.* 2011). D_{presc} was calculated on a 2-mm cubic voxel grid to be compatible with the treatment planning software. The dose prescription outside the prostate was set to 0 Gy. Modeling of TP and DPbN maps was performed with MATLAB (version R2015b, MathWorks, Natick, MA).

Treatment planning

Apart from the prostate, the following OARs were delineated on the mDIXON images according to the clinical guidelines of our institute: the rectum (up to either the sigmoid or pelvic joints) and anal sphincter, the femoral heads (including articular cartilage) and the bladder. The pseudo-CT, DPbN map and the OAR structure set were imported in a research version of Pinnacle planning software (Pinnacle 9.710, Philips Research, Hamburg, Germany) to establish an IMRT treatment plan. The treatment plan was optimized for a 10 MV step-and-shoot photon beam at seven angles: 210°, 260°, 310°, 0°, 50°, 100°, 150°. The minimum segment area of the multi-leaf collimator was set to 9 cm^2, the minimum number of monitor units per segment to four and the maximum number of segments per beam position to 10. Plan optimization with a maximum of 100 iterations was performed using a standard cost function in combination with two dose painting objective functions f_{min} and $f_{max,}$ summing over all voxels i in the prostate to penalize under- and overdosing of the prescribed dose respectively:

$$f_{\min}\left(D_{plan}\right) = w_{\min} \cdot \sum_i \min\left(\frac{D_{plan,i} - D_{presc,i}}{D_{presc,i}}, 0\right)^2$$
$$f_{\max}\left(D_{plan}\right) = w_{\max} \cdot \sum_i \max\left(\frac{D_{plan,i} - D_{presc,i}}{D_{presc,i}}, 0\right)^2 . \quad (2)$$

Here $w_{\min}$ and $w_{\max}$ are tunable weight factors, and $D_{\text{presc,i}}$ and $D_{\text{plan,i}}$ are the prescribed and planned dose to voxel *i*. The treatment plan was accepted if the dose-volume constraints for OARs in Table 1 were met.

Evaluation of the treatment plan was performed with a quality factor (QF), summing over the normalized absolute differences between planned and prescribed dose to each voxel *i* (Vanderstraeten *et al*. 2006):

$$QF = \frac{1}{n}\sum_i \left|\frac{D_{plan,i} - D_{presc,i}}{D_{presc,i}}\right| \cdot 100\% . \quad (3)$$

To get information about the capability of the treatment planning system to deliver high dose to small regions, a $QF_{95\%}$ was also calculated, which evaluates the treatment plan quality similar to equation (3), but only for voxels that were prescribed at least 95% of the maximum prescribed dose. Within the group of patients we tested both QF and $QF_{95\%}$ for significant difference between linear and square root dose mapping functions and between the first and second planning session with the paired Wilcoxon signed rank test at a 5% significance. We also performed a cross evaluation of the treatment plan quality with xQF and $xQF_{95\%}$, where the dose distribution of the treatment plan from the first planning session was compared with the prescribed dose distribution of the second session and vice versa. This allows to quantify

Table 1. *List of OARs that were delineated prior to treatment planning, together with the dose-volume constraints that were imposed to the treatment planning system to build a clinically acceptable treatment plan.* $V_{X\,Gy}$ *refers to the volume receiving X Gy.*

Organ at Risk	Dose-volume constraint
Rectum + 2 mm	$V_{80\,Gy} \leq 1$ cc
Rectal wall	$V_{64\,Gy} \leq 35$ %
	$V_{75\,Gy} \leq 10$ %
Anal sphincter	$D_{mean} \leq 45$ Gy
Bladder	$V_{80\,Gy} \leq 1$ cc
Femoral head	$D_{max} \leq 50$ Gy
Bowel loop	$D_{max} \leq 68$ Gy

to what extent treatment planning based on a different mp-MRI compromises the plan quality. Treatment plan quality (Q) was visualized with a Q-volume histogram (QVH), displaying Q-values defined as the ratio of planned over prescribed dose.

Between-session registration

For voxel-level evaluation of the TP, DPbN and IMRT map within the prostate a registration between both scan sessions is required. A B-spline deformable registration between the delineated prostates on the axial T2w scans of both mp-MRI sessions was used to match and resample the TP, DPbN and IMRT maps from the first planning session to the second. As a consequence of the registration step the DPbN and IMRT maps were upsampled to the higher resolution T2w grid, thereby introducing correlated data. We therefore resampled all maps again to the lower resolution 2-mm cubic voxel grid used for the treatment planning.

Repeatability analysis

Stability and repeatability of the planning pipeline was investigated on the TP, DPbN and IMRT map with the intraclass correlation coefficient (ICC), to quantify the variability between voxels relative to the measurement error (de Vet *et al.* 2006). In this study a one-way random model was used to measure consistency between two measurements, where single values were calculated for each individual patient (Shrout *et al*. 1979, McGraw *et al*. 1996, Raunig *et al.* 2015). The one-way random model was chosen since it effectively models the between-voxels variation as a fixed effect. This holds for an ICC calculated on a single patient where all voxels of interest of the patient are included instead of a random selection. To satisfy the requirement of normally distributed data, repeatability of the TP map was assessed on the logarithm of the TP odds ratio. Hemorrhage areas as a result of biopsies taken prior to the first MRI examination were identified on T1w images. Since these areas shrink from first to second MRI exam, we excluded these regions from the repeatability analysis. In order to investigate to what extent the variability of the mp-MRI image features propagates through the planning pipeline, the repeatability of the 30 individual imaging features that were input for the TP model was assessed with the ICC as well.

Results

To illustrate the planning pipeline, an example of the tT2w, ADC and K^{trans} intensity maps together with the calculated TP, DPbN and IMRT map is given in Figure 2. Slices are taken in the axial direction of the patient from a location in the prostate involving an area suspected of being tumor. The top row depicts the intensity maps of T2w, ADC and K^{trans}. The T2w map was normalized to the 75th percentile of the intensity values within the prostate. Within the prostate, image features from the intensity maps were combined to create the TP map in Figure 2D. A suspicion of tumor can be identified at the left side of the prostate, characterized by low T2w and ADC, and high K^{trans} values. The increased K^{trans} values at the transition zone of the prostate was disproven to be tumor from histopathology, in accordance with the TP map in panel B.

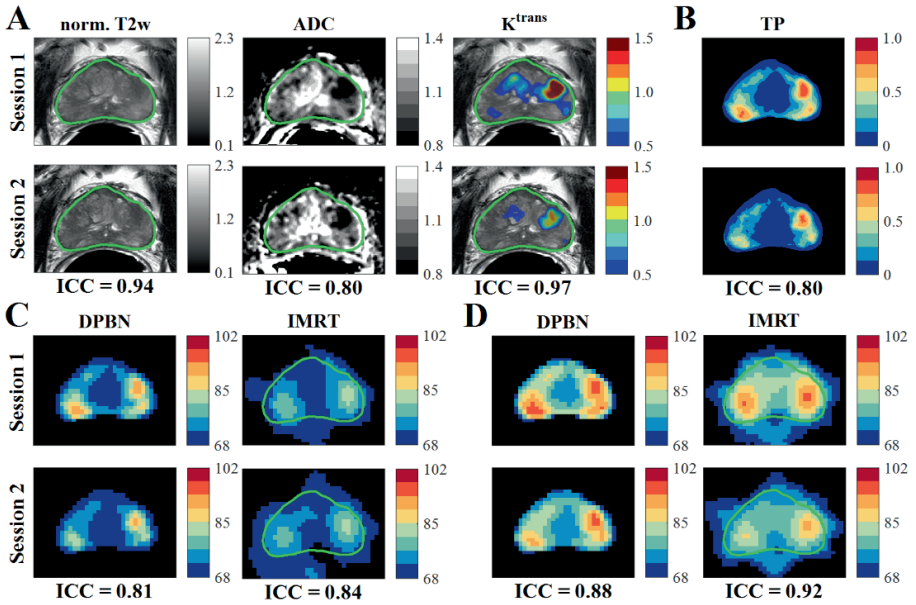


Figure 2. Example of an axial slice through the prostate. In each panel A-D the first (top row) and second planning session (bottom row) are compared. The ICC value is shown for each set of maps. Panel A depicts the mp-MRI: normalized T2w, ADC (in 10^{-3} mm^2/s) and K^{trans} map (in min^{-1}), where K^{trans} is shown as a colorwash on top of the normalized T2w scan. The DPbN map (in Gy) and IMRT plan (in Gy) are shown for linear (panel C) and square root (panel D) dose mapping functions. The delineated prostate is contoured in green.

After resampling of the TP map to the 2-mm cubic dose grid a linear and square root dose mapping function were applied to obtain the DPbN map. TP values of 0 and 1 were mapped to prescription dose levels between 68 and 102 Gy respectively. An example of the two dose mapping functions is shown in Figure 3. The TP histogram at the top shows a majority of voxels with low TP, and only a small fraction of voxels with TP above 0.5. In this example the maximum TP is 0.93, corresponding with 99.0 Gy for linear and 100.5 Gy for square root dose mappings.

An example of IMRT treatment plans in both planning sessions is given for linear dose mapping (Figure 2C) and square root dose mapping (Figure 2D). Smoothing of the prescribed dose due to scatter of photons and dose deposition limitations can be observed in the optimized plan. The quality of the IMRT plan was assessed with QF and $QF_{95\%}$ and visualized with a QVH in Figure 4. The majority of planned dose voxels was between 90% and 110% of the prescribed dose values, as can be observed from the gradient between Q-values of 0.9 and 1.1. For linear and square root dose mappings respectively, 3.4% and 1.5% of the prostate volume received less than 90% of the prescribed dose, while overdosing with more than 110% of the prescribed dose occurred in 2.3% and 6.2% of the voxels. QFs of 3.7% and 4.4% were observed, which is within the 5% treatment planning goal proposed by Duprez *et al.* (2011). $QF_{95\%}$ were 13.8% and 7.5%, indicating a higher agreement of planned with prescribed dose for square root dose mapping.

Comparable QVHs and quality factors were obtained for the other patients in this study, showing good quality of treatment plans based on DPbN. For linear dose mapping, QFs ranged between 2.3% and 4.1% (median 2.9%) and $QF_{95\%}$ values ranged between 5.5% and 17.3% (median 11.6%). For square root dose mapping, QFs between 2.7% and 4.4% (median 3.4%) were observed, whereas $QF_{95\%}$ values were between 4.1% and 11.7% (median 7.2%). Statistical testing with a Wilcoxon signed rank test of both QF and $QF_{95\%}$ revealed a significant difference between linear and square root dose mapping ($p < 0.001$). No significant differences between plans from session 1 and plans from session 2 were observed ($p = 0.80$ for QF and $p = 0.98$ for $QF_{95\%}$). xQF values ranged between 2.1% and 4.4% (median 3.3%) for linear dose mapping, and between 2.9% and 4.4% (median 3.9%) for square root dose mapping. $xQF_{95\%}$ values for linear dose mapping were between 4.8% and 18.1% (median 11.6%), and for square root dose mapping between 3.2% and 15.7% (median 8.0%).

Stability of automated treatment planning based on DPbN was tested with a repeatability analysis and expressed with the ICC. For ICC classification we considered excellent (above 0.75), good (0.6 – 0.75), fair (0.4 – 0.59) and poor (below 0.4) repeatability (Cicchetti *et al*. 1981). The evaluation of TP and DPbN map and IMRT dose distributions of the whole patient

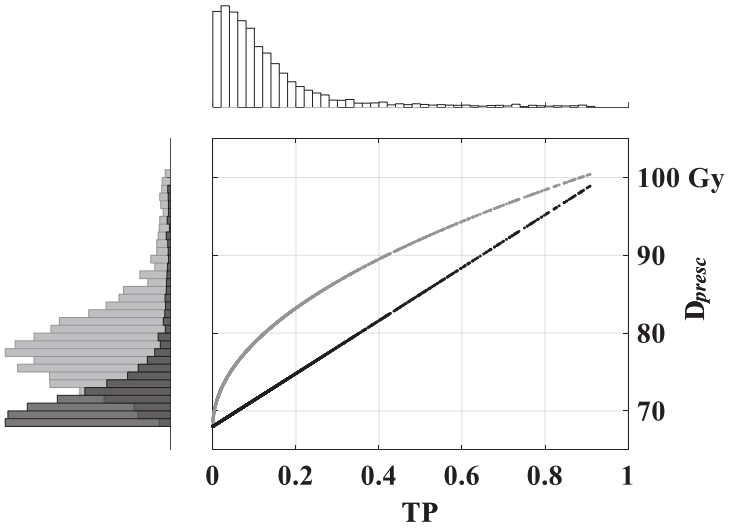


Figure 3. *Example of a linear (black) and square root (grey) dose mapping from tumor probability TP to prescribed dose D_{presc}. Normalized histograms for TP and D_{presc} are shown along the axis indicating the density of data points in the plot. Prescription dose values ranged from 68.0 to 99.0 and 100.5 Gy for linear and square root mappings.*

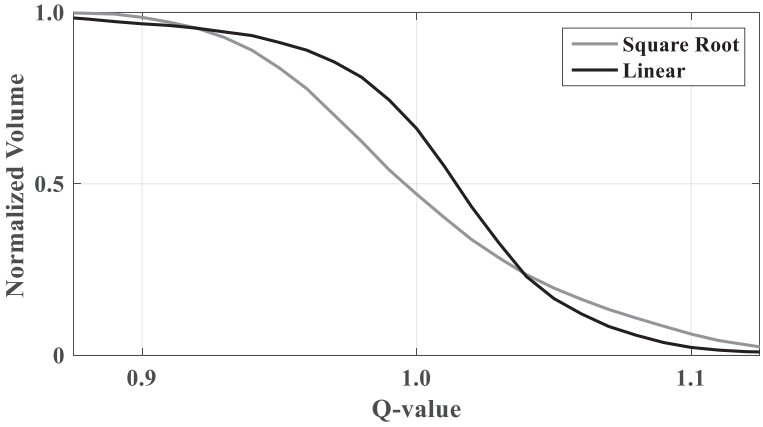


Figure 4. *Example of a QVH. Inverse cumulative histogram of the prostate volume receiving fraction Q of the prescribed dose. For linear dose mapping Q-values below 0.9 are observed in 3.4% of the prostate, and above 1.1 in 2.3% of the prostate. For square root dose mapping 1.5% of the prostate voxels had Q-values below 0.9, while 6.2% of the voxels had Q-values above 1.1.*

group with ICC resulted in excellent agreement at all three stages. Figure 5 shows the ICC values with median values of all patients for linear and square root dose mapping. The median ICC of TP maps was 0.82, ranging from 0.67 to 0.93. Median ICC values of 0.82 and 0.84 were found for the DPbN maps, using linear and square root dose mapping respectively, and ICC values ranged from 0.71 to 0.90. At IMRT stage a median ICC value of 0.88 was observed for linear dose mapping (range 0.71 – 0.93), while for square root dose mapping this was 0.94 (range 0.88 – 0.96). On individual level good to excellent agreement between both planning sessions was shown. Density scatter plots of the IMRT dose values based on both linear and square root dose mapping are provided in Figure 6. Dose values of the first planning session are plotted versus values from the second session and are accumulated over all patients. We observed a symmetric distribution around the diagonal for both dose mapping functions, with 95% of the voxel-to-voxel dose differences below 4.8 and 5.7 Gy for linear and square root dose mapping, respectively. To improve visualization of the low-count pixels in the scatter plot, we log transformed the intensity values.

Median ICC values for the repeatability of the 30 imaging features over the group of patients varied between 0.82 and 0.98. The five most repeatable features were all T2w Gaussian smoothed derivatives of the first (G_y) and second (G_{yy}) order, where y denotes the AP direction of the prostate. The ICC values of these five features ranged between 0.97 – 0.98: $G_y(3.8)$, $G_y(2.4)$, $G_y(6.0)$, $G_{yy}(6.0)$ and $G_{yy}(3.8)$. Here the values between brackets represent the scale (in mm) of the smoothing kernel. The five best performing features according to Dinh *et al.*

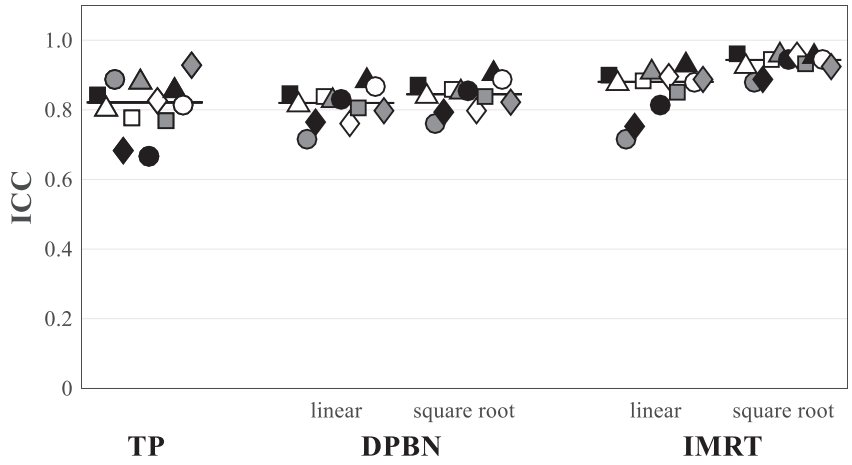


Figure 5. *ICC values of twelve patients at three stages of the treatment planning pipeline for both linear and square root dose mapping. Median ICC per stage is indicated with a horizontal bar. For each patient, identified by symbol-filling combination, repeatability can be traced through the planning pipeline.*

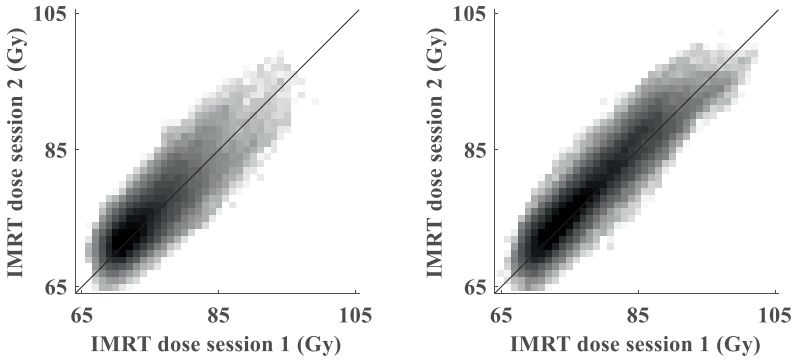


Figure 6. *Density scatter plots of dose values within the IMRT plan based on linear (left) and square root (right) dose mapping. Dose values are accumulated over twelve patients and plotted as session 1 versus session 2. Intensity values were log transformed to improve visualization of the low-count pixels.*

(2017) were ADC intensity, Prevalence map, K^{trans} intensity, $G_{xx}(6.0)$ and $G_{xx}(1.5)$, where x denotes the LR direction of the prostate. ICC values of these features were 0.84, 0.94, 0.85, 0.92 and 0.94, respectively.

Discussion

We investigated the repeatability of mp-MRI-based DPbN treatment planning for prostate cancer. Results from ICC analysis showed excellent repeatability for TP modelling, dose prescription and treatment planning, and QF analysis revealed good agreement between prescribed and planned dose. Simple polynomial dose mapping functions resulted in realistic prescription dose maps, and high repeatability of IMRT planning was observed. These results confirm that prostate cancer radiotherapy based on DPbN leads to stable treatment plans.

Dose prescription maps were derived from the TP maps using a polynomial dose mapping function. In this work we implemented linear and square root dose mappings, and left out squared dose mapping as described by Bowen *et al.* (2009). We calculated a prescription dose map for one patient based on squared dose mapping, but we observed a mapping of more than 90% of the voxels to dose values between 68 and 70 Gy. This almost homogenous dose

distribution can be explained with the skewed TP histogram, which results in an even more skewed prescription dose histogram with a squared dose mapping. For this reason, the squared dose mapping was not included in the test-retest study.

DPbN dose prescriptions with linear and square root dose mappings resulted in comparable repeatability of dose prescription maps. We found that the IMRT treatment plans based on square root dose mapping were more repeatable than the plans based on linear dose mapping. This might be explained by the steeper linear dose mapping functions, particularly at dose levels above 85 Gy, in combination with the limited capability of the treatment planning optimization algorithm to deliver high dose to small, isolated areas. Isolated high dose areas are assumed to give rise to uncertainty in the final treatment plan, which ultimately leads to lower repeatability when compared in a test-retest setting. In these areas also less agreement with the prescribed dose was obtained resulting in lower treatment plan quality. This was confirmed with higher $QF_{95\%}$ and $xQF_{95\%}$ values that were found for linear dose mapping plans compared to square root dose mapping plans. Nevertheless, QF values showed that treatment plans were on average in good agreement with the prescribed dose distributions, and xQF values that were all below 5% indicated that the differences in mp-MRI did not compromise play quality.

Feature repeatability analysis revealed a top five stable features all being Gaussian derivatives in the AP direction of the prostate at different smoothing levels. Feature ranking based on the TP model performance revealed a different top five: ADC intensity, tumor prevalence, K^{trans} normalized intensity, and two second order Gaussian derivatives of the T2w image in the LR direction of the prostate with different smoothing kernel (Dinh *et al.* 2017). The higher ranking of the Gaussian derivative features in the repeatability analysis is explained by smoothening of the stochastic noise which is the main cause of within-patient image variability. ICC values of ADC and K^{trans} intensity maps were 0.84 and 0.85 respectively, indicating high repeatability of these features. Comparable findings on repeatability of ADC and K^{trans} intensity features in MRI prostate imaging are reported in literature, as well as on CT and withß other tissue types. Toivonen *et al.* (2015) reported an ICC of 0.89 for ADC intensity in prostate cancer using MRI, although performed on a ROI basis. Alonzi *et al.* (2010) found an ICC between 0.81 – 0.84 for K^{trans}, and reported a within-patient coefficient of variation (wCV) of 13.9% – 15.8%. Koh *et al.* (2009) reported a high repeatability for ADC measurements in a two-center phase I clinical trial. Padhani *et al.* (2002) used the within-patient standard deviation (wSD) to quantify K^{trans} repeatability in muscle tissue in the pelvic region, and found values between 0.32 – 0.33 min^{-1}. K^{trans} was also shown to be reproducible between CT and MRI with an wSD of 0.03 min^{-1} for median K^{trans} values of 0.10 and 0.08 min^{-1} in MRI and CT respectively (Korporaal *et al.* 2011).

Evaluation of differences between planned and prescribed dose distributions during treatment planning occurred via visual comparison of DVH curves and isodose lines. Unacceptable under- and overdosage was controlled by tuning the weight factors of the dose painting objective functions. Remaining underdosing in treatment plans consistently occurred in high dose prescription voxels that are associated with near-certainty about tumor presence. The underdosing is a limitation of the dose painting objective functions in the sense that they penalized under- and overdosing for the whole prostate volume. The relative small fraction of underdosed high prescription voxels has a relatively minor contribution to the total cost function. As a result, the optimizer is not steered strongly enough to increase the dose to these regions. Figure 2 shows an example of this issue where a small lesion in the peripheral zone with a TP of 0.6 has a planned dose lower than the prescribed dose.

Improvement of the optimization algorithm for IMRT planning may be possible, for example by adding new cost functions. This would result in dose distributions that are more similar to the prescribed dose and would be reflected in smaller values of QF and particularly $QF_{95\%}$. This would also imply that the ICC of the IMRT dose distributions would approach the slightly lower ICC of the DPbN prescriptions. However, taking into account that dose distributions in treatment plans are more blurred compared to prescription dose distributions, higher ICC values for IMRT plans can be expected since dose blurring reduces both the inter-voxel heterogeneity and the part of intra-voxel differences caused by noise, thus resulting in higher repeatability in a test-retest situation.

Although the tumor probability of the last three patients was calculated on a coarser T2w-grid, no apparent differences in the ICC values were observed. We assume that the difference in voxel size at this resolution has no discernable influence on the final repeatability. Instead, resampling artifacts within the planning pipeline and deformable registrations between the planning sessions have a higher impact on the observed ICC values. For a test-retest study resampling and registrations steps are inevitable. Introduction of artifacts in the data could however be minimized with reduction of resampling steps within the planning pipeline.

Throughout the planning pipeline repeatability was above 0.75 and median values indicated that stable treatment plans can be realized. Although median values already show high repeatability, the stability of the planning pipeline on the individual level may be further improved. For example, with selection of image features based on not only performance but also their repeatability, which was not investigated in the current study. Furthermore, only simple polynomial dose mapping functions were implemented to prove the principle of DPbN. Instead of dose mapping functions, tumor control probability models could serve to relate tumor probability to prescribed dose based on radiobiological assumptions.

The pipeline holds the prospect of automation. In this work the prostate and OARs were delineated manually, a process that could be replaced with automatic segmentation software. Also, the treatment planning process used for this study involved manual interventions to ensure that the dose-volume constraints from Table 1 were met. Automatic plan generation software is already available and will be valuable in the development of a fully automatic dose painting by numbers MR-only treatment planning workflow.

Conclusions

Using test-retest mp-MRI, we have shown that DPbN treatment plans for prostate cancer can be realized with excellent repeatability. From tumor probability modelling based on mp-MRI towards treatment plan realization based on voxel-wise dose prescription a stable treatment planning pipeline was demonstrated. Using validated tumor probability modelling and simple dose mapping it was shown that despite day-to-day variations in imaging data consistent treatment plans were still obtained.

5

CHAPTER 5

Quantitative MRI changes during weekly ultra-hypofractionated prostate cancer radiotherapy with integrated boost

Marcel A. van Schie
Petra J. van Houdt
Ghazaleh Ghobadi
Floris J. Pos
Iris Walraven
Hans C.J. de Boer
Cornelis A.T. van den Berg
Robert Jan Smeenk
Linda G.W. Kerkmeijer
Uulke A. van der Heide

Frontiers in Oncology, Volume 9, p1264 (2019)

Abstract

Purpose

Quantitative MRI reflects tissue characteristics. As possible changes during radiotherapy may lead to treatment adaptation based on response, we here assessed if such changes during treatment can be detected.

Materials and methods

In the hypoFLAME trial patients received ultra-hypofractionated prostate radiotherapy with an integrated boost to the tumor in five weekly fractions. We analyzed T2 and ADC maps of 47 patients that were acquired in MRI exams prior to and during radiotherapy, and performed rigid registrations based on the prostate contour on anatomical T2-weighted images. We analyzed median T2 and ADC values in three regions of interest (ROIs): the central gland (CG), peripheral zone (PZ) and tumor. We analyzed T2 and ADC changes during treatment and compared patients with and without hormonal therapy. We tested changes during treatment for statistical significance with Wilcoxon signed rank tests. Using confidence intervals as recommended from test-retest measurements, we identified persistent T2 and ADC changes during treatment.

Results

In the CG, median T2 and ADC values significantly decreased 12% and 8% respectively in patients that received hormonal therapy, while in the PZ these values decreased 17% and 18%. In the tumor no statistically significant change was observed. In patients that did not receive hormonal therapy, median ADC values in the tumor increased with 20%, while in the CG and PZ no changes were observed. Persistent T2 changes in the tumor were found in 2 out of 24 patients, while none of the 47 patients had persistent ADC changes.

Conclusions

Weekly quantitative MRI could identify statistically significant ADC changes in the tumor in patients without hormonal therapy. On a patient level few persistent T2 changes in the tumor were observed. Long-term follow-up is required to relate the persistent T2 and ADC changes to outcome and evaluate the applicability of quantitative MRI for response-based treatment adaptation.

Introduction

Whole gland dose escalation for prostate cancer has shown to result in increased biochemical control rates but is associated with increased toxicity (Viani *et al.* 2009). Focal dose escalation may benefit patient outcome without compromising toxicity levels compared to conventional treatment. This hypothesis is currently tested in the FLAME trial (Lips *et al.* 2011) where patients received an integrated boost up to 95 Gy to the visible tumor in addition to a whole gland dose of 77 Gy in 35 treatment fractions. With advancing insight in prostate tumor radiobiology, hypofractionated prostate radiotherapy is increasingly performed (Brenner *et al.* 1999, Miralbell *et al.* 2012). With ultra-hypofractionation, the therapeutic ratio between tumor control and toxicity increases even further due to the low α/β ratio of prostate cancer. Several ultra-hypofractionation trials have demonstrated similar toxicity as compared to standard fractionation, with reduced treatment time (Widmark *et al.* 2016, Morgan *et al.* 2018, Brand *et al.* 2019, Widmark *et al.* 2019). Also, non-inferiority has already been demonstrated (Widmark *et al.* 2016, Widmark *et al.* 2019). For intermediate to high-risk disease, the combination of ultra-hypofractionation with a focal dose escalation to the tumor as conducted in the FLAME trial may even result in better outcomes. Therefore, ultra-hypofractionation was combined with a focal boost to the tumor to treat intermediate to high-risk prostate cancer in the hypoFLAME trial. In prostate cancer long term follow-up of at least five years is required to evaluate treatment outcome. If changes in the prostate occur at an early stage during treatment and are related to outcome, treatment adaptation for prostate cancer could be considered.

Quantitative MRI is known to reflect tissue characteristics. Diffusion weighted imaging (DWI) and T2 mapping are suitable quantitative MRI techniques to investigate tissue properties in the prostate (Bonekamp *et al.* 2018, Mai *et al.* 2019). Through DWI a quantitative apparent diffusion coefficient (ADC) map can be obtained that represents water diffusion between cells and allows to discriminate between malignant and benign prostate tissue. Furthermore, the ADC value of tumor tissue was found to relate to aggressiveness of the disease (Manetta *et al.* 2019). With T2 mapping a spatial distribution of T2 values can be calculated that are unique to biological tissues. T2 was for example found to correlate with hypoxia (Hoskin *et al.* 2007, Luttje *et al.* 2017). Since prostate tumors have different properties from benign prostate tissue, T2 mapping has the potential to discriminate between benign and malignant tissue.

Since quantitative MRI reflects tissue characteristics, tissue changes due to treatment may be visible on quantitative MRI as well. Therefore, quantitative MRI has the potential to generate imaging biomarkers for treatment response assessment. Before investigating this potential

role for quantitative MRI, the first step is to identify if any changes in the tumor during treatment can be detected on quantitative MRI.

To identify changes in the prostate during treatment, in the hypoFLAME trial we acquired quantitative MRI data at each weekly fraction of radiation and tracked quantitative MRI values during the course of treatment. Since concurrent hormonal therapy may affect these MRI values (Hötker *et al.* 2015), we also investigated the influence of hormonal therapy on tissue changes during radiotherapy.

Materials and methods

Patient characteristics

We collected data of 73 patients from two institutions who participated in the hypoFLAME trial (clinicaltrials.gov NCT02853110). All patients had biopsy-proven, clinically localized, intermediate to high-risk prostate cancer (Ash *et al.* 2000). Patients were excluded if they had a contraindication for performing an MRI examination, if no tumor nodule was visible on MRI or if placement of fiducial markers was unsafe. Other exclusion criteria were ≥ 5 mm seminal vesicle invasion, lymph node or distant metastasis, or an iPSA of more than 30 ng/mL. Also, patients that received previous pelvic irradiation or underwent transurethral resection of the prostate (TURP), or patients with an International Prostate Symptom Score (IPSS) > 15 or a World Health Organization (WHO) >2 were not included in the trial. We obtained approval from the institutional review boards and written informed consent from all included patients.

Treatment delivery

Patients were treated in the University Medical Center in Utrecht (UMCU, n = 36) and the Netherlands Cancer Institute in Amsterdam (NKI, n = 37). Dual-arc VMAT treatment was delivered once per week with 35 Gy in five fractions to the prostate, with an integrated focal boost up to 50 Gy to the visible tumor on MRI. Position verification of the prostate was performed prior to each radiation fraction using gold fiducial markers visible on cone-beam CT. In the UMCU 10 out of 36 patients received concurrent hormonal therapy for a period of 6 to 36 months, in the NKI these were 31 out of 37 patients. Hormonal therapy was typically started 2-6 weeks prior to the start of radiotherapy.

Scanning protocol

Prior to treatment patients received a planning CT scan and MRI exam, including a T2-weighted scan and a diffusion weighted imaging (DWI) scan. In the NKI also a T2 mapping sequence was performed. In both institutions patients were scanned on a 3T Philips Ingenia MRI scanner. Specifications of the scanned MRI sequences are listed in Table 1. To track changes in the prostate and tumor during treatment, a weekly repeat MRI exam was scanned at each treatment fraction that included the same image sequences as the pretreatment MRI exam.

Table 1. *Specifications of MRI sequences in the UMCU and NKI. FOV = Field of View, TE = echo time, TR = repetition time. For T2 mapping patients were consistently scanned with one of the reported voxel sizes.*

	UMCU	NKI
T2-weighted (TSE)		
Voxel size (mm^3)		
Acquired	0.6 x 0.7 x 3	0.7 x 0.7 x 3
Reconstructed	0.5 x 0.5 x 3	0.4 x 0.4 x 3
FOV (mm^3)	200 x 200 x 90 / 230 x 230 x 141–150	282 x 282 x 75–90
TE / TR (ms)	90–100 / 3770–8620	120 / 3690–7930
T2 mapping (multi-echo spin-echo)		
Voxel size (mm^3)		
Acquired		0.8 x 0.8 x 3 / 1.0 x 1.0 x 3
Reconstructed		0.4 x 0.4 x 3 / 0.6 x 0.6 x 3
FOV (mm^3)		170 x 170 x 60
TE / TR (ms)		32 / 2470–4150
Echo spacing (ms)		16
Echoes (n)		12
DWI (single-shot EPI)		
Voxel size (mm^3)		
Acquired	3.0 x 3.0 x 4	2.3 x 2.4 x 3
Reconstructed	2.5 x 2.5 x 4	1.1 x 1.1 x 3
FOV (mm^3)	256 x 256 x 66	256 x 256 x 60–66
TE / TR (ms)	62–93 / 3400–4940	62 / 2860–5410
b-values (s/mm^2)	0, 100, 300, 500, 800, 1000	0, 200, 800

Calculation of T2 and ADC maps

The DWI scans were acquired with different protocols as described in Table 1. For consistency between institutions we only considered b-values between 200 and 800 s/mm^2. In the NKI cohort we calculated the ADC maps using b = 200 and 800 s/mm^2, in the UMCU cohort we calculated the ADC maps using b = 300, 500 and 800 s/mm^2. In the NKI cohort we derived quantitative T2 maps from the T2 mapping sequence. For calculation of the T2 map we applied an in-house developed weighted logarithmic fitting algorithm to determine the T2 value per voxel in the image (Dinis Fernandes *et al.* 2019).

Image registration

We registered all images to the pretreatment images to allow for tracking of prostate and tumor changes during treatment. All registrations were performed rigidly with in-house developed software using mutual information as the cost function, and registrations were manually adapted whenever required. Within each MRI exam the b = 0 s/mm^2 image from the DWI was selected, since it contained most anatomical information, and registered to the T2-weighted image. We applied the transformation matrix obtained from registration to the ADC map to register it to the T2-weighted image. From the T2 echo image series the image with echo time closest to the echo time of the T2-weighted image (TE = 120 ms) was selected and registered to the T2-weighted image. We applied the transformation matrix to the T2 map to register it to the T2-weighted image. From each repeat MRI exam we registered the T2-weighted image to the pretreatment T2-weighted image.

Delineations

We delineated the prostate and the peripheral zone on T2-weighted MRI and labeled the remaining part of the prostate as central gland (CG). The delineation of the tumor was based on multi-parametric MRI. CG, PZ and tumor together are referred to as ROIs throughout this study.

Image analysis

We resampled the registered images to 1 mm isotropic voxels. This allowed for exclusion of an isotropic margin of 2 mm around each ROI that was considered to minimize the impact of residual registration errors. We extracted the median value within each ROI on T2 and ADC. We determined the population median value for each time point during treatment. Per patient we normalized the values to the pretreatment value to examine the relative behavior over time. We stratified by patients with and without hormonal therapy to investigate the influence on T2 and ADC changes during hypofractionated radiotherapy.

On a patient level we identified significant trends using confidence intervals for T2 and ADC defined by literature values. These confidence intervals were derived from test-retest measurements. For T2 we used a confidence interval of 11% as found by Van Houdt *et al.* (2018). For ADC we used a value of 47% as recommended by the Quantitative Imaging Biomarkers Alliance (QIBA) (Shukla-Dave *et al.* 2019). These confidence intervals separate real changes in T2 and ADC values from measurement imprecision with 95% confidence. We subsequently determined the number of patients in which T2 and ADC changes were outside the confidence intervals at any time point during treatment and were persistent until week 5.

Statistics

We performed Wilcoxon signed rank tests to identify if changes per ROI were statistically significant during treatment. We applied a Bonferroni correction to account for multiple testing (nine tests), considering $p < 0.0056$ as significance level. All image analysis and statistical tests were performed using MATLAB (MathWorks, Natick, MA, USA).

Results

Table 2 summarizes the number of patients per institution available for analysis. We did not perform analysis on 15 patients for whom less than three out of six MRI exams were scanned. Eleven patients were not analyzed since they were scanned with two different DWI scanning protocols during acquisition of pretreatment and repeat MRI. We could not analyze T2 values of four patients since pretreatment T2 maps were not acquired.

The T2-weighted images, T2 and ADC maps from one patient are shown in Figure 1 for all time points. A decrease in contrast within the prostate can be observed in all three image sequences over the course of treatment, which reduces the conspicuity of the tumor from the surrounding prostate tissue.

Median values of T2 and ADC in the CG, PZ and tumor during pretreatment imaging are shown in Table 3. We observed statistically significant differences in the CG and PZ between the ADC values in the UMCU cohort and the NKI cohort.

Table 2. *Number of patients per institution from which T2 and ADC maps were available for analysis, separated by hormonal therapy (HT or No HT).*

	UMCU	NKI	All
T2			
HT		21	21
No HT		3	3
All		24	24
ADC			
HT	4	24	28
No HT	15	4	19
All	19	28	47

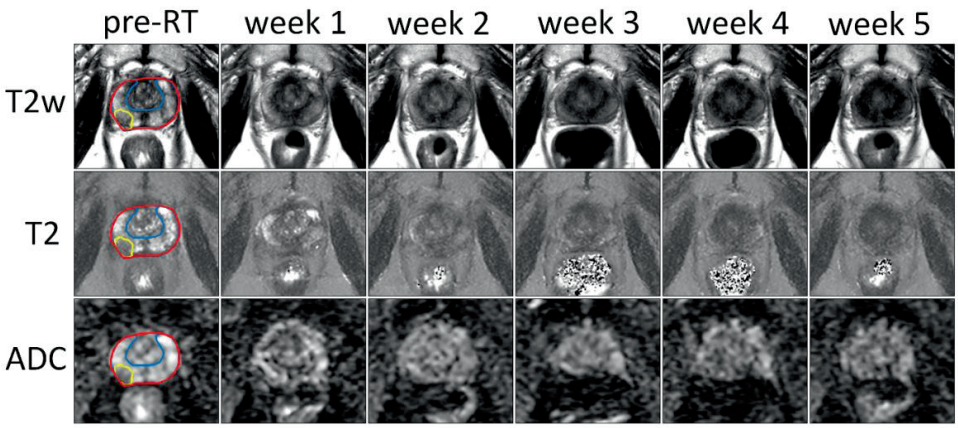


Figure 1. *Example of T2-weighted images, and T2 and ADC maps of the prostate prior to treatment (pre-RT) and at each repeat MRI exam (week 1 to week 5) of a patient treated at the NKI. The entire prostate, the boundary between PZ and CG and the tumor are delineated in red, blue and yellow respectively.*

T2 and ADC values normalized to the pretreatment values are shown in Figure 2. In the CG we observed a median decrease of 12% on T2 and 8% on ADC in patients that received hormonal therapy. T2 and ADC values at week 5 were significantly lower compared to pretreatment values. For patients that received no hormonal therapy, the median ADC value decreased 4% and this was not statistically significant.

Table 3. *Population median and interquartile range (between brackets) of median T2 (in ms) and ADC values (in 10^{-3} mm^2/s) in the CG, PZ and tumor on pretreatment quantitative MRI. Statistically significant differences between institutions are indicated in bold.*

	T2 (ms)	ADC (10^{-3} mm^2/s)	
	NKI	UMCU	NKI
CG	93 (19)	**1.30** (0.13)	**1.09** (0.18)
PZ	110 (26)	**1.37** (0.19)	**1.24** (0.24)
Tumor	80 (9)	1.07 (0.20)	0.90 (0.28)

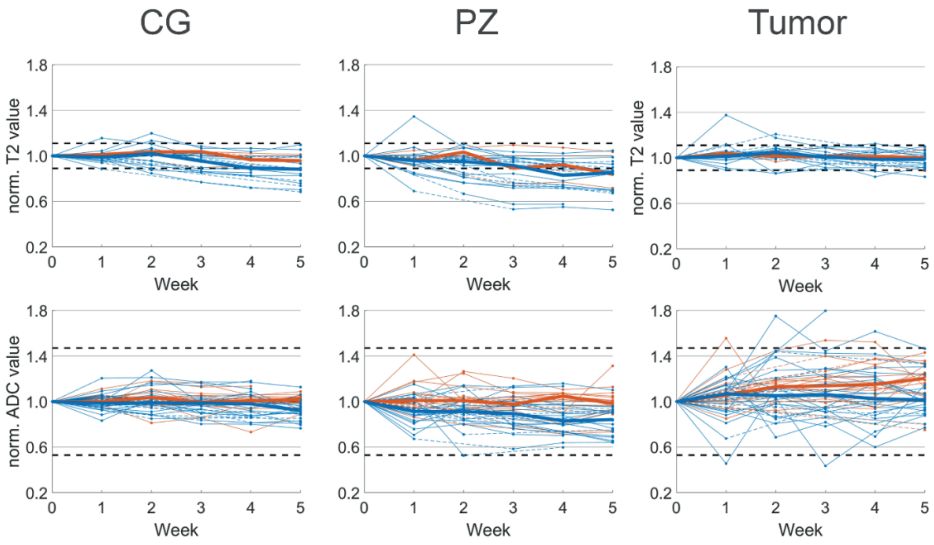


Figure 2. *Median normalized T2 (top) and ADC value (bottom) per patient with respect to pretreatment imaging (week 0). Median values of patients with (blue) and without (orange) hormonal therapy are plotted as solid lines. Interpolated values are displayed with dashed lines. Confidence intervals of 11% and 47% for T2 and ADC respectively are plotted as horizontal dashed lines.*

In the PZ we observed similar behavior. In patients with hormonal therapy the median T2 and ADC value decreased significantly with 17% and 18% respectively, while in patients without hormonal therapy we observed a non-significant decrease in ADC of 5%. In the tumor the behavior was different from CG and PZ. Median increases of 5% and 7% on T2 and ADC maps were found for patients with hormonal therapy, and these were not statistically significant. For patients without hormonal therapy, on ADC we observed a median increase of 20% that was statistically significant.

Due to the low number of patients that were scanned with a T2 mapping sequence and received no hormonal therapy, we did not test statistical significance of T2 changes in these patients. On an individual patient level we found that 14 out of 21 patients who received hormonal therapy, showed persistent T2 changes larger than 11% during treatment. These were 11 patients with persistent changes in the CG, 12 in the PZ and one in the tumor. For the three patients without hormonal therapy, two patients had persistent T2 changes, from which one showed changes in the CG, two in the PZ and one in the tumor. In total 67% of the 23 patients showed persistent T2 changes during treatment. In contrast, on ADC maps for both patients with and without hormonal therapy we observed no changes outside the confidence interval of 47%.

Discussion

In this study we analyzed changes in the prostate as observed on quantitative MRI during hypofractionated radiotherapy with an integrated boost to the tumor. Using repeated imaging we observed changes in median T2 and ADC values that depended on the use of hormonal therapy. The changes we observed can explain the reduced tumor conspicuity that is observed after primary radiotherapy. However, depending on hormonal therapy this can be explained by either normalization of tumor characteristics or by a decrease of normal prostate tissue values. For patients who received hormonal therapy, we observed a reduction of T2 and ADC values in the PZ, while values in the tumor did not change significantly. However, for patients who did not receive hormonal therapy, we found that ADC values increased significantly in the tumor, but not in the PZ.

The pretreatment ADC values were significantly different between the two institutions. This may be a consequence of the DWI scanning protocols. The b-values in both protocols were similar with b = 200 and 800 s/mm^2 in the NKI and b = 300, 500 and 800 s/mm^2 in the UMCU.

However, the acquisition voxel size in the UMCU protocol was 2.2 times larger than in the NKI protocol. This resulted in a different signal to noise ratio and could contribute to differences in ADC values (Scherrer *et al.* 2011).

In the literature a similar variation between ADC values was found. In one study median ADC values in the tumor of $1.08 \pm 0.39 \cdot 10^{-3}$ mm^2/s (mean ± SD) prior to treatment are reported (Sato *et al.* 2005). ADC values in the untreated healthy PZ were $1.8 \pm 0.4 \cdot 10^{-3}$ mm^2/s. Other studies found values of $1.6 \pm 0.2 \cdot 10^{-3}$ mm^2/s in the healthy prostate of untreated patients (Kim *et al.* 2009, Westphalen *et al.* 2012). Again, differences in DWI protocol as well as image reconstruction methods may have contributed to the existing variation.

We observed different trends in patients that did and did not receive hormonal therapy. Hormonal therapy however correlated with the institution where patients were treated. In the UMCU 4 out of the 19 patients received hormonal therapy, while in the NKI this was 24 out of the 28 patients. Because of this unbalanced distribution we could not separate hormonal therapy from institution to explain the differences in normalized ADC value behavior during treatment. This was also the reason we did not compare the T2 values for patients with and without hormonal therapy in the NKI cohort.

One study describes prostate and tumor changes on MRI during treatment. Foltz *et al*. (2013) reported an early treatment response in the entire prostate and CG, plus a progressive response in the PZ and tumor towards the end of treatment (Foltz *et al.* 2013). A statistically significant change in the tumor was found after six weeks on ADC. Early treatment response in the tumor was not observed on either T2 or ADC. While there were differences in the overall treatment duration, the frequency of imaging and the time between radiotherapy fractions compared to our study. our quantitative MRI results indicate similar behavior. We found progressive T2 changes in the PZ and late ADC changes in the tumor. This qualitative comparison is only indicative though, since the use of hormonal therapy was not reported in Foltz *et al*. (2013).

Here we analyzed the T2 and ADC changes in prostate and tumor only during treatment. Dinis Fernandes *et al.* reported late changes on quantitative MRI in recurrent prostate cancer patients that were scanned at least two years after primary treatment (Dinis Fernandes *et al.* 2019). Adjuvant hormonal therapy was given in 82% of the patients but ended at least one year before the MRI examination. Changes in CG and PZ regions on both T2 and ADC maps were found and reduced contrast between PZ and tumor on T2 maps was observed. Median T2 values in the CG, PZ and tumor decreased by 29%, 19% and 5%, while we observed statistically significant decreased values of 12% and 17% in the CG and PZ and no statistically significant change in the tumor. For ADC values a reduction of 5 – 9% in CG, PZ and tumor was observed two year after treatment. In our study we observed a decrease of 8% and 18% in the

CG and PZ in case of hormonal therapy, while an increase of 20% was found in the tumor in absence of hormonal therapy. Based on these findings we expect further reduction of T2 values in the CG and PZ after treatment, as well as post-treatment changes in ADC. Also, the treatment fractionation and both timing and duration of hormonal therapy may contribute to the discrepancies between both studies. Follow-up of patients in our study will be required to confirm if changes on T2 and ADC correlate with long-term biochemical recurrence free survival.

We implemented a rigid registration method to align all images to the pretreatment T2-weighted image. More accurate registration methods like deformable registration could be more appropriate when registering between MRI exams. Deformable registration would account for possible deformations of the prostate between MRI exams and allow for voxel-level analysis. However, as a result of treatment we experienced intensity changes on T2-weighted images that lead to incorrect deformations we were unable to manually adapt. Therefore, we applied rigid registrations instead and minimized registration inaccuracy via removal of an isotropic margin around each ROI, which required resampling of all images. Since we performed our analysis on ROI level, we expected limited impact of both the registration method and the image resampling on our results.

Conclusions

Using quantitative MRI, on a population level we were able to find significant ADC changes in the intraprostatic tumors of patients that did not receive hormonal therapy during hypofractionated radiotherapy. However, early during treatment, when treatment adaptation could be considered, no significant change was identified in the tumor. We did observe only two individual patients that showed persistent T2 changes in the tumor, while no individual patients showed persistent ADC changes in the tumor. On ADC we did observe several patients with early and progressive trends in the tumor although these trends were within the confidence intervals. If these trends are continued after treatment and exceed the confidence intervals, a possible relation between early treatment response and clinical outcome could be established. Follow-up is therefore desired for assessing the potential role of quantitative MRI for adaptation of hypofractionated radiotherapy based on early treatment response.

6

CHAPTER 6

General Discussion

In this thesis, the performance of mp-MRI based focal dose escalation to the visible tumor in the prostate was investigated, and potential applications of mp-MRI to further development of dose-escalated treatment were explored. Although intraprostatic lesions may present locally or heterogeneously distributed throughout the prostate, present-day radiotherapy treatment involves irradiation of the entire prostate gland with a uniform dose. Over the past decades, multiple planning studies have demonstrated the feasibility of a focal boost to the intraprostatic tumor (Singh *et al.* 2007, Bossart *et al.* 2016, Garibaldi *et al.* 2016). Recently, the multi-center phase III FLAME trial (NCT01168479) has shown a statistically significant improvement of biochemical disease-free survival in favor of the dose-escalated arm without increased toxicity (Kerkmeijer *et al.* 2021). Focal dose escalation treatment has become a likely candidate for clinical introduction to improve recurrence free survival rates among intermediate to high-risk prostate cancer patients.

The quality of the delivered focal treatment requires accurate contouring of the intraprostatic lesion, and an optimal planned focal dose escalation to achieve the highest local tumor control. By retrospective analysis of the clinically used delineations and focal boost plans in the FLAME study, it was demonstrated that both delineation practice and escalated dose planning could be further improved, and that automated prediction models are valuable tools to realize those improvements. Furthermore, the role for mp-MRI as direct input for dose painting by numbers (DPbN) treatment plans is promising, since treatment plan quality was found robust to noise and day-to-day variation during image acquisition. During the course of treatment on the contrary, quantitative MRI could not detect early responding tissue. Tumor-response driven adaptive strategies for prostate cancer based on interfraction imaging are therefore considered unlikely to be established in the near future.

Delineating on mp-MRI

Detection of the intraprostatic tumor requires both anatomical and functional information. Mp-MRI combines anatomical and functional imaging in a single examination and is therefore an excellent candidate for focal dose escalation strategies. The generally accepted PI-RADS V2 guidelines recommend acquisition of anatomical T2-weighted and functional DWI for the detection and staging of prostate cancer, with a secondary role for DCE MRI (Weinreb *et al.* 2016). These guidelines lead to improved detection rates and more consistent staging (Kasel-Seibert *et al.* 2016, Padhani *et al.* 2019). However, to date no consensus on delineation practice has been established. It has been reported in various studies that in the absence of delineation guidelines significant interobserver variability is observed (Bratan *et al.* 2013, Rischke *et al.* 2013, Anwar *et al.* 2014, Steenbergen *et al.* 2015). Rischke *et al.* observed

substantial agreement between observers on T2-weighted and DCE-MRI and moderate agreement on DWI, which were considered insufficient for radiotherapy purpose. Steenbergen *et al.* found a median interobserver delineation standard deviation of 2.3 mm for intraprostatic lesions delineated on mp-MRI. In this thesis it was demonstrated that in a large multi-institutional patient cohort apart from the observer variability also institutional bias exists regarding relative use of the different MR sequences to delineate the intraprostatic lesions.

Interobserver variability as well as different institutional weighting of mp-MRI suggest that depending on the observer and the institution, the agreement between delineation and actual tumor boundary may vary. Relative large disagreements could result in a suboptimal targeted focal treatment and may affect the expected improvement in biochemical disease free survival. Linking institutional delineation preferences regarding MRI sequence weighting with disease free survival may give evidence to base delineation guidelines on. Unfortunately, for such evaluation institutional confounders as patient cohort, T-stage and treatment plan generation may require a statistical power beyond that of the FLAME dataset.

Consensus guidelines on the use of mp-MRI sequences for delineation of the intraprostatic lesion are hypothesized to minimize interobserver variability and institutional bias. Another approach to improve delineation practice is the introduction of probabilistic models that predict tumor presence. Histopathology based prediction models may reveal hidden patterns and assist to improve contouring consistency. Groenendaal *et al.* observed an area under the receiver operator curve (AUC) value of 0.70 for predicting tumor presence based on ADC and K^{trans} image parameters, using a logistic regression model that was developed based on ground-truth histopathology from prostatectomy patients (Groenendaal *et al.* 2012). Dinh *et al.* extracted higher order image features and cross-validated a logistic regression model on histopathology data from two institutions yielding an AUC value of 0.78 (Dinh *et al.* 2017). They also related per voxel the model predicted tumor probability with the number of observers that included the voxel in the tumor delineation, and found a significant correlation between tumor probability and group consensus (Dinh *et al.* 2016). However due to a large uncertainty range of this correlation no probability threshold could be identified to distinguish tumor from benign tissue and therefore delineations cannot be derived from the probability maps directly.

Over the past years availability of computational power has enabled development of deep learning approaches for the detection and classification of prostate cancer on MRI. Convolutional neural networks (CNNs) are the most commonly investigated deep learning approach and have the potential to outperform classical logistic regression models. To date, according to a recent systematic review on the detection of significant prostate cancer using

computer models, CNNs were found to perform similar to logistic regression approaches (Castillo T. *et al.* 2020).

In this thesis the per-patient tumor probability prediction was determined using an adapted version of the model developed by Dinh *et al.* to check the quality of the manual tumor delineations of the FLAME trial (Dinh *et al*. 2017). Using the AUC score to quantify the agreement with the predicted tumor probability map, clinical delineations with low AUC values were flagged and reviewed retrospectively. It was found that considerable improvements could be realized upon review, suggesting that applying this methodology in clinical routine could be beneficial for delineation practice.

Although these prediction models improve identification of tumor tissue and contribute to more accurate manual tumor delineations, specific parts of the intraprostatic lesion are systematically missed upon delineation on mp-MRI (Rosenkrantz *et al.* 2013, de Visschere *et al.* 2016, Lewis *et al.* 2017). De Visschere *et al.* concluded that the majority of missed tumors were low grade and organ-confined. Van Houdt *et al.* observed a relation between histopathological features and visibility on mp-MRI. The invisibility of tumors on mp-MRI was associated with heterogeneous morphology and low tumor density (van Houdt *et al.* 2020). Since mp-MRI fails to identify certain tumor characteristics, recent studies also investigated the benefit of an additional PSMA-PET scan. Draulans *et al.* proposed optimal tracer-specific window levels for PSMA PET to reduce interobserver delineation variability of the intraprostatic lesion (Draulans *et al.* 2020). In a review study, Zamboglou *et al.* reported that multimodal imaging using PSMA PET and mp-MRI offers complementary information for intraprostatic tumor delineation (Zamboglou *et al.* 2018). The combination of both image modalities results in high sensitivity and specificity. The addition of diagnostic PET to mp-MRI is believed to improve the detection of the intraprostatic tumor, better characterize the tumor biological properties, and ultimately decrease interobserver delineation variability. Although further research is demanded, the combination of PET and mp-MRI better addresses the present-day requirements for identification of intraprostatic lesions and is considered a realistic scenario to improve delineation consistency.

Focal dose escalation

The FLAME trial demonstrated an increase in 5-year biochemical disease-free survival from 85% to 92% with focal boosting of the intraprostatic lesion (Kerkmeijer *et al.* 2021). Post hoc analysis suggests a strong positive correlation between escalated dose level and the 5-year biochemical disease-free survival. The dose response curve suggests there is room for further improvement of focal dose escalation when further increasing the dose to the intraprostatic

lesion. For the remaining 8% of the patients in the dose-escalated arm of the trial, possible explanations of biochemical recurrence could be underdiagnosis of latent metastases, radioresistance, missed tumor tissue in the delineation, or a too low escalated dose to the intraprostatic lesion. For the latter explanation, in this thesis an anatomy-based prediction model was shown to be able to predict an escalated dose in the tumor that could be achieved upon replanning. Since the FLAME trial has demonstrated that focal boosting is relevant to increase biochemical disease-free survival rates, the KBP prediction model that was evaluated in this thesis may be valuable in future clinical setting to optimally target the intraprostatic tumor.

For the treatment plans where unfavorable patient anatomy or tumor location caused an undertreatment of the intraprostatic tumor, solely focal escalated EBRT might not be the most suitable treatment option. Alternative strategies to achieve steeper dose gradients in the unfavorably located tumor may include minimal invasive combined EBRT and brachytherapy boost to the tumor as investigated in the TARGET trial (Sanmamed *et al.* 2020), or MR-based treatment planning and delivery on an MR-Linac, leading to reduced target volume margins (Murray J, *et al.* 2019). Insertion of an endorectal balloon reduces anorectal doses which gives a therapeutic window to increase radiation dose at the dorsal side of the prostate (Smeenk *et al.* 2011). As an alternative to an endorectal balloon, a hydrogel spacer can be simulated between prostate and rectum to separate prostate and rectal wall, and followed by implantation if shown beneficial for the patient (Vanneste *et al.* 2016, van Wijk *et al.* 2017).

The presented KBP model in this thesis was specifically developed to indicate achievable maximum tumor dose in the prostate and was tested on the treatment plans in the dose-escalated arm of the FLAME study. KBP models published so far were developed to predict DVH parameters of OARs surrounding the target volume (Wu *et al.* 2009, Yuan *et al.* 2012, Good *et al.* 2013, Wall *et al.* 2018). Such predictions could guide the clinical planning optimization process to reduce OAR doses. Incorporation of OAR dose DVH parameter prediction in the KBP model presented in this thesis is therefore considered as the next step to produce a plan QA tool dedicated for clinical focal dose escalation in prostate cancer.

Current state-of-the-art automated plan quality optimization approaches rely on KBP, protocol-based automatic iterative optimization, and multi-criteria optimization (MCO) (Hussein *et al.* 2019, Cozzi *et al.* 2019, Ge *et al.* 2019, Moore 2019). In protocol-based automatic iterative optimization, using a predefined protocol, the objectives and constraints are updated for the next iteration (Gintz *et al.* 2016, Kusters *et al.* 2017) In MCO optimization parameters are tuned to optimally balance the target coverage and spare organs at risk. Erasmus-iCycle is a fully automated MCO algorithm that allows to produce Pareto-optimal treatment plans (Breedveld *et al.* 2012). Presently, automated plan optimization is a standard

module of commercial treatment planning systems. Janssen *et al.* reviewed Pinnacle's AutoPlan module to conclude that even commercial auto-plan solutions should be audited with independent KBP methods (Janssen *et al.* 2019). Future automated planning may involve the combination of KBP and MCO, for which higher plan quality in less time was found as compared to both methods separately (Teichert *et al.* 2019).

Dose painting

Delineation guidelines and incorporation of a prediction model are hypothesized to ease clinical delineation practice with the purpose of escalating the radiation dose to a pre-specified region. The resulting manual delineation of the intraprostatic tumor reflects the binary decision of the observer based on the available clinical and image information. To meet closer with the actual tumor biology characteristics, DPbN omits manual contouring and allows to produce a dose prescription at the resolution of the functional images. The direct conversion from image parameters to dose prescription however, potentially allows image value uncertainties to impact the planned dose distribution. In this thesis it was shown that applying a straightforward first or second order polynomial mapping function to image-derived probability maps of repeated mp-MRI examinations, DPbN planning was robust to image value uncertainties and resulted in equivalent dose distributions. This conclusion is a precondition for the further investigation of DPbN and acceptance as an alternative strategy to contour-based dose escalation.

DPbN requires a mapping function between a (composite) image characteristic and prescription dose. Most studies have adopted a linear relationship, inspired by Vanderstraeten *et al.* (2006). Bowel *et al.* presented polynomial and sigmoidal mapping functions (Bowen *et al.* 2009). Such mathematical mapping functions are straightforward to implement and are applicable to image-derived parameter maps or composite parameter maps. On the other hand, these mapping functions are not validated and oversimplify the true dose-response. In addition, the upper dose limit is usually based on prescribed escalation dose levels from clinical trials, which does not necessarily guarantee sufficient tumor control in all tissue.

As an alternative to DPbN, prescription dose could be related to radiobiological tumor characteristics. In biological optimization generally the tumor control probability (TCP) is chosen as target objective to optimize the dose distribution for (Her *et al.* 2020). In addition, biological optimization may simultaneously minimize normal tissue complication probability (NTCP). Most planning studies incorporate a phenomenological TCP model with radiobiological parameter values derived from clinical data (Levegrün *et al.* 2002, Kim *et al.* 2006). Biofocused radiotherapy (BiRT) incorporates histopathology-validated machine

learning methods to derive radiobiological feature maps from imaging and produce a TCP objective map for biological optimization (Haworth *et al.* 2018).

Several studies have reported on the relation between pretreatment ADC values and pathological Gleason score as a predictor of biochemical recurrence. Ghobadi *et al.* postulated a patient-specific dose response with incorporation of the Gleason score in the linear-quadratic model (Ghobadi *et al.* 2016). Casares-Magaz *et al.* inserted cell densities derived from ADC values into a linear-quadratic model to estimate individual tumor control probability (TCP) levels (Casares-Magaz *et al.* 2016). Grönlund *et al.* combined these findings to demonstrate a formalism in which ADC values were related to Gleason score driven dose-responses and dose painted treatment plans were generated that yielded higher TCP and similar dose to normal tissue as compared to homogenous dose prescription (Grönlund *et al.* 2018, Grönlund *et al.* 2021). Although biological optimization is a promising method to yield dose painted treatment plans, uncertainties of radiobiological parameters and models need to be addressed first before considering patient studies.

Both DPbN and biological optimization strategies prescribe a heterogeneous dose distribution to the prostate gland. During treatment anatomical movements may introduce a relative displacement of the delivered dose, which in case of dose painted plans has consequences for the actual received dose by the prostate tissue. Probabilistic planning methods have been developed and evaluated that incorporate optimization uncertainties and perform a robust optimization (Shusharina *et al.* 2018, Miura *et al*. 2019, Bortfeld *et al.* 2021). Thereby image value uncertainties as well as modelling and geometric uncertainties can be accounted for in the generated treatment plan.

DPbN treatment planning in principle could also de-escalate the radiation dose based on image characteristics representing benign prostate tissue. Ultimately such de-escalation to benign tissue below conventional doses is considered beneficial for the patient, but it is challenging to prove this hypothesis without risking insufficient local tumor control. One randomized study for head and neck cancer treatment avoided this risk by defining a dose escalated region first (Heukelom *et al.* 2013). One treatment arm received focal escalated treatment, in the other arm the escalated dose was redistributed based on image characteristics. Thereby the de-escalated part of the escalated dose region still received a dose level above conventional prescription. Such exploratory studies are essential to investigate safe lower limits of radiation dose to tissues considered benign.

Imaging biomarkers

Quantitative imaging biomarkers in prostate radiotherapy are currently primarily used in research on treatment response monitoring and to a smaller extent on dose painting (Gurney-Champion *et al.* 2020). Regarding treatment response monitoring, biomarkers derived from quantitative MRI could prove valuable for image-based adaptive treatment strategies. In order to deliver adaptive therapy, actionable changes occurring early during the course of treatment are required. In this thesis it was investigated if MRI was capable of recording early treatment changes to the tumor and surrounding tissue. In an extreme-hypofractionated setting with an MRI at each of the five fractions, significant early changes to MRI parameters were not observed. Although this study was not powered to detect small or individual changes, results show that adaptive treatment based on the investigated MRI parameters is not evident.

In this thesis imaging biomarkers to evaluate prostate cancer treatment response were investigated in a hypofractionated radiotherapy schedule. In a conventional fractionation schedule primarily ADC has been subject of response monitoring research. Park *et al.* observed significant increase of tumor ADC value at three time points during and after treatment, with respect to the pretreatment ADC value (Park *et al.* 2012). Also, for other tumor sites ADC values were found to change during conventionally fractionated treatment: cervical cancer (Liu *et al.* 2009), rectal cancer (Lambrecht *et al.* 2012), and brain metastases (Mahmood *et al.* 2017). Both the high fraction dose as well as the relative long interval of one week between radiation fraction and MRI examination could be hypothesized to explain why early ADC changes were not observed in this thesis.

Although quantitative imaging for radiotherapy has been studied for decades, multiple issues need to be addressed before clinical introduction is to be considered. Dose response monitoring, timing and optional early treatment adaptation require imaging biomarkers changes that are predictive of patient outcome for a given radiation dose. For dose painting based on quantitative imaging biomarkers, establishing which voxels of the original tumor are related to poor outcome is challenging. In addition, sophisticated trial design is required to prevent unethical dose de-escalation. As long as the imaging biomarker of interest is not proven to be predictive, sub regions in the prostate may receive an unjustified dose de-escalation based on the imaging biomarker with less tumor control as compared to conventional treatment. In order to identify imaging biomarkers for treatment response monitoring or dose painting purpose, large multicenter studies are required to demonstrate clinical applicability (Gurney-Champion *et al.* 2020).

Hypofractionation

In this thesis an assessment of tumor delineations and focal escalated dose distributions was performed on treatment plans that were delivered in 35 fractions. The presented automatic evaluation tools could lead to more consistent delineation practice and higher escalated doses whenever such treatment is delivered in routine clinical practice. Over the past years moderately hypofractionated treatment has become standard of care and extreme hypofractionation with only five treatment fractions is expected to perform even better in terms of patient outcome and clinical workload. Therefore, for patients with intermediate to high-risk disease, trials have been performed that combined extreme hypofractionation and focal dose escalation, and promising preliminary results have been presented (Draulans *et al.* 2020, Murray *et al.* 2020, Nicholls *et al.* 2020). Supposing positive long-term results of these trials, the applicability of the presented work in this thesis to extremely hypofractionated focal dose escalation can already be considered.

The prescribed focal escalated dose distribution is a direct result of the delineation of the intraprostatic lesion. The accuracy of the dose that is actually delivered to the intraprostatic lesion is dependent on the ability of the treatment planning system to realize steep dose gradients and on anatomical changes during treatment. For the latter, the relative high number of treatment fractions with conventional fractionation causes blurring of the planned dose distribution upon treatment delivery and is therefore rather forgiving to small delineation inaccuracies. With fewer treatment fractions, the blurring effect will decrease and delineation inaccuracies will have more impact on the delivered dose. Therefore, precise definition of the tumor boundary is more relevant in an extreme-hypofractionated setting.

The KBP model presented in this thesis was developed using features from planned dose distributions for delivery with conventional fractionation. For application to hypofractionated dose escalation plans, the methodology would remain the same. Anatomical features need to be derived from a new dataset consisting of patients treated with hypofractionated focal dose escalation, and be related to the achieved dose to the tumor. In the FLAME consortium both the Hypo-FLAME (NCT02853110) and the Hypo-FLAME 2.0 (NCT04045717) study would be suitable datasets to build a tumor dose prediction model on. Such model could for example be a valuable tool to explain the discrepancy between the intended dose escalation of 50 Gy to the visible tumor and the realized median dose escalation of 44.7 Gy in the HypoFLAME study (Draulans *et al.* 2020).

The robustness of DPbN treatment plans to image uncertainties was evaluated up to the treatment planning stage and independent from the fractionation scheme. DPbN should therefore be equally applicable to extreme hypofractionated as conventional

hypofractionated focal dose escalation treatments. Since hypofractionated treatment is always preceded by online position verification, setup errors are avoided. Therefore, mainly blurring of the planned dose due to interfraction motion impacts the delivery of the DPbN plan.

The applicability of quantitative MRI for dose response monitoring and potential adaptive strategies in an extreme hypofractionated treatment setting was presented in this thesis. There were no strong indications for a dose response early during treatment on a population basis. Individual early changes were detected but within the uncertainty ranges of the image parameters. Additional research to the outcome of the investigated patients in this study may raise hypotheses on tissue response on an individual level. Further optimization of image protocols is also desired to reduce the uncertainty bandwidth and ease detection of significant early treatment changes.

Conclusions

The FLAME trial has demonstated that mp-MRI can be utilized to define the intraprostatic tumor and deliver a focal dose escalated treatment with improved biochemical disease-free survival. In this thesis it was concluded that delineation guidelines are desired to minimize institutional bias towards interpretation of mp-MRI and ultimately improve delineation consistency and accuracy. Prediction tools for tumor localization and dose escalation were shown to yield potential improvements to clinical cases and are therefore considered valuable tool for future clinical practice of focal dose escalation treatment.

Mp-MRI can furthermore find application in future dose escalation strategies. For DPbN treatment planning it was found that mp-MRI value uncertainties would not significantly impact planned dose distributions and thereby enable further investigation of DPbN as a more sophisticated dose escalation treatment strategy. During the course of extreme-hypofractionated treatment however, repeated MRI did not reveal early responding tumor tissue to potentially adapt treatment to. It is therefore not likely that these will result in actionable imaging biomarkers in the near future to allow for response-based treatment adaptations in hypofractionated radiotherapy of prostate cancer.

A

APPENDICES

References

Alonzi R, Taylor NJ, Stirling JJ, *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* **32**, 155–164 (2010).

Anwar M, Westphalen AC, Jung AJ, *et al.* Role of endorectal MR imaging and MR spectroscopic imaging in defining treatable intraprostatic tumor foci in prostate cancer: quantitative analysis of imaging contour compared to whole-mount histopathology. *Radiother Oncol*. **110**, 303–308 (2014).

Appenzoller LM, Michalski JM, Thorstad WL, Mutic S, Moore KL. Predicting dose-volume histograms for organs-at-risk in IMRT planning. *Med Phys*. **39**, 7446–7461 (2012).

Ash D, Flynn A, Battermann J, *et al.* ESTRO/EAU/EORTC recommendations on permanent seed implantation for localized prostate cancer. *Radiother Oncol*. **57**, 315–321 (2000).

Bammer R. Basic principles of diffusion-weighted imaging. *Eur J Radiol*. **45**, 169–184 (2002).

Barentsz JO, Richenberg J, Clements R, *et al.* ESUR prostate MR guidelines 2012. *Eur Radiol*. **22**, 746–757 (2012).

Bauman G, Haider M, van der Heide UA, Ménard C. Boosting imaging defined dominant prostatic tumors: a systematic review. *Radiother Oncol*. **107**, 274–281 (2013).

Benjamin LC, Tree AC, Dearnaley DP. The role of hypofractionated radiotherapy in prostate cancer. *Curr Oncol Rep*. **19**, 30 (2017).

Bentzen SM. Theragnostic imaging for radiation oncology: dose-painting by numbers. *Lancet Oncol*. **6**, 112–117 (2005).

Böhmer D, Wirth M, Miller K, Budach V, Heidenreich A, Wiegel T. Radiotherapy and hormone treatment in prostate cancer. *Dtsch Arztebl Int*. **113**, 235–241 (2016).

Bonekamp D, Kohl S, Wiesenfarth M, *et al.* Radiomic machine learning for characterization of prostate lesions with MRI: comparison to ADC values. *Radiology*. **289**, 128–137 (2018).

Bortfeld T, Shusharina N, Craft DL. Probabilistic definition of the clinical target volume - implications for tumor control probability modeling and optimization. *Phys Med Biol*. **66**, 01NT01 (2021).

Bossart EL, Stoyanova R, Sandler K, *et al.* Feasibility and initial dosimetric findings for a randomized trial using dose-painted multiparametric magnetic resonance imaging-defined targets in prostate cancer. *Int J Radiat Oncol Biol Phys*. **95**, 827–834 (2016).

Bowen SR, Flynn RT, Bentzen SM, Jeraj R. On the sensitivity of IMRT dose optimization to the mathematical form of a biological imaging-based prescription function. *Phys Med Biol*. **54**, 1483–1501 (2009).

Brand DH, Tree AC, Ostler P, *et al.* Intensity-modulated fractionated radiotherapy versus stereotactic body radiotherapy for prostate cancer (PACE-B): acute toxicity findings from an international, randomised, open-label, phase 3, non-inferiority trial. *Lancet Oncol*. **20**, 1531–1543 (2019).

Bratan F, Niaf E, Melodelima C, *et al.* Influence of imaging and histological factors on prostate cancer detection and localisation on multiparametric MRI: a prospective study. *Eur Radiol*. **23**, 2019–2029 (2013).

Breedveld S, Stochi PRM, Voet PWJ, Heijmen BJM. iCycle: integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. *Med Phys*. **39**, 951–963 (2012).

Brenner DJ, Hall EJ. Fractionation and protraction for radiotherapy of prostate carcinoma. *Int J Radiat Oncol Biol Phys*. **43**, 1095–1101 (1999).

Brix G, Semmler W, Port R, Schad LR, Layer G, Lorenz WJ. Pharmacokinetic parameters in CNS Gd-DTPA enhanced MR imaging. *J Comput Assist Tomogr*. **15**, 621–628 (1991).

Buckley DL, Kerslake RW, Blackband SJ, Horsman A. Quantitative analysis of multi-slice Gd-DTPA enhanced dynamic MR images using an automated simplex minimization procedure. *Magn Reson Med*. **32**, 646–651 (1994).

Cabarrus MC, Westphalen AC. Multiparametric magnetic resonance imaging of the prostate-a basic tutorial. *Transl Androl Urol*. **6**, 376–386 (2017).

Casares-Magaz O, van der Heide UA, Rørvik J, Steenbergen S, Muren LP. A tumour control probability model for radiotherapy of prostate cancer using magnetic resonance imaging-based apparent diffusion coefficient maps. *Radiother Oncol*. **119**, 111–116 (2016).

Castillo T. JM, Arif M, Niessen WJ, Schoots IG, Veenland JF. Automated classification of significant prostate cancer on MRI: a systematic review on the performance of machine learning applications. *Cancers (Basel)*. **12**, 1606 (2020).

Cellini N, Morganti AG, Mattiucci GC, *et al.* Analysis of intraprostatic failures in patients treated with hormonal therapy and radiotherapy: implications for conformal therapy planning. *Int J Radiat Oncol Biol Phys*. **53**, 595 – 599 (2002).

Chang AJ, Autio KA, Ill MR, Scher HI. "High-risk" prostate cancer: classification and therapy. *Nat Rev Clin Oncol.* **11**, 308–323 (2014).

Chen ME, Johnston DA, Tang K, Babaian RJ, Troncoso P. Detailed mapping of prostate carcinoma foci: biopsy strategy implications. *Cancer*. **89**, 1800–1809 (2000).

Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *Am J Ment Defic*. **86**, 127–137 (1981).

Cozzi L, Heijmen BJM, Muren LP. Advanced treatment planning strategies to enhance quality and efficiency of radiotherapy. *Phys Imaging Radiat Oncol*. **11**, 69–70 (2019).

Dearnaley DP, Sydes MR, Graham JD, *et al.* Escalated-dose versus standard-dose conformal radiotherapy in prostate cancer: first results from the MRC RT01 randomised controlled trial. *Lancet Oncol*. **8**, 475–487 (2007).

Dearnaley D, Syndikus I, Mossop H, *et al.* Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHiP trial. *Lancet Oncol*. **17**, 1047–1060 (2016).

Dickinson L, Ahmed HU, Allen C, *et al.* Scoring systems used for the interpretation and reporting of multiparametric MRI for prostate cancer detection, localization, and characterization: could standardization lead to improved utilization of imaging within the diagnostic pathway? *J Magn Reson Imaging*. **37**, 48–58 (2013).

Differding S, Sterpin E, Hermand N, *et al.* Radiation dose escalation based on FDG-PET driven dose painting by numbers in oropharyngeal squamous cell carcinoma: a dosimetric comparison between TomoTherapy-HA and RapidArc. *Radiat Oncol*. **12**, 59 (2017).

Dinh CV, Steenbergen P, Ghobadi G, *et al.* Magnetic resonance imaging for prostate cancer radiotherapy. *Phys Med*. **32**, 446–451 (2016).

Dinh CV, Steenbergen P, Ghobadi G, *et al.* Multicenter validation of prostate tumor localization using multiparametric MRI and prior knowledge. *Med Phys*. **44**, 949–961 (2017).

Dinis Fernandes C, van Houdt PJ, Heijmink SWTPJ, *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*. **50**, 269 –278 (2019).

Donnelly DW, Vis LC, Kearney T, *et al.* Quality of life among symptomatic compared to PSA-detected prostate cancer survivors - results from a UK wide patient-reported outcomes study. *BMC Cancer*. **19**, 947 (2019).

Draulans C, van der Heide UA, Haustermans K, *et al.* Primary endpoint analysis of the multicentre phase II hypo-FLAME trial for intermediate and high risk prostate cancer. *Radiother Oncol*. **147**, 92–98 (2020).

Dregely I, Prezzi D, Kelly-Morland C, Roccia E, Neji R, Goh V. Imaging biomarkers in oncology: Basics and application to MRI. *J Magn Reson Imaging*. **48**, 13–26 (2018).

Duprez F, de Neve W, de Gersem W, Coghe M, Madani I. Adaptive dose painting by numbers for head-and-neck cancer. *Int J Radiat Oncol Biol Phys*. **80**, 1045–1055 (2011).

European Society of Radiology (ESR). ESR position paper on imaging biobanks. *Insights Imaging*. **6**, 403–410 (2015).

Ferlay J, Colombet M, Soerjomataram I, *et al.* Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer*. **144**, 1941–1953 (2019).

Foltz WD, Wu A, Chung P, *et al.* Changes in apparent diffusion coefficient and T2 relaxation during radiotherapy for prostate cancer. *J Magn Reson Imaging*. **37**, 909–916 (2013).

Fonteyne V, Villeirs G, Speleers B, *et al.* Intensity-modulated radiotherapy as primary therapy for prostate cancer: report on acute toxicity after dose escalation with simultaneous integrated boost to intraprostatic lesion. *Int J Radiat Oncol Biol Phys*. **72**, 799–807 (2008).

Garibaldi E, Delmastro E, Gabriele D, *et al.* Clinical and technical feasibility of ultra-boost irradiation in dominant intraprostatic lesion by tomotherapy: preliminary experience and revision of literature. *Panminerva Med*. **58**, 16–22 (2016).

Ge Y, Wu QJ. Knowledge based planning for intensity modulated radiation therapy: a review of data driven approaches. *Med Phys*. **46**, 2760–2775 (2019).

Ghobadi G, De Jong J, Hollmann BG, *et al.* Histopathology-derived modeling of prostate cancer tumor control probability: implications for the dose to the tumor and the gland. *Radiother Oncol*. **119**, 97–103 (2016).

Giganti F, Moore CM. A critical comparison of techniques for MRI-targeted biopsy of the prostate. *Transl Androl Urol*. **6**, 432–443 (2017).

Gintz D, Latifi K, Cuadell J, *et al.* Initial evaluation of automated treatment planning software. *J Appl Clin Med Phys*. **17**, 331–346 (2016).

Gomez-Iturriaga A, Casquero F, Urresola A, *et al.* Dose escalation to dominant intraprostatic lesions with MRI-transrectal ultrasound fusion high-dose-rate prostate brachytherapy. Prospective phase II trial. *Radiother Oncol*. **119**, 91–96 (2016).

Good D, Lo J, Lee WR, Wu QJ, Yin FF, Das SK. A knowledge-based approach to improving and homogenizing intensity modulated radiation therapy planning quality among treatment centers: an example application to prostate cancer planning. *Int J Radiat Oncol Biol Phys*. **87**, 176–181 (2013).

Groenendaal G, Borren A, Moman MR, *et al.* Pathological validation of a model based on diffusion-weighted imaging and dynamic contrast-enhanced MRI for tumor delineation in the prostate peripheral zone. *Int J Radiat Oncol Biol Phys*. **82**, e537–e544 (2012).

Grönlund E, Johansson S, Nyholm T, Thellenberg C, Ahnesjö A. Dose painting of prostate cancer based on Gleason score correlations with apparent diffusion coefficients. *Acta Oncol*. **57**, 574–581 (2018).

Grönlund E, Almhagen E, Johansson S, Traneus E, Ahnesjö A. Robust maximization of tumor control probability for radicality constrained radiotherapy dose painting by numbers of head and neck cancer. *Phys Imaging Radiat Oncol*. **12**, 56–62 (2019).

Grönlund E, Almhagen E, Johansson S, *et al.* Robust treatment planning of dose painting for prostate cancer based on ADC-to-Gleason score mappings – what is the potential to increase the tumor control probability? *Acta Oncol*. **60**, 199–206 (2021).

Gurney-Champion OJ, Mahmood F, van Schie M, *et al.* Quantitative imaging for radiotherapy purposes. *Radiother Oncol*. **146**, 66–75 (2020).

van Haaren PMA, Bel A, Hofman P, van Vulpen M, Kotte ANTJ, van der Heide UA. Influence of daily setup measurements and corrections on the estimated delivered dose during IMRT treatment of prostate cancer patients. *Radiother Oncol*. **90**, 291–298 (2009).

Haworth A, Sun Y, Ebert M, *et al.* Incorporating a novel radiomics framework for Biologically optimised prostate RadioTherapy (BiRT). *Med Phys*. **45**, e446 (2018).

Heijmink SWTPJ, Fütterer JJ, Hambrock T, *et al.* Prostate cancer: body-array versus endorectal coil MR imaging at 3 T--comparison of image quality, localization, and staging performance. *Radiology*. **244**, 184–195 (2007).

Her EJ, Haworth A, Rowshanfarzad P, Ebert MA. Progress towards patient-specific, spatially-continuous radiobiological dose prescription and planning in prostate cancer IMRT: an overview. *Cancers (Basel)*. **12**, 854 (2020).

Heukelom J, Hamming O, Bartelink H, *et al.* Adaptive and innovative Radiation Treatment FOR improving Cancer treatment outcomE (ARTFORCE); a randomized controlled phase II trial for individualized treatment of head and neck cancer. *BMC Cancer*. **13**, 84 (2013).

Hoskin PJ, Carnell DM, Taylor NJ, *et al.* Hypoxia in prostate cancer: correlation of BOLD-MRI with pimonidazole immunohistochemistry-initial observations. *Int J Radiat Oncol Biol Phys*. **68**, 1065–1071 (2007).

Hötker AM, Mazaheri Y, Zheng J, *et al.* Prostate cancer: assessing the effects of androgen-deprivation therapy using quantitative diffusion-weighted and dynamic contrast-enhanced MRI. *Eur Radiol*. **25**, 2665–2672 (2015).

van Houdt PJ, Agarwal HK, van Buuren LD, *et al.* Performance of a fast and high-resolution multi-echo spin-echo sequence for prostate T2 mapping across multiple systems. *Magn Reson Med*. 79, 1586–1594 (2018).

van Houdt PJ, Ghobadi G, Schoots IG, *et al.* Histopathological features of MRI-invisible regions of prostate cancer lesions. *J Magn Reson Imaging*. **51**, 1235–1246 (2020).

Huang W, Chen Y, Fedorov A, *et al.* The impact of arterial input function determination variations on prostate dynamic contrast-enhanced magnetic resonance imaging pharmacokinetic modeling: a multicenter data analysis challenge. *Tomography*. **2**, 56–66 (2016).

Hussein M, Heijmen BJM, Verellen D, Nisbet A. Automation in intensity-modulated radiotherapy treatment planning – a review of recent innovations. *Br J Radiol*. **91**, 20180270 (2018).

Janssen TM, Kusters M, Wang Y, *et al.* Independent knowledge-based treatment planning QA to audit Pinnacle autoplanning. *Radiother Oncol*. **133**, 198–204 (2019).

Jung SI, Donati OF, Vargas HA, Goldman D, Hricak H, Akin O. Transition zone prostate cancer: incremental value of diffusion-weighted endorectal MR imaging in tumor detection and assessment of aggressiveness. *Radiology*. **269**, 493–503 (2013).

Kasel-Seibert M, Lehmann T, Aschenbach R, *et al.* Assessment of PI-RADS v2 for the detection of prostate cancer. *Eur J Radiol*. **85**, 726–731 (2016).

Kayhan A, Fan X, Oommen J, Oto A. Multi-parametric MR imaging of transition zone prostate cancer: imaging features, detection and staging. *World J Radiol*. **2**, 180–187 (2010).

Kerkmeijer LGW, Maspero M, Meijer GJ, van der Voort van Zyp JRN, de Boer HCJ, van den Berg CAT. Magnetic resonance imaging only workflow for radiotherapy simulation and planning in prostate cancer. *Clin Oncol (R Coll Radiol)*. **30**, 692–701 (2018).

Kerkmeijer LGW, Groen VH, Pos FJ, *et al.* Focal boost to the intraprostatic tumor in external beam radiotherapy for patients with localized prostate cancer: results from the FLAME randomized phase III trial. *J Clin Oncol*. **39**, 787–796 (2021).

Khalifa F, Soliman A, El-Baz A, *et al.* Models and methods for analyzing DCE-MRI: a review. *Med Phys*. **41**, 124301 (2014).

Kim CK, Park BK, Lee HM. Prediction of locally recurrent prostate cancer after radiation therapy: incremental value of 3T diffusion-weighted MRI. *J Magn Reson Imaging*. **29**, 391–397 (2009).

Kim Y, Tomé WA. Risk-adaptive optimization: selective boosting of high-risk tumor subvolumes. *Int J Radiat Oncol Biol Phys*. **66**, 1528–1542 (2006).

Kishan AU, King CR. Stereotactic body radiotherapy for low- and intermediate-risk prostate cancer. *Semin Radiat Oncol*. **27**, 268–278 (2017).

Koh DM, Blackledge M, Collins DJ, *et al.* Reproducibility and changes in the apparent diffusion coefficients of solid tumours treated with combretastatin A4 phosphate and bevacizumab in a two-centre phase I clinical trial. *Eur Radiol*. **19**, 2728–2738 (2009).

Korporaal JG, van den Berg CAT, van Osch MJP, Groenendaal G, van Vulpen M, van der Heide UA. Phase-based arterial input function measurements in the femoral arteries for quantification of dynamic contrast-enhanced (DCE) MRI and comparison with DCE-CT. *Magnet Reson Med*. **66**, 1267–1274 (2011).

Kusters JMAM, Bzdusek K, Kumar P, *et al.* Automated IMRT planning in Pinnacle. *Strahlenther Onkol*. **193**, 1031–1038 (2017).

Lambrecht M, Vandecaveye V, de Keyzer F, *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*. **82**, 863–870 (2012).

Langer DL, van der Kwast TH, Evans AJ, Trachtenberg J, Wilson BC, Haider MA. Prostate cancer detection with multi-parametric MRI: logistic regression analysis of quantitative T2, diffusion-weighted imaging, and dynamic contrast-enhanced MRI. *J Magn Reson Imaging*. **30**, 327–334 (2009).

van Leenders GJLH, van der Kwast TH, Grignon DJ, *et al.* The 2019 International Society of Urological Pathology (ISUP) consensus conference on grading of postatic carcinoma. *Am J Surg Pathol*. **44**, e87–e99 (2020).

Lehrer EJ, Kishan AU, Yu JB, *et al.* Ultrahypofractionated versus hypofractionated and conventionally fractionated radiation therapy for localized prostate cancer: a systematic review and meta-analysis of phase III randomized trials. *Radiother Oncol*. **148**, 235–242 (2020).

Levegrün S, Jackson A, Zelefsky MJ, *et al.* Risk group dependence of dose–response for biopsy outcome after three-dimensional conformal radiation therapy of prostate cancer. *Radiother Oncol*. **63**, 11–26 (2002).

Lewis S, Besa C, Rosen A, *et al.* Multiparametric magnetic resonance imaging for transition zone prostate cancer: essential findings, limitations, and future directions. *Abdom Radiol*. **42**, 2732–2744 (2017).

van Lin ENJT, Fütterer JJ, Heijmink SWTPJ, *et al.* IMRT boost dose planning on dominant intraprostatic lesions: gold marker-based three-dimensional fusion of CT with dynamic contrast-enhanced and 1H-spectroscopic MRI. *Int J Radiat Oncol Biol Phys*. **65**, 291–303 (2006).

Ling CC, Humm J, Larson S, *et al.* Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. *Int J Radiat Oncol Biol Phys*. **47**, 551–560 (2000).

Lips IM, van der Heide UA, Kotte ANTJ, van Vulpen M, Bel A. Effect of translational and rotational errors on complex dose distributions with off-line and on-line position verification. *Int J Rad Oncol Biol Phys*. **74**, 1600–1608 (2009).

Lips IM, van der Heide UA, Haustermans K, *et al.* Single blind randomized phase III trial to investigate the benefit of a focal lesion ablative microboost in prostate cancer (FLAME-trial): study protocol for a randomized controlled trial. *Trials*. **12**, 255 (2011).

Liu Y, Bai R, Sun H, Liu H, Wang D. Diffusion-weighted magnetic resonance imaging of uterine cervical cancer. *J Comput Assist Tomogr*. **33**, 858–862 (2009).

Luttje MP, van Buuren LD, Luijten PR, van Vulpen M, van der Heide UA, Klomp DWJ. Towards intrinsic R2* imaging in the prostate at 3 and 7 tesla. *Magn Reson Imaging*. **42**, 16–21 (2017).

Mahmood F, Johannesen HH, Geertsen P, Hansen RH. Ultra-early apparent diffusion coefficient change indicates irradiation and predicts radiotherapy outcome in brain metastases. *Acta Oncol*. **56**, 1651–1653 (2017).

Mai J, Abubrig M, Lehmann T, *et al.* T2 mapping in prostate cancer. *Invest Radiol*. **54**, 146–152 (2019).

Manetta R, Palumbo P, Gianneramo C, *et al.* Correlation between ADC values and Gleason score in evaluation of prostate cancer: multicentre experience and review of the literature. *Gland Surg*. **8**, S216–S222 (2019).

Maurer MH, Heverhagen JT. Diffusion weighted imaging of the prostate-principles, application, and advances. *Transl Androl Urol*. **6**, 490–498 (2017).

McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. **1**, 30–46 (1996).

McMahon SJ. The linear quadratic model: usage, interpretation and challenges. *Phys Med Biol*. **64**, 01TR01 (2018).

Miralbell R, Roberts SA, Zubizarreta E, Hendry JH. Dose-fractionation sensitivity of prostate cancer deduced from radiotherapy outcomes of 5,969 patients in seven international institutional datasets: α/β = 1.4 (0.9-2.2) Gy. *Int J Radiat Oncol Biol Phys*. **82**, e17–e24 (2012).

Miura H, Doi Y, Ozawa S, *et al.* Volumetric modulated arc therapy with robust optimization for larync cancer. *Phys Med*. **58**, 54–58 (2019).

Monninkhof EM, van Loon JWL, van Vulpen M, *et al.* Standard whole prostate gland radiotherapy with and without lesion boost in prostate cancer: toxicity in the FLAME randomized controlled trial. *Radiother Oncol*. **127**, 74–80 (2018).

Moore KL, Brame RS, Low DA, Mutic S. Experience-based quality control of clinical intensity-modulated radiotherapy planning. *Int J Radiat Oncol Biol Phys*. **81**, 545–551 (2011).

Moore KL. Automated radiotherapy treatment planning. *Semin Radiat Oncol*. **29**, 209–218 (2019).

Morgan PB, Hanlon AL, Horwitz EM, Buyyounouski MK, Uzzo RG, Pollack A. Radiation dose and late failures in prostate cancer. *Int J Radiat Oncol Biol Phys*. **67**, 1074–1081 (2007).

Morgan SC, Hoffman K, Loblaw DA, *et al.* Hypofractionated radiation therapy for localized prostate cancer: an ASTRO, ASCO, and AUA evidence-based guideline. *J Clin Oncol*. **36**, JCO1801097 (2018).

Moris L, Cumberbatch MG, van den Broeck T, *et al.* Benefits and risks of primary treatments for high-risk localized and locally advanced prostate cancer: an international multidisciplinary systematic review. *Eur Urol*. **77**, 614–627 (2020).

Mottet N, Bellmunt J, Bolla M, *et al.* EAU-ESTRO-SIOG guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent. *Eur Urol*. **71**, 618–629 (2017).

Murase K. Efficient method for calculating kinetic parameters using T1-weighted dynamic contrast-enhanced magnetic resonance imaging. *Magn Reson Med*. **51**, 858–862 (2004).

Murray J, Tree AC. Prostate cancer – advantages and disadvantages of MR-guided RT. *Clin Transl Radiat Oncol*. **18**, 68–73 (2019).

Murray JR, Tree AC, Alexander EJ, *et al.* Standard and hypofractionated dose escalation to intraprostatic tumor nodules in localized prostate cancer: efficacy and toxicity in the DELINEATE trial. *Int J Radiat Oncol Biol Phys*. **106**, 715–724 (2020).

Nederlandse Kankerregistratie (NKR), IKNL. (2019).

Nicholls L, Suh YE, Chapman E, *et al.* Stereotactic radiotherapy with focal boost for intermediate and high-risk prostate cancer: initial results of the SPARC trial. *Clin Transl Radiat Oncol*. **25**, 88–93 (2020).

Nwankwo O, Mekdash H, Sihono DSK, Wenz F, Glatting G. Knowledge-based radiation therapy (KBRT) treatment planning versus planning by experts: validation of a KBRT algorithm for prostate cancer treatment planning. *Radiat Oncol*. **10**, 111 (2015).

Oldenhuis CNAM, Oosting SF, Gietema JA, de Vries EGE. Prognostic versus predictive value of biomarkers in oncology. *Eur J Cancer*. **44**, 946–953 (2008).

Olsson LE, Johansson M, Zackrisson B, Blomqvist LK. Basic concepts and applications of functional magnetic resonance imaging for radiotherapy of prostate cancer. *Phys Imaging Radiat Oncol*. **9**, 50–57 (2019).

Ou Y, Shen D, Zeng J, Sun L, Moul J and Davatzikos C. Sampling the spatial patterns of cancer: optimized biopsy procedures for estimating prostate cancer volume and Gleason Score. *Med Image Anal*. **13**, 609–620 (2009).

Owrangi AM, Greer PB, Glide-Hurst CK. MRI-only treatment planning: benefits and challenges. *Phys Med Biol*. **63**, 05TR01 (2018).

Ozer S, Langer DL, Liu X, *et al.* Supervised and unsupervised methods for prostate cancer segmentation with multispectral MRI. *Med Phys*. **37**, 1873–1883 (2010).

Padhani AR, Hayes C, Landau S and Leach MO. Reproducibility of quantitative dynamic MRI of normal human tissues. *NMR Biomed*. **15**, 143–153 (2002).

Padhani AR, Weinreb J, Rosenkrantz AB, Villeirs G, Turkbey B, Barentsz J. Prostate imaging-reporting and data systems steering committee: PI-RADS v2 status update and future directions. *Eur Urol*. **75**, 385–396 (2019).

Park SY, Kim CK, Park BK, *et al.* Early changes in apparent diffusion coefficient from diffusion-weighted MR imaging during radiotherapy for prostate cancer. *Int J Radiat Oncol Biol Phys*. **83**, 749–755 (2012).

Peeters ST, Heemsbergen WD, Koper PC, *et al.* Dose-response in radiotherapy for localized prostate cancer: results of the Dutch multicenter randomized phase III trial comparing 68 Gy of radiotherapy with 78 Gy. *J Clin Oncol*. **24**, 1990–1996 (2006).

Pickett B, Vigneault E, Kurhanewicz J, Verhey L, Roach M. Static field intensity modulation to treat a dominant intra-prostatic lesion to 90 Gy compared to seven field 3-dimensional radiotherapy. *Int J Radiat Oncol Biol Phys*. **44**, 921–929 (1999).

Pollack A, Zagars GK, Starkschall G, *et al.* Prostate cancer radiation dose response: results of the M. D. Anderson phase III randomized trial. *Int J Radiat Oncol Biol Phys*. **53**, 1097–1105 (2002).

Pucar D, Hricak H, Shukla-Dave A, *et al.* Clinically significant prostate cancer local recurrence after radiation therapy occurs at the site of primary tumor: magnetic resonance imaging and step-section pathology evidence. *Int J Radiat Oncol Biol Phys*. **69**, 62–69 (2007).

Raunig DL, McShane LM, Pennello G, *et al.* Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment. *Stat Methods Med Res*. **24**, 27–67 (2015).

Rawla P. Epidemiology of Prostate Cancer. *World J Oncol*. **10**, 63–89 (2019).

Rischke HC, Nestle U, Fechter T, *et al.* 3 Tesla multiparametric MRI for GTV-definition of dominant intraprostatic lesions in patients with prostate cancer--an interobserver variability study. *Radiat Oncol*. **8**, 183 (2013).

Ritter M. Rationale, conduct, and outcome using hypofractionated radiotherapy in prostate cancer. *Semin Radiat Oncol*. **18**, 249–256 (2008).

Roach M, Hanks G, Thames H Jr, *et al.* Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: recommendations of the RTOG-ASTRO Phoenix consensus conference. *Int J Radiat Oncol Biol Phys*. **65**, 965–974 (2006).

Rosenkrantz AB, Kim S, Lim RP, *et al.* Prostate cancer localization using multiparametric MR imaging: comparison of Prostate Imaging Reporting and Data System (PI-RADS) and Likert scales. *Radiology*. **269**, 482–492 (2013).

Rouvière O, Papillard M, Girouin N, *et al.* Is it possible to model the risk of malignancy of focal abnormalities found at prostate multiparametric MRI? *Eur Radiol*. **22**, 1149–1157 (2012).

Rudolph MM, Baur ADJ, Haas M, *et al.* Validation of the PI-RADS language: predictive values of PI-RADS lexicon descriptors for detection of prostate cancer. *Eur Radiol*. **30**, 4262–4271 (2020).

Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, Hawkes DJ. Non-rigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging*. **18**, 712–721 (1999).

Sanmamed N, Lee J, Berlin A, *et al.* Tumor-targeted dose escalation for localized prostate cancer using MR-guided HDR brachytherapy (HDR) or integrated VMAT (IB-VMAT) boost: dosimetry, toxicity and health related quality of life. *Radiother Oncol*. **149**, 240–245 (2020).

Sato C, Naganawa S, Nakamura T, *et al.* Differentiation of noncancerous tissue and cancer lesions by apparent diffusion coefficient values in transition and peripheral zones of the prostate. *J Magn Reson Imaging*. **21**, 258–262 (2005).

Schabel MC, Parker DL. Uncertainty and bias in contrast concentration measurements using spoiled gradient echo pulse sequences. *Phys Med Biol*. **53**, 2345–2373 (2008).

Scherrer B, Gholipour A, Warfield SK. Super-resolution in diffusion-weighted imaging. *Med Image Comput Comput Assist Interv*. **14**, 124–132 (2011).

van Schie MA, Dinh CV, van Houdt PJ, *et al.* Contouring of prostate tumors on multiparametric MRI: evaluation of clinical delineations in a multicenter radiotherapy trial. *Radiother Oncol*. **128**, 321–326 (2018).

Schreibmann E, Fox T. Prior-knowledge treatment planning for volumetric arc therapy using feature-based database mining. *J Appl Clin Med Phys*. **15**, 4596 (2014).

Shiraishi S, Moore KL. Knowledge-based prediction of three-dimensional dose distributions for external beam radiotherapy. *Med Phys*. **43**, 378 (2016).

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. **86**, 40–48 (1979).

Shukla-Dave A, Obuchowski NA, Chenevert TL, *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*. **49**, e101–e121 (2019).

Shusharina N, Craft D, Chen YL, Shih H, Bortfeld T. The clinical target distribution: a probabilistic alternative to the clinical target volume. *Phys Med Biol*. **63**, 155001 (2018).

Singh AK, Guion P, Sears-Crouse N, *et al.* Simultaneous integrated boost of biopsy proven, MRI defined dominant intra-prostatic lesions to 95 Gray with IMRT: early results of a phase I NCI study. *Radiat Oncol*. **2**, 36 (2007).

Smeenk RJ, van Lin ENJT, van Kollenburg P, McColl GM, Kunze-Busch M, Kaanders JHAM. Endorectal balloon reduces anorectal doses in post-prostatectomy intensity-modulated radiotherapy. *Radiother Oncol*. **101**, 465–470 (2011).

Song T, Li N, Zarepisheh M, *et al.* An automated treatment plan quality control tool for intensity-modulated radiation therapy using a voxel-weighting factor-based re-optimization algorithm. *PLoS ONE*. **11**, e0149273. (2016).

Steenbergen P, Haustermans K, Lerut E, *et al.* Prostate tumor delineation using multiparametric magnetic resonance imaging: inter-observer variability and pathology validation. *Radiother Oncol*. **115**, 186–190 (2015).

Teichert K, Currie G, Küfer KH, *et al.* Targeted multi-criteria optimisation in IMRT planning supplemented by knowledge based model creation. *Oper Res Health Care*. **23**, 100185 (2019).

Thorwarth D, Eschmann SM, Paulsen F, Alber M. Hypoxia dose painting by numbers: a planning study. *Int J Radiat Oncol Biol Phys*. **68**, 291–300 (2007).

Tofts PS. Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *J Magn Reson Imaging*. **7**, 91–101 (1997).

Tofts PS, Brix G, Buckley DL, *et al.* Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusable tracer: standardized quantities and symbols. *J Magn Reson Imaging*. **10**, 223–232 (1999).

Toivonen J, Merisaari H, Pesola M, *et al.* Mathematical models for diffusion-weighted imaging of prostate cancer using b values up to 2000 s/mm(2): correlation with Gleason score and repeatability of region of interest analysis. *Magnet Reson Med*. **74**, 1116–1124 (2015).

Trotti A, Colevas AD, Setser A, *et al.* CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. *Semin Radiat Oncol*. **13**, 176–181 (2003).

Vanderstraeten B, Duthoy W, De Gersem W, de Neve W, Thierens H. [18F]fluoro-deoxy-glucose positron emission tomography ([18F]FDG-PET) voxel intensity-based intensity-modulated radiation therapy (IMRT) for head and neck cancer. *Radiother Oncol*. **79**, 249–258 (2006).

Vanneste BGL, Hoffmann AL, van Lin EN, van de Voorde L, Pinkawa M, Lambin P. Who will benefit most from hydrogel rectum spacer implantation in prostate cancer radiotherapy? A model-based approach for patient selection. *Radiother Oncol*. **121**, 118–123 (2016).

Viani GA, Stefano EJ, Afonso SL. Higher-than-conventional radiation doses in localized prostate cancer treatment: a meta-analysis of randomized, controlled trials. *Int J Radiat Oncol Biol Phys*. **74**, 1405–1418 (2009).

de Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol*. **59**, 1033–1039 (2006).

Vigneault E, Mbodji K, Racine LG, *et al.* Image-guided high-dose-rate brachytherapy boost to the dominant intraprostatic lesion using multiparametric magnetic resonance imaging including spectroscopy: results of a prospective study. *Brachytherapy*. **15**, 746–751 (2016).

de Visschere PJL, Briganti A, Fütterer JJ, *et al.* Role of multiparametric magnetic resonance imaging in early detection of prostate cancer. *Insights Imaging*. **7**, 205–214 (2016).

Viswanath SE, Bloch NB, Chappelow JC, *et al.* Central gland and peripheral zone prostate tumors have significantly different quantitative imaging signatures on 3 Tesla endorectal, in vivo T2-weighted imagery. *J Magn Reson Imaging*. **36**, 213–224 (2012).

Vogelius IR, Bentzen SM. Meta-analysis of the alpha/beta ratio for prostate cancer in the presence of an overall time factor: bad news, good news, or no news? *Int J Radiat Oncol Biol Phys*. **85**, 89–94 (2013).

Vogelius IR, Bentzen SM. Dose response and fractionation sensitivity of prostate cancer after external beam radiation therapy: a meta-analysis of randomized trials. *Int J Radiat Oncol Biol Phys*. **100**, 858–865 (2018).

Vos PC, Hambrock T, Hulsbergen-van de Kaa CA, Fütterer JJ, Barentsz JO, Huisman HJ. Computerized analysis of prostate lesions in the peripheral zone using dynamic contrast enhanced MRI. *Med Phys*. **35**, 888–899 (2008).

Vos PC, Barentsz JO, Karssemeijer N, Huisman HJ. Automatic computer-aided detection of prostate cancer based on multiparametric magnetic resonance imaging analysis. *Phys Med Biol*. **57**, 1527–1542 (2012).

Wall PDH, Carver RL, Fontenot JD. An improved distance-to-dose correlation for predicting bladder and rectum dose-volumes in knowledge-based VMAT planning for prostate cancer. *Phys Med Biol*. **63**, 015035 (2018).

Wang Y, Zolnay A, Incrocci L, *et al.* A quality control model that uses PTV-rectal distances to predict the lowest achievable rectum dose, improves IMRT planning for patients with prostate cancer. *Radiother Oncol*. **107**, 352–357 (2013).

Wang Y, Heijmen BJM, Petit SF. Prospective clinical validation of independent DVH prediction for plan QA in automatic treatment planning for prostate cancer patients. *Radiother Oncol*. **125**, 500–506 (2017).

Weinreb JC, Barentsz JO, Choyke PL, *et al.* PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol*. **69**, 16–40 (2016).

Westphalen AC, Reed GD, Vinh PP, Sotto C, Vigneron DB, Kurhanewicz J. Multiparametric 3T endorectal MRI after external beam radiation therapy for prostate cancer. *J Magn Reson Imaging*. **36**, 430–437 (2012).

Widmark A, Gunnlaugsson A, Beckman L, *et al.* Extreme hypofractionation versus conventionally fractionated radiotherapy for intermediate risk prostate cancer: early toxicity results from the Scandinavian randomized phase III trial "HYPO-RT-PC". *Int J Radiat Oncol Biol Phys*. **96**, 938–939 (2016).

Widmark A, Gunnlaugsson A, Beckman L, *et al.* Ultra-hypofractionated versus conventionally fractionated radiotherapy for prostate cancer: 5-year outcomes of the HYPO-RT-PC randomised, non-inferiority, phase 3 trial. *Lancet*. **394**, 385–395 (2019).

van Wijk Y, Vanneste BGL, Walsh S, *et al.* Development of a virtual spacer to support the decision for the placement of an implantable rectum spacer for prostate cancer radiotherapy: comparison of dose, toxicity and cost-effectiveness. *Radiother Oncol*. **125**, 107–112 (2017).

Woo S, Suh CH, Kim SY, Cho JY, Kim SH. Diagnostic performance of Prostate Imaging Reporting and Data System Version 2 for detection of prostate cancer: a systematic review and diagnostic meta-analysis. *Eur Urol*. **72**, 177–188 (2017).

Wu B, Ricchetti F, Sanguineti G, *et al.* Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Med Phys*. **36**, 5497–5505 (2009).

Yan D, Chen S, Krauss DJ, Chen PY, Chinnaiyan P, Wilson GD. Tumor voxel dose-response matrix and dose prescription function derived using ^{18}F-FDG PET/CT images for adaptive dose painting by number. *Int J Radiat Oncol Biol Phys*. **104**, 207–218 (2019).

Yuan L, Ge Y, Lee WR, Yin FF, Kirkpatrick JP, Wu QJ. Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans. *Med Phys*. **39**, 6868–6878 (2012).

Zamboglou C, Eiber M, Fassbender TR, *et al.* Multimodal imaging for radiation therapy planning in patients with primary prostate cancer. *Phys Imaging Radiat Oncol*. **8**, 8–16 (2018).

Zhang L, Tang M, Chen S, Lei X, Zhang X, Huan Y. A meta-analysis of use of Prostate Imaging Reporting and Data System Version 2 (PI-RADS V2) with multiparametric MR imaging for the detection of prostate cancer. *Eur Radiol*. **27**, 5204–5214 (2017).

Summary

For patients with intermediate to high-risk prostate cancer, focal dose escalation to the intraprostatic lesion has been hypothesized to improve local control with no additional normal tissue complications as compared to conventional uniform dose prescription. Recently, the FLAME trial demonstrated that focal dose escalation significantly improves 5-year biochemical disease-free survival rates, without increased toxicities or reduced quality of life. Since a clinical benefit of focal dose-escalated treatment was found, a technical analysis of the FLAME dataset is useful for the further optimization of focal dose escalation strategies. In Chapter 2 and 3 of this thesis the FLAME dataset is analyzed and prediction models are applied to evaluate the soundness of tumor delineations and escalated dose levels. Chapter 4 and 5 of this thesis describe the applicability of multiparametric (mp-)MRI for dose painted treatment, and the value for early response-based adaptive treatment.

In **Chapter 2** the soundness of **clinical delineations** of the intraprostatic tumor in the FLAME trial is evaluated. It was found that besides interobserver variability, significant institutional bias exists regarding the weighting of T2-weighted, DWI and DCE-MRI sequences to delineate the intraprostatic lesion. It is therefore recommended to involve weighting of MRI sequences in the development of delineation guidelines for intraprostatic lesions on mp-MRI. In addition, a tumor prediction model was shown to be capable of identifying manual delineations that needed to be corrected. This demonstrates the added benefit of prediction models in the clinical workflow.

The realized level of dose escalation during treatment planning depends to a large extent on the anatomy of the patient and the complexity of the dose optimization process. In **Chapter 3** dose volume histogram (DVH) parameters of the **planned dose distributions** of the FLAME trial are reported, showing how well a focal dose escalation up to 95 Gy could be planned. In the dose-escalated arm of the trial, a median $D_{98\%}$ of 84.7 Gy was observed, and 99% of the patients received a significant dose escalation above 82.4 Gy. During treatment planning, an unfavorable tumor location with respect to healthy surrounding organs results in a lower than intended planned dose escalation. To assist during the treatment planning procedure, a regression model was developed that estimates the achievable dose escalation to the tumor as a function of patient anatomy. To validate the model, a random subset of dose-escalated treatment plans was replanned. For this subset a strong correlation ($r = 0.89$) was found between predicted increase in tumor $D_{98\%}$ and replanned tumor $D_{98\%}$. It was concluded that the model may assist future clinical dose escalated treatment planning to reach the highest achievable dose to the tumor.

With mp-MRI both anatomical and functional information is acquired. For this reason, an mp-MRI examination is recommended for identifying the intraprostatic lesion and for defining the target for focal dose escalation. Currently, most research on mp-MRI focusses on improved identification of the intraprostatic lesion. With the positive outcome of the FLAME trial, new applications of mp-MRI can be considered for further development of focal dose escalation for prostate cancer.

While in the FLAME trial a uniform escalated dose was prescribed to the delineated intraprostatic lesion, in reality the tumor can be spatially heterogeneous and the boundary is not discrete. Circumventing the binary process of manual contouring and to reflect the biological characteristics of the tumor, dose painting by numbers (DPbN) prescribes a heterogeneous dose distribution to the prostate gland and is directly related to imaging features. The omission of manual interaction however potentially allows measurement uncertainties in image data to impact the planned dose distribution. In **Chapter 4**, a repeatability analysis is performed to quantify the applicability of **dose painting by numbers**. The intra-class correlation (ICC) coefficients used as repeatability measure, improved from 0.84 to 0.93 from acquired imaging data towards planned dose distributions. It was concluded that variation in imaging data had a minimal impact on the realized planned dose distributions, which suggests that DPbN treatment planning is a realistic alternative to contour-based dose painting.

While mp-MRI has proven its use for focal dose escalated treatment and was shown suitable for DPbN treatment planning as well, during the delivery of the treatment, mp-MRI may also record radiation response of the tumor. In **Chapter 5** the value of **quantitative MRI biomarkers during radiotherapy** for the purpose of image-based treatment adaptation was investigated. T2 and ADC maps of six MRI examinations during hypofractionated treatment were analyzed. No early treatment-induced changes to the tumor on a population basis were found. It is therefore unlikely that in the near future actionable imaging biomarkers will be discovered for early-treatment adaptations in hypofractionated prostate radiotherapy.

In this thesis a technical evaluation of focal dose escalation of prostate cancer was performed and the role of mp-MRI for future optimization of this treatment strategy was investigated. The technical evaluation demonstrates the need for MRI-based delineation guidelines of the intraprostatic tumor, and that models could contribute to more consistent tumor delineations and optimization of the escalated dose to the tumor. The presented findings with respect to future optimization of this treatment strategy show that mp-MRI is applicable in the planning stage of the treatment for the purpose of dose painting, and leave little room for early adaptive treatment based on mp-MRI.

Samenvatting

Voor patiënten met matig- tot hoog-risico prostaatkanker wordt het integreren van een lokale dosisverhoging (bestralingsboost) op de tumor in de prostaat gezien als geschikte bestralingsstrategie om zonder extra bijwerkingen van het gezonde weefsel een verbeterde lokale controle te behalen. Onlangs heeft de FLAME studie aangetoond dat een lokale dosisverhoging een significante verbetering op de 5-jaars biochemisch recidief-vrije overleving geeft, zonder toename in toxiciteit en vermindering van kwaliteit van leven. Omdat met deze behandeling de klinische uitkomst van de patiënt verbetert, is een technische analyse van de FLAME dataset nuttig voor de verdere optimalisering van bestralingsbooststrategieën. In hoofdstuk 2 en 3 van dit proefschrift is de FLAME dataset onderzocht en worden voorspellingsmodellen toegepast om de betrouwbaarheid van klinische tumor intekeningen en dosis van de bestralingsboost te beoordelen. Hoofdstuk 4 en 5 van dit proefschrift beschrijven de toepassing van multiparametrische (mp-)MRI voor een 'dose painting' behandeling, en de rol die het kan spelen voor vroegtijdige aanpassing van de behandeling op basis van reacties van het weefsel.

In **hoofdstuk 2** wordt de betrouwbaarheid van **klinische intekeningen** van de prostaattumor in de FLAME studie beoordeeld. Het blijkt dat naast individuele intekenvariatie er op afdelingsniveau een significante trend bestaat met betrekking tot de voorkeur voor T2-gewogen, DWI en DCE-MRI sequenties ten behoeve van het intekenen van de prostaattumor. Het is daarom aanbevolen om de weging van de verschillende sequenties mee te nemen in het opstellen van intekenrichtlijnen voor de prostaattumor op basis van mp-MRI. Daarnaast bleek een tumor voorspellingsmodel in staat om handmatige intekeningen te herkennen die aangepast moesten worden. Dit resultaat laat de toegevoegde waarde van voorspellingsmodellen in de kliniek zien.

Tijdens het opstellen van het behandelplan hangt de hoogte van de lokale dosisverhoging voor een groot deel af van de anatomie van de patiënt en de complexiteit van het optimalisatieproces. In **hoofdstuk 3** worden dosis volume histogram (DVH) parameters van **geplande dosisverdelingen** uit de FLAME studie vermeld, die aangeven in hoeverre een focale bestralingboost tot aan 95 Gy is gepland op de tumor in de prostaat. In de bestralingsboostarm van de studie werd een mediane $D_{98\%}$ gevonden van 84.7 Gy, waarbij 99% van de patiënten een significante dosisverhoging boven de 82.4 Gy kregen. Tijdens het opstellen van het bestralingsplan kan een ongunstige tumorlocatie ten opzichte van gezonde organen leiden tot een bestralingsboost dosis die lager uitvalt dan gewenst. Om deze oorzaak aan te pakken, is een regressiemodel ontwikkeld dat de haalbare dosisverhoging in de prostaattumor als functie van de anatomie van de patiënt kan inschatten. Om het model te valideren is een willekeurige

selectie van bestralingplannen met een focale bestralingsboost opnieuw gepland. Voor deze bestralingsplannen werd een sterke correlatie (r = 0.89) gevonden tussen de voorspelde toename in tumor $D_{98\%}$ en de gerealiseerde toename tijdens de herplanning. Deze resultaten tonen aan dat het model een toekomstige klinische toepassing kan vinden in de behandeling van prostaatkanker met een focale bestralingsboost.

Met mp-MRI wordt zowel anatomische als functionele informatie verkregen. Om deze reden is een mp-MRI onderzoek aangeraden voor het identificeren van de tumor in de prostaat en het definiëren van het doelvolume waarop een lokale dosisverhoging wordt gegeven. Op dit moment richt onderzoek met mp-MRI zich vooral op een verbeterde herkenning van de tumor in de prostaat. Met de positieve uitkomst van de FLAME studie kunnen ook andere toepassingen van mp-MRI worden onderzocht voor verdere ontwikkeling van prostaatbestraling met een lokale dosisverhoging.

Hoewel in de FLAME studie een uniforme dosisverhoging is voorgeschreven op de ingetekende prostaattumor, kan de tumor in werkelijkheid heterogeen en niet scherp begrensd zijn. Om de binaire beslissing van handmatig intekenen te omzeilen en om beter aan te sluiten bij de biologische eigenschappen, schrijft dose painting by numbers (DPbN) een onregelmatige dosisverdeling op de prostaat voor die direct gerelateerd is aan beeldkarakteristieken. Het weglaten van de handmatige intekening laat echter toe dat onzekerheden in de beelden de geplande dosisverdeling beïnvloeden. In **hoofdstuk 4** is een herhaalbaarheidsstudie uitgevoerd om de toepassing van **dose painting by numbers** te kwantificeren. De intra-class correlation (ICC) coëfficiënt als maat voor herhaalbaarheid, verbeterde van 0.84 naar 0.93 voor verkregen MRI beelden en geplande dosisverdelingen. Hieruit volgt dat variatie in beelden een minimale invloed op de gerealiseerde dosisverdeling heeft, wat DPbN aandraagt als realistisch alternatief voor contour-gebaseerde lokale dosisverhoging.

Terwijl mp-MRI zijn gebruik voor bestraling met een lokale dosisverhoging heeft bewezen en tevens geschikt blijkt voor DPbN behandeling, zou mp-MRI ook toepassing kunnen vinden tijdens de bestraling, door het in beeld brengen van respons op de radiotherapie in de tumor. In **hoofdstuk 5** is de waarde van **kwantitatieve MRI biomarkers** voor het aanpassen van de behandeling op basis van beeldinformatie onderzocht. T2 en ADC beelden van zes MRI onderzoeken tijdens een gehypofractioneerde behandeling werden geanalyseerd. Op populatieniveau werden geen veranderingen van de tumor gevonden als gevolg van de behandeling. Het is daarom onwaarschijnlijk dat in de nabije toekomst MRI biomarkers worden ontdekt die vroege aanpassing aan de gehypofractioneerde bestralingsdosis van prostaatkanker mogelijk maken.

In dit proefschrift is een technische evaluatie van de bestraling van prostaatkanker met een mp-MRI gebaseerde focale bestralingsboost uitgevoerd en is de rol van mp-MRI voor toekomstige optimalisatie van deze behandelstrategie onderzocht. De technische evaluatie toont aan dat er behoefte is aan intekenrichtlijnen voor de tumor in de prostaat op basis van mp-MRI en dat modellen kunnen bijdragen aan een verhoogde consistentie van tumor intekeningen en optimalisatie van de focale bestralingsboost. De gepresenteerde bevindingen omtrent toekomstige optimalisatie van deze behandelstrategie laten zien dat mp-MRI goed toepasbaar is in de planningsfase van de behandeling, zoals voor 'dose painting' doeleinden, en geven weinig aanleiding tot gebruik van mp-MRI voor het tussentijds aanpassen van de behandeling.

List of Publications

Publications in this thesis

Repeatability of dose painting by numbers treatment planning in prostate cancer radiotherapy based on multiparametric magnetic resonance imaging
M.A. van Schie, P. Steenbergen P, C.V. Dinh, G. Ghobadi, P.J. van Houdt, F.J. Pos, S.W.T.J.P. Heijmink, H.G. van der Poel, S. Renisch, T. Vik and U.A. van der Heide
Physics in Medicine & Biology, **62**(14), 5575–5588 (2017)
DOI: 10.1088/1361-6560/aa75b8

Contouring of prostate tumors on multiparametric MRI: evaluation of clinical delineations in a multicenter radiotherapy trial
M.A. van Schie, C.V. Dinh, P.J. van Houdt, F.J. Pos, S.W.T.J.P. Heijmink, L.G.W. Kerkmeijer, A.N.T.J. Kotte, R. Oyen, K. Haustermans and U.A. van der Heide
Radiotherapy & Oncology, **128**(2), 321–326 (2018)
DOI: 10.1016/j.radonc.2018.04.015

Quantitative MRI changes during weekly ultra-hypofractionated prostate cancer radiotherapy with integrated boost
M.A. van Schie, P.J. van Houdt, G. Ghobadi, F.J. Pos, I. Walraven, H.C.J. de Boer, C.A.T. van den Berg, R.J. Smeenk, L.G.W. Kerkmeijer and U.A. van der Heide
Frontiers in Oncology, **4**(9), 1264 (2019)
DOI: 10.3389/fonc.2019.01264

Knowledge-based assessment of focal dose escalation treatment plans in prostate cancer
M.A. van Schie, T.M. Janssen, D. Eekhout, I. Walraven, F.J. Pos, P. de Ruiter, A.N.T.J. Kotte, E.M. Monninkhof, L.G.W. Kerkmeijer, C. Draulans, R. de Roover, K. Haustermans, M. Kunze-Busch, R.J. Smeenk and U.A. van der Heide
International Journal of Radiation Oncology • Biology • Physics, **108**(4), 1055–1062 (2020)
DOI: 10.1016/j.ijrobp.2020.06.072

Other publications

Voxelized ray-tracing simulation dedicated to multi-pinhole molecular breast tomosynthesis
B. Wang, J. Roosmalen, L. Piët, **M.A. van Schie**, F.J. Beekman and M.C. Goorden
Biomedical Physics & Engineering Express, **3**(4), 045021 (2017)
DOI: 10.1088/2057-1976/aa8012

Quantitative imaging for radiotherapy purposes
O.J. Gurney-Champion, F. Mahmood, **M.A. van Schie**, R. Julian, B. George, M.E.P. Philippens, U.A. van der Heide, D. Thorwarth and K.R. Redalen
Radiotherapy & Oncology, **146**, 66–75 (2020)
DOI: 10.1016/j.radonc.2020.01.026

Conference presentations

Repeatability of dose painting by numbers in prostate cancer radiotherapy based on multiparametric MRI
M.A. van Schie*, P. Steenbergen, C.V. Dinh, G. Ghobadi, P.J. van Houdt, F.J. Pos, S.W.T.J.P. Heijmink, H.G. van der Poel, S. Renisch, T. Vik and U.A. van der Heide
5th MR in RT symposium, Sydney, Australia (2017)
*Oral presentation

Evaluation of prostate tumor delineations on multiparametric MRI in a multicenter radiotherapy trial
M.A. van Schie*, C.V. Dinh, P.J. van Houdt, F.J. Pos, S.W.T.J.P. Heijmink, L.G.W. Kerkmeijer, A.N.T.J. Kotte, R. Oyen, K. Haustermans and U.A. van der Heide
ESTRO 37, Barcelona, Spain (2018)
*Oral presentation

Prostate tumor prediction model for automated delineation QA on multiparametric MRI
M.A. van Schie*, C.V. Dinh, P.J. van Houdt, F.J. Pos, S.W.T.J.P. Heijmink, L.G.W. Kerkmeijer, A.N.T.J. Kotte, R. Oyen, K. Haustermans and U.A. van der Heide
6th MR in RT Symposium, Utrecht, The Netherlands (2018)
*Oral presentation

Automated treatment plan evaluation for focal dose escalation in prostate cancer radiotherapy
M.A. van Schie*, T.M. Janssen, D. Eekhout, I. Walraven, A.N.T.J. Kotte, L.G.W. Kerkmeijer, C. Draulans, K. Haustermans, R.J. Smeenk and U.A. van der Heide
Institute QuantiVision Conference, Amsterdam, The Netherlands (2019)
*Poster presentation

Treatment plan quality assessment for focal dose escalation in prostate cancer
M.A. van Schie*, T.M. Janssen, D. Eekhout, I. Walraven, A.N.T.J. Kotte, L.G.W. Kerkmeijer, C. Draulans, K. Haustermans, R.J. Smeenk and U.A. van der Heide
ESTRO 38, Milan, Italy (2019)
*Poster presentation

Dankwoord

Mijn dank gaat uit naar alle collega's en naasten die in welke vorm dan ook betrokken zijn geweest bij mijn promotieonderzoek en de voltooiing van dit proefschrift. Een aantal personen wil ik hiervoor met name bedanken.

- Uulke, onze overleggen zorgde er altijd voor dat ik weer met frisse moed aan de slag ging. Ik wil je zeker ook bedanken voor je inzet in de slotfase om mij te motiveren de onderzoeksresultaten tot een mooi geheel te bundelen.
- Linda, ik heb veel aan je kennis en kunde van het vakgebied gehad en waardeer de kritische blik waarmee de presentatie van resultaten nog beter tot zijn recht kwam.
- Floris, dank voor het inzichtelijk maken van de ontwikkelingen in het veld en de manier waarop je me liet nadenken over de implicaties van mijn onderzoek.
- Roelant, ongeveer gelijktijdig begonnen heb ik in die vier jaar de vele zowel inhoudelijke als 'off-topic' gesprekken als erg prettig ervaren. Je nuchtere insteek kon vaak heel verhelderend zijn. Bedankt dat je mijn paranimf wilt zijn.
- Ivo, al op de middelbare school gingen wij samen met Thomas onafscheidelijk door het leven als triumviraat. Mede door de gedeelde interesses voelt onze vriendschap als vanzelfsprekend. Bedankt dat je mijn paranimf wilt zijn.
- Cuong, thank you for giving me a swift start of my PhD trajectory and the confidence of finishing without your supervision. I appreciate your willingness to chat about work related issues as well as personal developments outside the walls of the NKI.
- Ghazaleh en Petra, bij jullie op de kamer heerste een prettige werksfeer met de juiste balans tussen focus en gezelligheid. Daarnaast waren jullie altijd bereid om mee te denken over mijn onderzoek, waarvoor ik jullie wil bedanken.
- Collega's van de afdeling radiotherapie, in het bijzonder de Uulke Groep, bedankt voor de vele goede herinneringen aan de jaren dat ik met jullie heb samengewerkt, naar congressen ben geweest en feesten heb gevierd.
- Familie en vrienden, ik wil jullie bedanken voor de oprechte interesse die jullie toonden in mijn onderzoek en de vele gelegenheden die er waren om maar niet altijd met mijn onderzoek bezig te zijn.
- Marion, ik heb grote waardering voor de manier waarop je begripvol en geduldig bent. Vanuit jouw perspectief wist je vaak serieuze inzichten te geven wat mijn onderzoek alleen maar ten goede kwam. Bedankt voor alle mooie momenten die mij positieve energie gaven om dit proefschrift te voltooien!

Curriculum Vitae

Marcel van Schie werd op 14 mei 1991 geboren in Singapore. Hij groeide op in Leidschendam, alwaar hij in 2009 zijn VWO diploma behaalde. Datzelfde jaar begon hij aan de bachelor Technische Natuurkunde met aansluitend de master Applied Physics aan de TU Delft. Tijdens de master volgde hij de specialisatie Nuclear Science and Engineering, waarbij het afstudeerproject aan het Reactor Instituut Delft zich focuste op de invloed van geometrische onnauwkeurigheden van een multi-pinhole micro-SPECT systeem op de kwaliteit van de gereconstrueerde beelden. Het afsluitende stageproject van de masteropleiding vond plaats bij de radiotherapie-afdeling van het MCH Westeinde in Den Haag, waar een ketentest werd opgezet voor de bestraling van prostaatkanker met HDR brachytherapie. De daar ontstane interesse voor het vakgebied radiotherapie leidde in 2015 tot het uitvoeren van een promotieonderzoek aan het NKI op de afdeling radiotherapie. Het daaruit voortvloeiende proefschrift dat in 2021 is voltooid, beschrijft de inzet van multiparametrische MRI voor de bestraling van prostaatkanker met een geïntegreerde focale bestralingsboost op de zichtbare tumor. Onderwijl is hij in 2020 als project-fysicus op de afdeling radiotherapie van het HMC in Leidschendam medeverantwoordelijk geweest voor de invoering van een nieuw treatment planning system, waarna hij in januari 2021 op dezelfde afdeling is begonnen aan de opleiding tot Klinisch Fysicus in de radiotherapie.